POST-GRADUATE DEGREE PROGRAMME (CBCS)

M.SC. IN MATHEMATICS

SEMESTER-III

PAPER : DSE 3.3
(Pure Stream)

Operator Theory

Measure Theory

Self-Learning Material

DIRECTORATE OF OPEN AND DISTANCE LEARNING
UNIVERSITY OF KALYANI
Kalyani, Nadia
West Bengal

Course Content Writers	
Course Name	**Writer**
Block - I : Operator Theory Units 1-5	Dr. Pulak Sahoo Department of Mathematics University of Kalyani
Block - I : Operator Theory Units 6-10	Dr. Sanjib Kumar Datta Department of Mathematics University of Kalyani
Block - II : Measure Theory	Ms. Audrija Choudhury Assistant Professor Department of Mathematics Directorate of Open and Distance Learning University of Kalyani

August, 2023

Directorate of Open and Distance Learning, University of Kalyani
Published by the Directorate of Open and Distance Learning
University of Kalyani, 741235, West Bengal
Printed by East India Photo Composing Centre, 209A, Bidhan Sarani, Kolkata-700006

Director's Message

Satisfying the varied needs of distance learners, overcoming the obstacle of Distance and reaching the unreached students are the three fold functions catered by Open and Distance Learning (ODL) systems. The onus lies on writers, editors, production professionals and other personnel involved in the process to overcome the challenges inherent to curriculum design and production of relevant Self Learning Materials (SLMs). At the University of Kalyani a dedicated team under the able guidance of the Hon'ble Vice-Chancellorhas invested its best efforts, professionally and in keeping with the demands of Post Graduate CBCS Programmes in Distance Mode to devise a self-sufficient curriculum for each course offered by the Directorate of Open and Distance Learning (DODL), University of Kalyani.

Development of printed SLMs for students admitted to the DODL within a limited time to cater to the academic requirements of the Course as per standards set by Distance Education Bureau of the University Grants Commission, New Delhi, India under Open and Distance Mode UGC Regulations, 2020 had been our endeavor. We are happy to have achieved our goal.

Utmost care and precision have been ensured in the development of the SLMs, making them useful to the learners, besides avoiding errors as far as practicable. Further suggestions from the stakeholders in this would be welcome.

During the production-process of the SLMs, the team continuously received positive stimulations and feedback from Professor **(Dr.) Amalendu Bhunia, Hon'ble Vice-Chancellor, University of Kalyani**, who kindly accorded directions, encouragements and suggestions, offered constructive criticism to develop it with in proper requirements. We gracefully, acknowledge his inspiration and guidance.

Sincere gratitude is due to the respective chairpersons as well as each and every member of PGBOS (DODL), University of Kalyani. Heartfelt thanks are also due to the Course Writers-faculty members at the DODL, subject-experts serving at University Post Graduate departments and also to the authors and academicians whose academic contributions have enriched the SLMs. We humbly acknowledge their valuable academic contributions. I wouldespecially like to convey gratitude to all other University dignitaries and personnel involved either at the conceptual or operational level of the DODL of University of Kalyani.

Their persistent and coordinated efforts have resulted in the compilation of comprehensive, learner-friendly, flexible texts that meet the curriculum requirements of the Post Graduate Programme through Distance Mode.

All the Self Learning Materials are self writing and collected from e-book, journals and websites.

Professor (Dr.) Sanjib Kumar Datta
Director
Directorate of Open and Distance Learning
University of Kalyani

Post Graduate Board of Studies (PGBOS) Members of Department of Mathematics, Directorate of Open and Distance Learning (DODL), University of Kalyani.

Sl. No.	Name & Designation	Role
1	Prof. (Dr.) Samares Pal, Professor & Head, Department of Mathematics, University of Kalyani.	Chairperson
2	Prof. (Dr.) Pulak Sahoo, Professor, Department of Mathematics, University of Kalyani.	Member
3	Dr. Sahidul Islam, Associate Professor, Department of Mathematics, University of Kalyani.	Member
4	Prof. (Dr.) Sushanta Kumar Mohanta, Professor, Department of Mathematics, West Bengal State University.	External Nominated Member
5	Ms. Audrija Choudhury, Assistant Professor, Department of Mathematics, DODL, University of Kalyani.	Member
6	Prof. (Dr.) Sanjib Kumar Datta, Director, DODL, University of Kalyani.	Convener

Discipline Specific Elective Paper

PURE STREAM

DSE 3.3
Marks : 100 (SEE : 80; IA : 20); Credit : 6

Operator Theory (Marks : 50 (SEE: 40; IA: 10))
Measure Theory (Marks : 50 (SEE: 40; IA: 10))

Syllabus

Block I

- **Unit 1:** Conjugate Space: Definition of conjugate space, determination of conjugate spaces of $\mathbb{R}^n$, l_p for $1 \leq p < \infty$. Representation theorem for bounded linear functionals on $C[a, b]$ (Statement only). Some idea about the spaces $BV[a, b]$ and $B[a, b]$. Determination of conjugate spaces of $C[a, b]$ and some other finite and infinite dimensional spaces.
- **Unit 2:** Weak convergence and weak* convergence: Definition, characterization of weak convergence and weak* convergence, sufficient condition for the equivalence of weak* convergence and weak convergence in the dual space.
- **Unit 3:** Reflexive spaces: Definition of reflexive space, canonical mapping, relation between reflexivity and separability, some consequences of reflexivity.
- **Unit 4:** Bounded linear operator, uniqueness theorem, adjoint of an operator and its properties.
- **Unit 5:** Self-adjoint, compact, normal, unitary and positive operators, norm of self -adjoint operator, group of unitary operator, square root of positive operator-characterization and basic properties
- **Unit 6:** Projection operator and their sum, product & permutability, invariant subspaces, closed linear transformation, closed graph theorem and open mapping theorem.
- **Unit 7:** Unbounded operator: Basic properties, Cayley transform, change of measure principle, spectral theorem.
- **Unit 8:** Compact map: Basic properties, compact symmetric operator, Rayleigh principle, Fisher's principle, Courant's principle, Mercer's theorem, positive compact operator.
- **Unit 9:** Strongly continuous semigroup: Strongly continuous semigroup of operator and contraction, infinitesimal generator
- **Unit 10:** Hille-Yosida theorem, Lumer-Phillips lemma, Trotter's theorem, Stone's theorem.

Block II

- **Unit 11:** Measures: Class of Sets, Measures, The extension Theorems.
- **Unit 12:** Caratheodory extension of measure, Completeness of measure; Lebesgue-Stieltjes measures
- **Unit 13:** Integration: Measurable transformations, Induced measures, distribution functions, Integration.
- **Unit 14:** More on Convergence
- **Unit 15:** Lp-spaces: Lp-Spaces, Dual spaces, Banach and Hilbert spaces.
- **Unit 16:** Product of two measure spaces, Fubini's theorem
- **Unit 17:** Decomposition and Differentiations: The Lebesgue-Radon-Nikodym theorem
- **Unit 18:** Signed and Complex Measures
- **Unit 19:** Differentiation on absolute Continuity, Lebesgue differentiation Theorem,
- **Unit 20:** Functions of Bounded variations, Riesz representation Theorem.

Contents

Unit 1

Conjugate Space

Course Structure

- Conjugate Space: Definition of conjugate space, determination of conjugate spaces of $\mathbb{R}^n$, l_p for $1 \leq p < \infty$. Representation theorem for bounded linear functionals on $C[a,b]$ (Statement only). Some idea about the spaces $BV[a,b]$ and $B[a,b]$. Determination of conjugate spaces of $C[a,b]$ and some other finite and infinite dimensional spaces.

1.1 Introduction

Suppose that X and Y be two normed linear spaces over the same scalar field $K(= \mathbb{R}$ or $\mathbb{C})$. Then, the collection $B(X,Y)$ of all bounded linear operators $T : X \longrightarrow Y$ is a linear space under the operations addition and scalar multiplication defined as follows:

$$\begin{aligned} & (T_1 + T_2)(x) = T_1(x) + T_2(x) \\ \text{and} \quad & (\lambda T)(x) = \lambda T(x), \ \forall x \in X, \lambda \in K. \end{aligned}$$

The zero element in this linear space is the operator 0 such that

$$0x = 0, \ \forall x \in X.$$

It can be shown that $B(X,Y)$ is a normed linear space, where for every $T \in B(X,Y)$,

$$\|T\| = \sup\{\|Tx\| : \|x\| \leq 1\}.$$

Further, if Y is a Banach space, then $B(X,Y)$ is also a Banach space.

Definition 1.1.1. Let X be a normed linear space over the scalar field $K(= \mathbb{R}$ or $\mathbb{C})$. Then, the space $B(X,K)$ of all the bounded linear functionals defined on X is called the conjugate space or the dual space of X and is denoted by X^*. Since K is a Banach space under the absolute value norm, it follows that $X^* = B(X,K)$ is a Banach space.

Definition 1.1.2. Let (X, d) and (Y, ρ) be two metric spaces. If there exist a one-one correspondence between the elements of X and Y such that the distance between any two elements of X is the same as the distance between the corresponding elements of Y, then the mapping is called an isometry. In this case, the spaces X and Y are called isometric spaces.

Definition 1.1.3. Let X and Y be two normed linear spaces over the same scalar field K. Let $T : X \longrightarrow Y$ be a linear operator. T is called an isometric isomorphism between X and Y if T is bijective and preserves norm, i.e., $\|Tx\| = \|x\|, \forall x \in X$. In this case, the spaces X and Y are called isometrically isomorphic.

If a normed linear space X is isometrically isomorphic to a normed linear space Y, then from the stand point of functional analysis, the spaces X and Y are identical.

Theorem 1.1.4. The conjugate space of $\mathbb{R}_n$ is $\mathbb{R}_n$.

Proof. We know that $\mathbb{R}_n$ is the collection of all n-tuples of real numbers. Let $e_1 = (1, 0, \dots, 0)$, $e_2 = (0, 1, 0, \dots, 0), \cdots, e_n = (0, 0, \dots, 0, 1)$ be a basis of $\mathbb{R}_n$.

Then, any element $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}_n$ can be written as

$$x = \sum_{k=1}^{n} x_k e_k.$$

Let $f \in \mathbb{R}_n^*$. Then

$$f(x) = \sum_{k=1}^{n} x_k f(e_k) = \sum_{k=1}^{n} x_k \alpha_k$$

where $\alpha_k = f(e_k), k = 1, 2, \dots, n$.

Thus, for each $f \in \mathbb{R}_n^*$, we obtain an element $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n) \in \mathbb{R}_n$. This defines an operator

$$\begin{aligned} T : \mathbb{R}_n^* &\longrightarrow \mathbb{R}_n \\ \text{given by } T(f) &= \lambda, f \in \mathbb{R}_n^* \text{ and } \lambda = (\lambda_1, \lambda_2, \dots, \lambda_n) \in \mathbb{R}_n, \end{aligned}$$

provided for any $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}_n$, $f(x) = \sum_{k=1}^{n} x_k \lambda_k$.

We now show that T is a bijective linear operator which preserves norm.

We first show that T is linear. Let $f, g \in \mathbb{R}_n^*$ and α be a scalar. Let

$$\begin{aligned} & T(f) = b = (b_1, b_2, \dots, b_n) \\ \text{and} \quad & T(g) = c = (c_1, c_2, \dots, c_n). \end{aligned}$$

If $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}_n$, then

$$f(x) = \sum_{k=1}^{n} x_k b_k \quad \text{and} \quad g(x) = \sum_{k=1}^{n} x_k c_k.$$

Now

$$\begin{aligned} (f + g)(x) &= f(x) + g(x) \\ &= \sum_{k=1}^{n} x_k b_k + \sum_{k=1}^{n} x_k c_k \\ &= \sum_{k=1}^{n} x_k (b_k + c_k) \end{aligned}$$

and

$$
\begin{aligned}
(\alpha f)(x) &= \alpha f(x) \\
&= \alpha \sum_{k=1}^{n} x_k b_k \\
&= \sum_{k=1}^{n} x_k(\alpha b_k).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
& T(f+g) = b + c = T(f) + T(g) \\
\text{and} \qquad & T(\alpha f) = \alpha b = \alpha T(f).
\end{aligned}
$$

This shows that T is linear.

Next, we show that T is injective. Let $f, g \in \mathbb{R}_n^*$ and $T(f) = T(g) = c = (c_1, c_2, \dots, c_n)$.

If $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}_n$, then

$$
f(x) = \sum_{k=1}^{n} x_k c_k \quad \text{and} \quad g(x) = \sum_{k=1}^{n} x_k c_k
$$

that is $f(x) = g(x)$. This shows that $f = g$ and hence T is injective.

We now show that T is surjective.

Let $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n) \in \mathbb{R}_n$. We define a functional f on $\mathbb{R}_n$ as follows:

If $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}_n$, then

$$
f(x) = \sum_{k=1}^{n} x_k \lambda_k.
$$

It can be easily shown that f is linear. By Cauchy Schwarz inequality

$$
\begin{aligned}
|f(x)| = |\sum_{k=1}^{n} x_k \lambda_k| &\leq \sum_{k=1}^{n} |x_k \lambda_k| \\
&\leq \left(\sum_{k=1}^{n} |x_k|^2 \right)^{\frac{1}{2}} \left(\sum_{k=1}^{n} |\lambda_k|^2 \right)^{\frac{1}{2}} \\
&= \|x\| \left(\sum_{k=1}^{n} |\lambda_k|^2 \right)^{\frac{1}{2}}. \qquad (1.1.1)
\end{aligned}
$$

This means that f is bounded and hence $f \in \mathbb{R}_n^*$. It is clear from the definition of f that $T(f) = \lambda$. Hence, T is surjective.

Next we show that T preserves norm. Let $f \in \mathbb{R}_n^*$ and $T(f) = \lambda$, where $\lambda := (\lambda_1, \lambda_2, \dots, \lambda_n) \in \mathbb{R}_n$. Then, $f(x) = \sum_{k=1}^{n} x_k \lambda_k$ for $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}_n$.

Proceeding similarly as (1.1.1), we obtain

$$\begin{aligned} |f(x)| &\leq \|x\| \left(\sum_{k=1}^{n} |\lambda_k|^2 \right)^{\frac{1}{2}} \\ \text{i.e., } \frac{|f(x)|}{\|x\|} &= \left(\sum_{k=1}^{n} |\lambda_k|^2 \right)^{\frac{1}{2}} = \|\lambda\|. \end{aligned}$$

Therefore,

$$\|f\| = \sup \left\{ \frac{|f(x)|}{\|x\|} : \|x\| \neq 0 \right\} \leq \|\lambda\|. \tag{1.1.2}$$

Choosing $x = \lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$ we obtain

$$\|f\| \geq \frac{|f(\lambda)|}{\|\lambda\|} = \left(\sum_{k=1}^{n} |\lambda_k|^2 \right)^{\frac{1}{2}} = \|\lambda\|. \tag{1.1.3}$$

Combining (1.1.2) and (1.1.3) we get

$$\|f\| = \|\lambda\| = \|T(f)\|.$$

Hence, T preserves norm. Thus, $T : \mathbb{R}_n^* \longrightarrow \mathbb{R}_n$ is a bijective linear operator which preserves norm. T is therefore an isometric isomorphism of $\mathbb{R}_n^*$ onto $\mathbb{R}_n$. Hence the conjugate space of $\mathbb{R}_n$ is $\mathbb{R}_n$. This proves the theorem. □

Note 1.1.5. The linear space $\mathbb{C}^n$ equipped with the norm given by

$$\|x\|_2 = \left(\sum_{i=1}^{n} |x_i|^2 \right)^{\frac{1}{2}}, \ x = (x_1, x_2, \ldots, x_n) \in \mathbb{C}^n$$

is a complex Banach space. The space $\mathbb{C}^n$ is called the n-dimensional unitary space.

Note 1.1.6. The linear space $K^n(\mathbb{R}_n \text{ or } \mathbb{C}_n)$ is a Banach space with each of the norms

$$\begin{aligned} \|x\|_1 &= \sum_{i=1}^{n} |x_i| \\ \text{and } \|x\|_\infty &= \max\{|x_i| : 1 \leq i \leq n\} \end{aligned}$$

where $x = (x_1, x_2, \ldots, x_n) \in K^n$. $\|.\|_1$ is called l_1-norm and $\|.\|_\infty$ is called the sup norm on K^n.

Note 1.1.7. It is to be noted that $\|.\|_1$, $\|.\|_2$ and $\|.\|_\infty$ are norms on the linear space $K^n(\mathbb{R}_n \text{ or } \mathbb{C}_n)$. We now introduce the general class of norms K^n to which these norms relate.

Let $p > 0$ be a real number. Define

$$\begin{aligned} \|.\|_p : K^n &\longrightarrow \mathbb{R} \text{ by} \\ \|x\|_p &= \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}}, \ x = (x_1, x_2, \ldots, x_n) \in K^n. \end{aligned}$$

It is easy to verify that $\|.\|_p$ for $1 \leq p < \infty$ actually defines a norm on K^n. We denote the normed linear space $(K^n, \|.\|_p)$ by $l^p(n)$.

Theorem 1.1.8. The conjugate space of $l^p(n)$ is $l^q(n)$, where $1 < p, q < \infty$ and $\dfrac{1}{p} + \dfrac{1}{q} = 1$.

Proof. Let $(e_1, e_2, \ldots, e_n)$ be a natural basis for $l^p(n)$. Then, any $x = (x_1, x_2, \ldots, x_n) \in l^p(n)$ can be expressed uniquely in the form

$$x = \sum_{k=1}^{n} x_k e_k.$$

Since $l^p(n)$ is a finite dimensional normed linear space, every linear functional on $l^p(n)$ is continuous. Thus, if f is continuous linear functional defined on $l^p(n)$, then

$$f(x) = \sum_{k=1}^{n} x_k f(e_k), \;\; x = (x_1, x_2, \ldots, x_n) \in l^p(n).$$

Clearly, $(f(e_1), f(e_2), \ldots, f(e_n)) \in l^q(n)$. Thus for each $f \in (l^p(n))^*$ we obtain an element $u = (f(e_1), f(e_2), \ldots, f(e_n)) \in l^q(n)$.

This defines an operator $T : (l^p(n))^* \longrightarrow l^q(n)$ given by $T(f) = u$, $f \in (l^p(n))^*$ and $u = (f(e_1), f(e_2), \ldots, f(e_n)) \in l^q(n)$. It can be easily seen that T is linear and bijective.

We now show that T preserves norm. Since f is bounded, it follows by Hölder's inequality that

$$\begin{aligned} |f(x)| &< \sum_{k=1}^{n} |x_k f(e_k)| \\ &\leq \left(\sum_{k=1}^{n} |x_k|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^{n} |f(e_k)|^q\right)^{\frac{1}{q}} \\ &\leq \|x\|_p \left(\sum_{k=1}^{n} |f(e_k)|^q\right)^{\frac{1}{q}}. \end{aligned}$$

This implies that

$$\begin{aligned} \|f\| &= \sup\left\{\frac{|f(x)}{\|x\|_p} : \|x\|_p \neq 0\right\} \\ &\leq \left(\sum_{k=1}^{n} |f(e_k)|^q\right)^{\frac{1}{q}} \qquad (1.1.4) \\ &= \|u\|_q = \|T(f)\|_q. \qquad (1.1.5) \end{aligned}$$

Choosing $x = x_0 = (\lambda_1, \lambda_2, \ldots, \lambda_n) \in l^p(n)$, where

$$\begin{aligned} \lambda_k &= \frac{|f(e_k)|}{f(e_k)}, \;\; f(e_k) \neq 0 \\ &= 0, \;\; \text{otherwise}. \end{aligned}$$

Then $\|x_0\|_p = (\sum_{k=1}^{n} |\lambda_k|^p)^{\frac{1}{p}} = (\sum_{k=1}^{n} |f(e_k)|^q)^{\frac{1}{p}}$, since $\dfrac{1}{p} + \dfrac{1}{q} = 1$. Also,

$$f(x_0) = \sum_{k=1}^{n} \lambda_k f(e_k) = \sum_{k=1}^{n} |f(e_k)|^q.$$

Therefore,

$$\|f\| \geq \frac{|f(x_0)|}{\|x_0\|_p} = \left(\sum_{k=1}^{n} |f(e_k)|^q\right)^{\frac{1}{q}}, \quad \left[\because \frac{1}{p} + \frac{1}{q} = 1\right].$$

This implies that

$$\|f\| \geq \|u\|_q = \|T(f)\|_q. \tag{1.1.6}$$

Combining (1.1.5) and (1.1.6) we get

$$\|f\| = \|Tf\|_q.$$

Hence T preserves norm. Thus, T is an isometric isomorphism of $(l^p(n), \|.\|_p)^*$ onto $(l^q(n), \|.\|_q)$ and hence the cojugate space of $l^p(n)$ is $l_q(n)$. This proves the theorem. □

Remark 1.1.9. Note that, if $p = 2$, then $q = 2$. Also, $l^2(n) \equiv \mathbb{R}^n$ or $\mathbb{C}^n$ according as the field K is $\mathbb{R}$ or $\mathbb{C}$. As such, the conjugate space of $\mathbb{R}^n$ is $\mathbb{R}^n$ and that of $\mathbb{C}^n$ is $\mathbb{C}^n$.

Exercise 1.1.10. 1. Prove that the conjugate space of $\mathbb{C}^n$ is $\mathbb{C}^n$.

2. Prove that the conjugate space of $(\mathbb{C}^n, \|.\|_\infty)$ is the space $(\mathbb{C}^n, \|.\|_1)$.

1.1.1 Schauder Basis

Due to restriction to finite linear combinations, classical vector space bases are not always suitable for the analysis of infinite dimensional spaces. Therefore, it is natural in some way to consider generalised basic concepts. In 1927, J. Schauder introduced the notion of Schauder basis in a Banach space, which is defined as follows.

Definition 1.1.11. Let X be a normed linear space over the scalar $K(\mathbb{R}$ or $\mathbb{C})$. A sequence (x_n) in X is called a Schauder basis for X if $\|x_n\| = 1$ for $n = 1, 2, \ldots$, and each $x \in X$ can be expressed as $x = \sum_{n=1}^{\infty} \alpha_n x_n$ where the series converges in the norm of X, and the scalar α_n are uniquely determined by x.

Example 1.1.12. For $n \in \mathbb{N}$, let $e_n = (0, 0, \ldots, 1, 0, \ldots) \in K^n$. Then (e_n) is a Schauder basis of $l_p, 1 \leq p < \infty$ and c_0. We call (e_n) the unit vector of l_p and c_0 respectively.

Remark 1.1.13. Assume that X is a Banach space and (e_n) is a basis of X. Then,

i) (e_n) is linearly independent.

ii) span$\{e_n : n \in \mathbb{N}\}$ is dense in X, in particular X is separable.

iii) every element x is uniquely determined by the sequence $\lambda(\alpha_n)$ so that $x = \sum_{n=1}^{\infty} \alpha_n e_n$.

Remark 1.1.14. The space l_∞ is not separable and therefore has no Schauder basis. Every orthonormal basis in a separable Hilbert space is a Schauder basis.

Remark 1.1.15. Each basis in a Banach space is a Schauder basis.

Definition 1.1.16 (Signum Function)**.** If α is a complex number, then

$$\begin{aligned} \text{sgn}\alpha &= \frac{\alpha}{|\alpha|}, \text{ if } \alpha \neq 0 \\ &= 0, \quad \alpha = 0. \end{aligned}$$

From the above definition, we have the following two properties of signum function.

i) $|\text{sgn } \alpha|= 0$ if $\alpha = 0$ and $|\text{sgn } \alpha|= 1$ if $\alpha \neq 0$.

ii) $\alpha\text{sgn } \overline{\alpha} = 0$ if $\alpha = 0$ and if $\alpha \neq 0$, then $\alpha\text{sgn } \overline{\alpha} = \dfrac{\alpha\overline{\alpha}}{|\alpha|} = |\alpha|$.

Theorem 1.1.17. The conjugate space of l_p is l_q, where $1 < p, q < \infty$ and $\dfrac{1}{p} + \dfrac{1}{q} = 1$.

Proof. Let $x = (x_n) \in l_p$ so that $\displaystyle\sum_{n=1}^{\infty} |x_n|^p < \infty$. Let $(e_1, e_2, \ldots)$ be a Schauder basis for l_p. Then, $x = (x_n) \in l_p$ can be written as

$$x = \sum_{k=1}^{\infty} x_k e_k.$$

Let $f \in l_p^*$. Then using the linearity and continuity of f, we have

$$f(x) = \sum_{k=1}^{\infty} x_k f(e_k) = \sum_{k=1}^{\infty} x_k \alpha_k$$

where $\alpha_k = f(e_k), k = 1, 2, \ldots$.

We now define an operator $T : l_p^* \longrightarrow l_q$ by

$$T(f) = (\alpha_1, \alpha_2, \ldots) \tag{1.1.7}$$

and we show that T is an isometric isomorphism of l_p^* onto l_q.

First we show that T is well-defined. Let $x = (\beta_1, \beta_2, \ldots, \beta_n, 0, 0, \ldots) \in l_p$ where

$$\begin{aligned} \beta_k &= |\alpha_k|^{q-1}\text{sgn}\overline{\alpha}_k, \text{ if } 1 \leq k \leq n \\ &= 0, \text{ if } k > n. \end{aligned}$$

Then, $|\beta_k|^p = |\alpha_k|^{p(q-1)} = |\alpha_k|^q$, since $\dfrac{1}{p} + \dfrac{1}{q} = 1$.

Also, $\alpha_k\beta_k = \alpha_k|\alpha_k|^{q-1}\text{sgn}\overline{\alpha}_k = |\alpha_k|^q$.

Therefore, $\|x\| = \left(\displaystyle\sum_{k=1}^{n} |\beta_k|^p\right)^{\frac{1}{p}} = \left(\displaystyle\sum_{k=1}^{n} |\alpha_k|^q\right)^{\frac{1}{p}}$. Since $x = \sum_{k=1}^{n} \beta_k e_k$, we have

$$f(x) = \sum_{k=1}^{n} \beta_k f(e_k) = \sum_{k=1}^{n} \beta_k \alpha_k = \sum_{k=1}^{n} |\alpha_k|^q.$$

Now, for all $x \in l_p$ we have

$$\begin{aligned} |f(x)| &\leq \|f(x)\|\|x\| \\ \text{i.e., } \sum_{k=1}^{n} |\alpha_k|^q &\leq \|f\| \left(\sum_{k=1}^{n} |\alpha_k|^q\right)^{\frac{1}{p}} \\ \text{i.e., } \left(\sum_{k=1}^{n} |\alpha_k|^q\right)^{\frac{1}{q}} &\leq \|f\| < \infty. \end{aligned}$$

Since the last inequality is true for arbitrary positive integer n, letting $n \to \infty$ we get

$$\left(\sum_{k=1}^{n} |\alpha_k|^q\right)^{\frac{1}{q}} \leq \|f\| < \infty. \tag{1.1.8}$$

This shows that $(\alpha_k) \in l_q$ and hence T is well-defined.

From (1.1.7) it follows that $f = 0$ if $T(f) = 0$ so that Ker $T = \{0\}$. Hence T is one-one.

To prove that T is onto, we suppose that $(\beta_k) \in l_q$. Define the functional $g : l_p \longrightarrow K$ by $g(x) = \sum_{k=1}^{\infty} x_k\beta_k, x = (x_i) \in l_p$. Obviously g is linear and

$$\begin{aligned} |g(x)| &= \left|\sum_{k=1}^{\infty} x_k\beta_k\right| \\ &\leq \sum_{k=1}^{\infty} |x_k\beta_k| \\ &\leq \left(\sum_{k=1}^{\infty} |x_k|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^{\infty} |\beta_k|^q\right)^{\frac{1}{q}} \\ &= \|x\| \left(\sum_{k=1}^{\infty} |\beta_k|^q\right)^{\frac{1}{q}}. \end{aligned}$$

This shows that g is bounded. Since $e_k \in l_p$ for $k = 1, 2, \ldots$ we get $g(e_k) = \beta_k$ for all k and so $T(g) = (\beta_k)$ and hence T is onto.

Next we show that T preserves norm. Since $T(f) \in l_q$, from (1.1.7) and (1.1.8) we obtain

$$\|T(f)\| = \left(\sum_{k=1}^{\infty} |\alpha_k|^q\right)^{\frac{1}{q}} \leq \|f\|. \tag{1.1.9}$$

To prove the reverse inequality, let us take $x \in l_p$ so that $x = \sum_{k=1}^{\infty} x_k e_k$. Hence,

$$f(x) = \sum_{k=1}^{\infty} x_k f(e_k) = \sum_{k=1}^{\infty} x_k \alpha_k.$$

Using Holder's inequality we have

$$\begin{aligned} |f(x)| &\leq \sum_{k=1}^{\infty} |x_k\alpha_k| \\ &\leq \left(\sum_{k=1}^{\infty} |x_k|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^{\infty} |\alpha_k|^q\right)^{\frac{1}{q}} \\ &= \|x\| \left(\sum_{k=1}^{\infty} |\alpha_k|^q\right)^{\frac{1}{q}}. \end{aligned}$$

Therefore

$$\begin{aligned} \|f\| &= \sup\left\{\frac{|f(x)|}{\|x\|} : \|x\| \neq 0\right\} \\ &\leq \left(\sum_{k=1}^{\infty} |\alpha_k|^q\right)^{\frac{1}{q}} = \|T(f)\|. \end{aligned} \tag{1.1.10}$$

Combining (1.1.9) and (1.1.10) we get

$$\|f\| = \|T(f)\|.$$

From the definition of T it is clear that T is linear.

Therefore $T : l_p^* \longrightarrow l_q$ is an isomorphism. Hence the conjugate space of l_p is l_q. This proves the theorem. □

Note 1.1.18. From theorem (1.1.17), we note the following.

i) If $x = (x_n) \in l_p$ and $f \in l_p^*$, then f has the unique representation of the form

$$f(x) = \sum_{k=1}^{\infty} x_k f(e_k).$$

ii) The norm of $f \in l_p^*$ is given by

$$\|f\| = \left(\sum_{k=1}^{\infty} |f(e_k)|^q\right)^{\frac{1}{q}}.$$

Theorem 1.1.19. The conjugate space of l_1 is l_∞ .

Proof. We note that $l_1 = \left\{x = (x_n) : \sum_{n=1}^{\infty} |x_n| < \infty\right\}$ and $l_\infty = \left\{x = (x_n) : \sup_{1 \leq n < \infty} |x_n| < \infty\right\}$. Let (e_n) be a Schauder basis for l_1. Then any $x = (x_n) \in l_1$ can be expressed as $x = \sum_{k=1}^{\infty} x_k e_k$. Let $f \in l_1^*$. Then using the linearity and continuity of f we have

$$f(x) = \sum_{k=1}^{\infty} x_k f(e_k) = \sum_{k=1}^{\infty} x_k \alpha_k$$

where $\alpha_k = f(e_k), k = 1, 2, \ldots$ We now define an operator $T : l_1^* \longrightarrow l_\infty$ by

$$T(f) = (\alpha_1, \alpha_2, \ldots) \tag{1.1.11}$$

and show that T is an isometric isomorphism of l_1^* onto l_∞.

First we show that T is well-defined. For that, let $x = (\beta_1, \beta_2, \ldots, \beta_n, 0, 0, \ldots) \in l_1$ where

$$\begin{aligned} \beta_k &= \text{sgn}\,\overline{\alpha}_k, \text{ if } k = n \\ &= 0, \text{ if } k \neq n. \end{aligned}$$

Then, $\|x\| = |\text{sgn}\overline{\alpha}_k| = 1$. Also, $\alpha_n \beta_n = \alpha_n \text{sgn}\overline{\alpha}_n = |\alpha_n|$. Since $x = \sum_{k=1}^{n} \beta_k e_k$, we have

$$f(x) = \sum_{k=1}^{n} \beta_k f(e_k) = \sum_{k=1}^{n} \beta_k \alpha_k = \alpha_n \beta_n$$

and hence

$$|f(x)| = |\alpha_n| \leq \|f\| \|x\| = \|f\|.$$

Therefore,

$$\sup_{1 \leq n < \infty} |\alpha_n| \leq \|f\| \tag{1.1.12}$$

which implies that $(\alpha_n) \in l_\infty$ and hence T is well-defined.

From (1.1.11) it follows that $f = 0$ if $T(f) = 0$ so that $\text{Ker}T = \{0\}$. Hence T is one-one.

To prove that T is onto, we suppose that $(\beta_k) \in l_\infty$. Define the functional $g : l_1 \longrightarrow K$ by $g(x) = \sum_{k=1}^{\infty} x_k \beta_k, x = (x_n) \in l_1$. Obviously g is linear and

$$\begin{aligned} |g(x)| &= \left| \sum_{k=1}^{\infty} x_k \beta_k \right| \\ &\leq \sum_{k=1}^{\infty} |x_k \beta_k| \\ &\leq \left\{ \max_k \beta_k \right\} \sum_{k=1}^{\infty} |x_k| \\ &= \left\{ \max_k \beta_k \right\} \|x\|. \end{aligned}$$

This shows that g is bounded. Since $e_k \in l_1$ for $k = 1, 2, \ldots$ we get $g(e_k) = \beta_k$ for all k. So. $T(g) = (\beta_k)$ and hence T is onto.

Next we show that T preserves norm. Since $T(f) \in l_\infty$, from (1.1.11) and (1.1.12) we get

$$\sup_{1 \leq k < \infty} |\alpha_k| = \|T(f)\| \leq \|f\|. \tag{1.1.13}$$

To prove the reverse inequality, let $x = (x_n) \in l_1$ so that $x = \sum_{k=1}^{\infty} x_k e_k$. Hence,

$$f(x) = \sum_{k=1}^{\infty} x_k f(e_k) = \sum_{k=1}^{\infty} x_k \alpha_k.$$

So,

$$\begin{aligned}
|f(x)| &= \left|\sum_{k=1}^{\infty} x_k \alpha_k\right| \\
&\leq \sum_{k=1}^{\infty} |x_k \alpha_k| \\
&\leq \left\{\sup_k |\alpha_k|\right\} \sum_{k=1}^{\infty} |x_k| \\
&= \left\{\sup_k |\alpha_k|\right\} \|x\|.
\end{aligned}$$

So,

$$\begin{aligned}
\|f\| &= \sup\left\{\frac{|f(x)|}{\|x\|} : \|x\| \neq 0\right\} \\
&\leq \sup_k |\alpha_k| = \|T(f)\|. \qquad (1.1.14)
\end{aligned}$$

Combining (1.1.15) and (1.1.16) we get $\|f\| = \|T(f)\|$. From the definition of T, it is clear that T is linear.

Therefore, $T : l_1^* \longrightarrow l_\infty$ is an isomorphism. Hence, the conjugate space of l_1 is l_∞. This proves the theorem. □

Exercise 1.1.20. Prove that the conjugate space of C_0 is l_1.

1.1.2 Conjugate Space of C[a,b]

In order to determine the conjugate space of C[a,b], the class of all real valued continuous functions defined on [a,b] we shall require some results which are note to us.

1.1.3 Functions of Bounded Variation

Let a function $f(x)$ be defined in the closed interval $[a,b]$ and $a = x_0 < x_1 < \cdots < x_k < x_{k+1} < \cdots < x_n = b$ be a partition of $[a,b]$ into a finite number of subintervals $[x_k, x_{k+1}], k = 0, 1, 2, \cdots, n-1$. If $V = \sum_{k=0}^{n-1} |f(x_{k+1}) - f(x_k)|$, then V is called the variation corresponding to the particular partition. $\sup\{V\}$ is known as the total variation of $f(x)$ on $[a,b]$ and is denoted by $\underset{[a,b]}{V}(f)$ or simply by $V(f)$ when there is no confusion about the interval $[a,b]$ in consideration. If $V(f) < +\infty$, then the function $f(x)$ is said to be of bounded variation on $[a,b]$. The following results are known.

Theorem 1.1.21. If a function $f(x)$ is of bounded variation over $[a,b]$, then it is bounded there.

Theorem 1.1.22. If $f(x)$ and $g(x)$ are of bounded variation over $[a,b]$, then $f(x) \pm g(x)$ is also of bounded variation over $[a,b]$ and

$$V(f \pm g) \leq V(f) + V(g);$$

if c is a constant then cf is also of bounded variation over $[a,b]$ and

$$V(cf) = |c|V(f).$$

Theorem 1.1.23. A function $f(x)$ is of bounded variation over $[a,b]$ if and only if it can be expressed as the difference of two increasing functions.

1.1.4 Riemann Stieltjes Integral

Let $f(x)$ and $\phi(x)$ be two bounded functions defined on $[a,b]$ and $a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b$ be a partition of $[a,b]$. Let $t_{k+1} \in [x_k, x_{k+1}], k = 0, 1, 2, \cdots, n-1$ and $\alpha = \sum_{k=0}^{n-1} f(t_{k+1})[\phi(x_{k+1}) - \phi(x_k)]$. Then α is known as the Stieltjes sum. If α tends to a finite limit I as $\max|x_{k+1} - x_k| \to 0$ and if this limit is independent of the mode of subdivision of $[a,b]$ and the choice of the points t_k, then the limit I is known as Riemann Stieltjes integral of $f(x)$ with respect to $\phi(x)$ and is denoted by $\int_a^b f(x)d\phi(x)$.

The following results are known.

Theorem 1.1.24. If $f_1(x)$ and $f_2(x)$ are integrable on $[a,b]$ with respect to $\phi(x)$ and c_1, c_2 are constants, then $c_1 f_1(x) + c_2 f_2(x)$ is also integrable on $[a,b]$ w.r.t. $\phi(x)$ and

$$\int_a^b [c_1 f_1(x) + c_2 f_2(x)]d\phi(x) = c_1 \int_a^b f_1(x)d\phi(x) + c_2 \int_a^b f_2(x)d\phi(x).$$

Theorem 1.1.25. If $f(x)$ is integrable w.r.t. both $\phi_1(x)$ and $\phi_2(x)$ over $[a,b]$ and c_1, c_2 are constants, then $f(x)$ is integrable w.r.t $c_1 f_1(x) + c_2 f_2(x)$ over $[a,b]$ and

$$\int_a^b f(x)d[c_1 f_1(x) + c_2 f_2(x)] = c_1 \int_a^b f(x)d\phi_1(x) + c_2 \int_a^b f(x)d\phi_2(x).$$

Theorem 1.1.26. If one of the integrals $\int_a^b f(x)d\phi(x)$ and $\int_a^b \phi(x)df(x)$ exists, then the other integral also exists and

$$\int_a^b f(x)d\phi(x) + \int_a^b \phi(x)df(x) = f(b)\phi(b) - f(a)\phi(a).$$

Theorem 1.1.27. If $f(x)$ is continuous on $[a,b]$ and $g(x)$ is of bounded variation over $[a,b]$, then $\int_a^b f(x)dg(x)$ exists and

$$\left|\int_a^b f(x)dg(x)\right| \leq V(g). \sup_{a \leq x \leq b} |f(x)|.$$

1.1.5 The space BV[a,b]

We consider the set BV[a,b] which contains the class of all real valued functions which are of bounded variation over $[a,b]$. We define the sum $\phi = f_1 + f_2$ of two elements $f_1, f_2 \in$BV[a,b] by

$$\phi(t) = (f_1 + f_2)(t) = f_1(t) + f_2(t).$$

If α is a scalar then the scalar multiple of the element $f \in$BV[a,b] by α denoted by $\psi = \alpha f$ is defined by

$$\psi(t) = (\alpha f)(t) = \alpha f(t).$$

Also the function $f(t)$ such that $f(t) = 0, \forall t \in [a,b]$ is the zero element of BV[a,b]. The negative of $f \in$BV[a,b] is $(-f)(t) = -f(t)$. It is clear that $-f \in$BV[a,b]. We can now easily verify that all the axioms of a linear space are satisfied. For $f \in$BV[a,b] we define

$$\|f\| = |f(a)| + V(f)$$

and verify that the axioms of a norm are satisfied.

i) Clearly $\|f\| \geq 0$. If $f(t) = 0, \forall t \in [a,b]$ then obviously $\|f\| = 0$. Conversely, suppose that $\|f\| = 0$. Then $f(a) = 0$ and $V(f) = 0$. Let $t \in (a,b)$. Then

$$\begin{aligned}
|f(t) - f(a)| + |f(b) - f(t)| &\leq V(f) = 0 \\
\text{i.e., } f(t) = f(a) &\quad \text{and} \quad f(t) = f(b) \\
\text{i.e., } f(t) &= 0, \forall t \in [a,b].
\end{aligned}$$

ii) If α is a scalar then

$$\begin{aligned}
\|\alpha f\| &= |\alpha f(a)| + V(\alpha f) \\
&= |\alpha||f(a)| + |\alpha|V(f) \\
&= |\alpha|[|f(a)| + V(f)] \\
&= |\alpha|\|f\|.
\end{aligned}$$

iii) If $f, g \in$BV[a,b], then

$$\begin{aligned}
\|f + g\| &= |(f+g)(a)| + V(f+g) \\
&= |f(a) + g(a)| + V(f+g) \\
&\leq |f(a)| + |g(a)| + V(f) + V(g) \\
&= \|f\| + \|g\|.
\end{aligned}$$

BV[a,b], is therefore, a normed linear space and so a metric space.

1.1.6 The space B[a,b]

Let B[a,b] denotes the class of all bounded real valued functions defined on $[a,b]$. The sum $\phi = f_1 + f_2$ of two elements $f_1, f_2 \in$B[a,b] and the scalar multiple $\psi = \alpha f$ of the element $f \in$B[a,b] by the scalar α are defined by

$$\begin{aligned}
\phi(t) &= f_1(t) + f_2(t) \\
\text{and } \psi(t) &= \alpha f(t).
\end{aligned}$$

The negative $-f$ of an element $f \in$B[a,b] is

$$(-f)(t) = -f(t).$$

The function $f(t)$ such $f(t) = 0, \forall t \in [a,b]$ is the zero element of B[a,b].

With these definitions of addition and scalar multiplication it is easy to see that $B[a,b]$ is a real linear space. For $f \in B[a,b]$ we define

$$\|f\| = \sup_{a \leq t \leq b} |f(t)|.$$

It may be verified that the axioms of a norm are satisfied. $B[a,b]$, the class of all bounded real valued functions defined on $[a,b]$ is therefore a metric space where the distance $\rho(f,g)$ between two elements f, g of $B[a,b]$ is given by

$$\rho(f,g) = \|f - g\| = \sup_{a \leq t \leq b} |f(t) - g(t)|.$$

We now show that the convergence in $B[a.b]$ is equivalent to uniform convergence.

Let (f_n) be a sequence of elements of $B[a,b]$ which converges to an element $f \in B[a,b]$. Then $\rho(f_n, f) \to 0$ as $n \to \infty$. Let $\epsilon > 0$ be arbitrary. Then there is a positive integer N such that

$$\begin{aligned}
& \rho(f_n, f) < \epsilon \text{ if } n \geq N \\
\text{i.e.,} \quad & \sup_{a \leq t \leq b} |f_n(t) - f(t)| < \epsilon, \text{ if } n \geq N \\
\text{i.e.,} \quad & |f_n(x) - f(x)| < \epsilon \, \forall \, t \in [a,b], \text{ if } n \geq N.
\end{aligned}$$

This however means that the sequence $(f_n(t))$ converges uniformly to $f(t)$ in $B[a,b]$.

Conversely, we suppose that the sequence $(f_n(t))$ of bounded functions converges uniformly to the bounded function $f(t)$ over $[a,b]$. Let $\epsilon > 0$ be arbitrary. Then there is a positive integer N(depending only on ϵ) such that

$$\begin{aligned}
& |f_n(t) - f(t)| < \epsilon, \text{ if } n \geq N \text{ and } \forall t \in [a,b] \\
\text{i.e.,} \quad & \sup_{a \leq t \leq b} |f_n(t) - f(t)| < \epsilon \text{ if } n \geq N \\
\text{i.e.,} \quad & \rho(f_n, f) \leq \epsilon \text{ if } n \geq N \\
\text{i.e.,} \quad & f_n \to f \text{ in } B[a,b].
\end{aligned}$$

Thus convergence in $B[a,b]$ is equivalent to uniform convergence. It is to be noted that $C[a,b]$ is a subspace of $B[a,b]$.

Theorem 1.1.28. (Riesz Representation Theorem for $C[a,b]$) Let $f \in (C[a,b])^*$, that is, f is continuous linear functional defined on $C[a,b]$. Then there is a function $g(t) \in$BV[a,b] such that

$$f(x) = \int_a^b x(t)dg(t), \forall x(t) \in C[a,b]$$

and $\|f\| = V(g)$.

Proof. The proof of the theorem is out of the scope of this study material. □

To prove the next theorem we shall require some results from real analysis.

Result 1.1.29. Let $f(x)$ be monotone increasing in $[a,b]$ and $x_0 \in [a,b]$. Then

$$\begin{aligned}
f(x_0+) &= \lim_{x \to x_0+} f(x) \text{ exists and} \\
f(x_0+) &= \inf\{f(x) : x_0 < x < b\}.
\end{aligned}$$

Result 1.1.30. If $f(x)$ is increasing in $[a,b]$ and $c \in [a,b]$, then

$$\lim_{h \to 0} \frac{1}{h} \int_c^{c+h} f(t)dt = \lim_{x \to c+} f(x) = f(c+0).$$

Result 1.1.31. Let $f(x)$ be of bounded variation over $[a,b]$ and $c \in [a,b]$. Then,

$$\lim_{h \to 0} \frac{1}{h} \int_c^{c+h} f(t)dt = \lim_{x \to c+} f(x) = f(c+0).$$

Proof. Since $f(x)$ is of bounded variation over $[a, b]$, we have

$$f(x) = \phi(x) - \psi(x)$$

where both $\phi(x)$ and $\psi(x)$ are increasing functions on $[a, b]$. Then,

$$\begin{aligned}
\lim_{h\to 0} \frac{1}{h} \int_c^{c+h} f(t)dt &= \lim_{h\to 0} \frac{1}{h} \int_c^{c+h} [\phi(t) - \psi(t)]dt \\
&= \lim_{h\to 0} \frac{1}{h} \int_c^{c+h} \phi(t)dt - \lim_{h\to 0} \frac{1}{h} \int_c^{c+h} \psi(t)dt \\
&= \phi(c+0) - \psi(c+0) \\
&= f(c+0).
\end{aligned}$$

□

Result 1.1.32. Suppose that $f(x)$ is of bounded variation over $[a, b]$ and $c \in [a, b]$. Then,

$$\lim_{h\to 0} \frac{1}{h} \int_{c-h}^{c} f(t)dt = \lim_{x\to c-} f(x) = f(c-0).$$

Definition 1.1.33. For $f, g \in$BV[a,b], we define $f \sim g$ if for all $x(t) \in C[a, b]$

$$\int_a^b x(t)df(t) = \int_a^b x(t)dg(t).$$

It can be easily shown that '$\sim$' is an equivalence relation in BV[a,b].

Lemma 1.1.34. Let $f(t) \in$BV[a,b] and $f \sim 0$. Then for any $c \in [a, b]$,

$$f(a) = f(b) = f(c+0) = f(c-0)$$

where $f(c+0) = \lim_{x\to c+} f(x)$ and $f(c-0) = \lim_{x\to c-} f(x)$

Proof. Since $f \sim 0$, we have for all $x(t) \in C[a, b]$

$$\int_a^b x(t)df(t) = 0.$$

Choosing $x(t) = 1$ we get

$$0 = \int_a^b df(t) = f(b) - f(a)$$

$$\text{i.e.,} f(a) = f(b).$$

Since $f(x)$ is of bounded variation on $[a, b]$, we have

$$\lim_{h\to 0} \frac{1}{h} \int_c^{c+h} f(t)dt = f(c+0).$$

We now show that $f(a) = f(c+0)$. The argument to show that this is also equal to $f(c-0)$ is quite similar and hence omitted. We consider the function

$$\begin{aligned}
g(t) &= 1, \text{ if } a \leq t \leq c \\
&= 1 - \frac{t-c}{h}, \text{ if } c < t \leq c+h \\
&= 0, \text{ if } c+h < t \leq b.
\end{aligned}$$

We note that $g(t)$ is continuous in $[a,b]$. Then we have

$$\begin{aligned}
0 &= \int_a^b g(t)df(t) \\
&= \int_a^c g(t)df(t) + \int_c^{c+h} g(t)df(t) + \int_{c+h}^b g(t)df(t) \\
&= \int_a^c df(t) + \int_c^{c+h} g(t)df(t) \\
&= f(c) - f(a) + \int_c^{c+h} g(t)df(t). \qquad (1.1.15)
\end{aligned}$$

By the formula for integration by parts of Riemann-Stieltjes integral we obtain

$$\begin{aligned}
\int_c^{c+h} g(t)df(t) + \int_c^{c+h} f(t)dg(t) &= f(c+h)g(c+h) - f(c)g(c) \\
\text{i.e., } \int_c^{c+h} g(t)df(t) &= \frac{1}{h}\int_c^{c+h} f(t)dt - f(c).
\end{aligned}$$

So from (1.1.15) we obtain

$$f(a) = \lim_{h\to 0} \frac{1}{h}\int_c^{c+h} f(t)dt.$$

Letting $h \to 0$ we obtain

$$f(a) = \lim_{h\to 0} \frac{1}{h}\int_c^{c+h} f(t)dt = f(c+0).$$

In a similar way we can show that

$$f(b) = f(c-0).$$

This proves the lemma. $\square$

Definition 1.1.35. The function $f(t)$ ∈BV[a,b] is called normalised if $f(a) = 0$ and $\lim_{t\to t_0+} f(t) = f(t_0), \forall t_0 \in (a,b)$, i.e., if f is continuous from the right. The collection of all normalised functions of bounded variation is denoted by NBV[a,b]. It is easy to see that NBV[a,b] is a subspace of BV[a,b].

Lemma 1.1.36. Let $f_1, f_2 \in$ BV[a,b]. If f_1, f_2 are normalised and $f_1 \sim f_2$, then $f_1 = f_2$.

Proof. Since $f_1 \sim f_2$ we have $f_1 - f_2 \sim 0$. So, by lemma (1.1.34) we have

$$\begin{aligned}
& (f_1 - f_2)(b) = (f_1 - f_2)(a) \\
\text{i.e.,} \quad & f_1(b) - f_2(b) = f_1(a) - f_2(a) = 0 \\
\text{i.e.,} \quad & f_1(b) = f_2(b).
\end{aligned}$$

Further, for any $c \in (a,b)$, we have

$$\begin{aligned}
& (f_1 - f_2)(c+0) = (f_1 - f_2)(a) = 0 \\
\text{i.e.,} \quad & f_1(c+0) - f_2(c+0) = 0 \\
\text{i.e.,} \quad & f_1(c+0) = f_2(c+0).
\end{aligned}$$

Since f_1 and f_2 are continuous from the right, it follows that $f_1(c) = f_2(c), \forall c \in (a,b)$ and hence $f_1 = f_2$. This proves the lemma. $\square$

Lemma 1.1.37. Let $f(t)$ ∈BV[a,b]. Then there exist a function $g(t)$ ∈NBV[a,b] such that

$$f \sim g \text{ and } V(g) \leq V(f).$$

Proof. The proof of the lemma is beyond the scope of this study material. □

Theorem 1.1.38. The spaces NBV[a,b] and $(C[a,b])^*$ are isometrically isomorphic.

Proof. Let $g(t)$ ∈NBV[a,b]. For $x(t) \in C[a,b]$, let $f(x) = \int_a^b x(t)dg(t)$. Let $x_1(t), x_2(t)$ ∈C[a,b] and λ_1, λ_2 be scalars. Then

$$\begin{aligned} f(\lambda_1 x_1 + \lambda_2 x_2) &= \int_a^b [\lambda_1 x_1(t) + \lambda_2 x_2(t)]dg(t) \\ &= \lambda_1 \int_a^b x_1(t)dg(t) + \lambda_2 \int_a^b x_2(t)dg(t) \\ &= \lambda_1 f(x_1) + \lambda_2 f(x_2). \end{aligned}$$

This shows that f is linear. Further by theorem 1.1.27

$$\begin{aligned} |f(x)| &= \left| \int_a^b x(t)dg(t) \right| \\ &\leq \sup_{a \leq t \leq b} |x(t)|.V(g) \\ &= V(g).\|x\|. \end{aligned}$$

Therefore f is bounded and

$$\|f\| \leq V(g). \tag{1.1.16}$$

Thus for every $g(t)$ ∈NBV[a,b] we obtain an element $f \in (C[a,b])^*$. This defines an operator

$T : NBV[a,b] \longrightarrow (C[a,b])^*$ given by $T(g) = f$ where $g \in NBV[a,b]$ and

$$f(x) = \int_a^b x(t)dg(t), \forall x(t) \in C[a,b].$$

Let $g_1, g_2 \in NBV[a,b]$ and λ_1, λ_2 be scalars. Further let $T(g_1) = f_1$ and $T(g_2) = f_2$. Then for all $x(t) \in C[a,b]$

$$\begin{aligned} f_1(x) &= \int_a^b x(t)dg_1(t) \text{ and} \\ f_2(x) &= \int_a^b x(t)dg_2(t). \end{aligned}$$

Since

$$\begin{aligned} \int_a^b x(t)d[\lambda_1 g_1(t) + \lambda_2 g_2(t)] &= \lambda_1 \int_a^b x(t)dg_1(t) + \lambda_2 \int_a^b x(t)dg_2(t) \\ &= \lambda_1 f_1(x) + \lambda_2 f_2(x) \\ &= (\lambda_1 f_1 + \lambda_2 f_2)(x), \end{aligned}$$

we have

$$\begin{aligned} T(\lambda_1 g_1 + \lambda_2 g_2) &= \lambda_1 f_1 + \lambda_2 f_2 \\ &= \lambda_1 T(g_1) + \lambda_2 T(g_2). \end{aligned}$$

This shows that T is linear. We now show that T is one-one.

Let $g_1, g_2 \in NBV[a,b]$ and $T(g_1) = T(g_2)$. Then for all $x(t) \in C[a,b]$,

$$\int_a^b x(t) dg_1(t) = \int_a^b x(t) dg_2(t).$$

This however means that $g_1 \sim g_2$ and so by Lemma 1.1.36, $g_1 = g_2$. Thus T is one-one.

We now show that T is surjective. Let $f \in (C[a,b])^*$. Then by Theorem 1.1.28 there is a function $h(t) \in BV[a,b]$ such that

$$\begin{aligned} f(x) &= \int_a^b x(t) dh(t) \ \forall x(t) \in C[a,b] \\ \text{i.e.,} \qquad \|f\| &= V(h). \end{aligned} \tag{1.1.17}$$

By Lemma 1.1.37, there exist a unique $g(t) \in NBV[a,b]$ such that

$$h \sim g \text{ and } V(g) \leq V(h). \tag{1.1.18}$$

Thus for all $x(t) \in C[a,b]$ we obtain

$$\int_a^b x(t) dg(t) = \int_a^b x(t) dh(t) = f(x).$$

So, $T(g) = f$ and hence T is surjective. By (1.1.16), (1.1.17) and (1.1.18) we get

$$\begin{aligned} \|f\| \leq V(g) &\leq V(h) = \|f\| \\ \text{i.e., } \|f\| = V(g) &= |g(a)| + V(g) \ \ [\text{since } g(a) = 0] \\ &= \|g\| \\ \text{i.e., } \|T(g)\| &= \|g\|. \end{aligned}$$

This shows that T preserves norm. T, is therefore, an isomorphic isomorphism of $NBV[a,b]$ onto $(C[a,b])^*$ and so the conjugate space of $C[a,b]$ is $NBV[a,b]$. This proves the theorem. $\square$

Unit 2

The weak and weak* Convergence

Course Structure

- Weak convergence and weak* convergence: Definition, characterization of weak convergence and weak* convergence, sufficient condition for the equivalence of weak* convergence and weak convergence in the dual space.

We are all familiar with the convergence in norm in normed linear spaces. Now, we shall introduce two new types of convergence called weak convergence and weak* convergence.

Definition 2.0.1. A sequence (x_n) in a normed linear space X is said to be weakly convergent in X, if there is a point $x \in X$ such that

$$\lim_{n\to\infty} f(x_n) = f(x) \text{ for all } f \in X^*.$$

In this case we write $x_n \xrightarrow{w} x$ and call x as the weak limit of the sequence (x_n).

The convergence in the normed linear space X will now be called strong convergence, that is, $y_n \to y$ strongly in X if and only if $d(y, y_n) \to 0$, i.e., $\|y - y_n\| \to 0$ as $n \to \infty$.

Theorem 2.0.2. Let $x_n \xrightarrow{w} x$ in a normed linear space X. Then

i) the weak limit x of (x_n) is unique;

ii) every subsequence of (x_n) converges weakly to x;

iii) the sequence $(\|x_n\|)$ is bounded.

Proof. i) Suppose $x_n \xrightarrow{w} x$ and $x_n \xrightarrow{w} y$ in X. Then for each $f \in X^*$ we have

$$\lim_{n\to\infty} f(x_n) = f(x)$$
$$\text{and} \quad \lim_{n\to\infty} f(x_n) = f(y).$$

Since $f(x_n)$ is a sequence of scalars, its limit is unique. Hence, $f(x) = f(y)$. Since this is true for all $f \in X^*$, we have $x = y$. Thus the weak limit is unique.

ii) Since $x_n \xrightarrow{w} x$ in X, for each $f \in X^*$ we have

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

Since $f(x_n)$ is a convergent sequence of scalars, every subsequence $f(x_{n_k})$ converges for every $f \in X^*$ and has the same limit as the sequence. Therefore,

$$\lim_{k\to\infty} f(x_{n_k}) = f(x) \text{ and hence } x_{n_k} \xrightarrow{w} x.$$

iii) Since $x_n \xrightarrow{w} x$ in X, for each $f \in X^*$ we have

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

Since $f(x_n)$ is a convergent sequence of scalars, it is bounded for all f. Hence,

$$|f(x_n)| \leq C_f, \text{ for all } f,$$

where C_f is a positive constant depending on f. We define

$$F_{x_n}(f) = f(x_n), \forall f \in X^*.$$

Then

$$|F_{x_n}(f)| = |f(x_n)| \leq C_f \; \forall n.$$

This shows that for any $f \in X^*$, the sequence $(F_{x_n}(f))$ is bounded. Since X^* is a Banach space, the principle of uniform boundedness implies that $(\|F_{x_n}\|)$ is bounded. Now,

$$\begin{aligned} \|F_{x_n}\| &= \sup\left\{\frac{|F_{x_n}(f)|}{\|f\|} : \|f\| \neq 0\right\} \\ &= \sup\left\{\frac{|f(x_n)|}{\|f\|} : \|f\| \neq 0\right\} \\ &= \|x_n\|. \end{aligned}$$

Thus the sequence $(\|x_n\|)$ is bounded. This completes the proof.

□

Theorem 2.0.3. In any normed linear space, strong convergence implies weak convergence with the same limit but not conversely.

Proof. Let (x_n) converges strongly to x. Then

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

For arbitrary $f \in X^*$ we have

$$\begin{aligned} |f(x_n) - f(x)| &= |f(x_n - x)| \\ &\leq \|f\|\|x_n - x\| \\ &\to 0 \text{ as } n \to \infty. \end{aligned}$$

Therefore, $x_n \xrightarrow{w} x$. □

The converse is not true as shown by the following example.

Example 2.0.4. We consider the Schauder basis $e_1 = \{1, 0, 0, \ldots\}, e_2 = \{0, 1, 0, \ldots\}, \ldots$ in l_2. Let $x = (x_n) \in l_2$ such that $\sum_{n=1}^{\infty} |x_n|^2 < \infty$. Then,

$$x = \sum_{k=1}^{\infty} x_k \, \mathrm{e}_k$$

and hence $f(x) = \sum_{k=1}^{\infty} x_k f(e_k) = \sum_{k=1}^{\infty} x_k \alpha_k$ where $\alpha_k = f(e_k), k = 1, 2, \ldots$ and $(\alpha_k) \in l_2$. Since $(\alpha_k) \in l_2$ we have $\alpha_n \to 0$ as $n \to \infty$ and therefore $f(e_n) \to 0$ as $n \to \infty$.

This shows that $e_n \xrightarrow{w} \theta$ for all $f \in l_2^*$. Now for $n \neq m$, $\|e_n - e_m\|^2 = 2 \neq 0$ and so the sequence (e_n) cannot converge strongly to any element.

This proves the theorem.

Theorem 2.0.5. In a finite dimensional normed linear space X the notion of strong convergence and weak convergence are equivalent.

Proof. Since strong convergence implies weak convergence in any normed linear space, it is also true in a finite dimensional normed linear space. So it is enough if we prove that weak convergence implies strong convergence in a finite dimensional normed linear space.

Let X be a finite dimensional normed linear space and (x_n) be a sequence of elements in X such that $x_n \xrightarrow{w} x_0$. Since X is finite dimensional, there exist a finite number of linearly independent elements $e_1, e_2, \cdots, e_k$ in X such that $x \in X$ can be represented as

$$x = \alpha_1 e_1 + \alpha_2 e_2 + \cdots + \alpha_k e_k$$

where $\alpha_1, \alpha_2, \cdots, \alpha_k$ are scalars.

Therefore, we can write

$$\begin{aligned} x_n &= \alpha_1^{(n)} e_1 + \alpha_2^{(n)} e_2 + \cdots + \alpha_k^{(n)} e_k, \;\; n = 1, 2, 3, \ldots \\ \text{and} \;\; x_0 &= \alpha_1^{(0)} e_1 + \alpha_2^{(0)} e_2 + \cdots + \alpha_k^{(0)} e_k. \end{aligned}$$

We now define functionals $f_1, f_2, \cdots, f_k$ over X as follows:

If $x = \alpha_1 e_1 + \alpha_2 e_2 + \cdots + \alpha_k e_k \in X$, then $f_i(x) = \alpha_i, i = 1, 2, \cdots, k$.

Clearly each f_i is linear. Since X is finite dimensional, each f_i is bounded and so continuous. Hence $f_i \in X^*$ for $i = 1, 2, \cdots, k$. Now,

$$f_i(x_n) = \alpha_i^{(n)} \text{ and } f_i(x_0) = \alpha_i^{(0)}.$$

Since $x_n \xrightarrow{w} x_0$, we have $f_i(x_n) \to f_i(x_0)$ and so $\alpha_i^{(n)} \to \alpha_i^{(0)}$ as $n \to \infty$ for $i = 1, 2, \ldots, k$.

Let $M = \max \|e_i\|, i = 1, 2, \cdots, k$ and $\epsilon > 0$ be arbitrary. Then there exist a positive integer n_0 such that

$$|\alpha_i^{(n)} - \alpha_i^{(0)}| < \frac{\epsilon}{MK} \text{ for } n \geq n_0 \text{ and } i = 1, 2, \cdots, k.$$

Then for $n \geq n_0$,

$$\begin{aligned} \|x_n - x_0\| &= \|\sum_{i=1}^{k}(\alpha_i^{(n)} - \alpha_i^{(0)})e_i\| \\ &\leq \sum_{i=1}^{k}|\alpha_i^{(n)} - \alpha_i^{(0)}|.M \\ &< \epsilon. \end{aligned}$$

Therefore, (x_n) converges strongly to x_0. This completes the proof. □

Theorem 2.0.6. In a normed linear space X, $x_n \xrightarrow{w} x$ if and only if

i) the sequence $(\|x_n\|)$ is bounded and

ii) for every element f of a subset M of X^* which is everywhere dense in X^*, we have $f(x_n) \to f(x)$.

Proof. We first assume that $x_n \xrightarrow{w} x$. Then (i) follows from (iii) of theorem 2.0.2 and (ii) follows from the definition of weak convergence.

Next we suppose that 2.0.6 and 2.0.6 hold. We have to show that $f(x_n) \to f(x)$ for arbitrary $f \in X^*$. Let $c > 0$ be a number such that $\|x_n\| < c$ for all n and also $\|x\| < c$. Let $f \in X^*$ be arbitrary. Since M is everywhere dense in X^*, corresponding to $\epsilon > 0$ there exists $f_j \in M$ such that

$$\|f_j - f\| < \frac{\epsilon}{3c}.$$

Since $f_j \in M$, by 2.0.6 we have

$$f_j(x_n) \to f_j(x).$$

So there exist N such that for all $n > N$

$$|f_j(x_n) - f_j(x)| < \frac{\epsilon}{3}.$$

So for all $n > N$,

$$\begin{aligned} |f(x_n) - f(x)| &\leq |f(x_n) - f_j(x_n)| + |f_j(x_n) - f_j(x)| + |f_j(x) - f(x)| \\ &< \|f - f_j\|\|x_n\| + \frac{\epsilon}{3} + \|f_j - f\|\|x\| \\ &< \frac{\epsilon}{3c}.c + \frac{\epsilon}{3} + \frac{\epsilon}{3c}.c \\ &= \epsilon \end{aligned}$$

Since this is true for arbitrary $f \in X^*$, it follows that $x_n \xrightarrow{w} x$. This proves the theorem. □

Theorem 2.0.7. Let $T \in B(X,Y)$. If $x_n \xrightarrow{w} x_0$ in X, then $T(x_n) \xrightarrow{w} T(x_0)$ in Y.

Proof. Let $K(= \mathbb{R}$ or $\mathbb{C})$ denote the scalar field of X and Y. Let $f \in Y^*$. Then clearly $T : X \longrightarrow Y$ and $f : Y \longrightarrow K$. We define the composite map $fT : X \longrightarrow K$ by $(fT)(x) = f(Tx)$. Let $x_1, x_2 \in X$ and $\alpha, \beta \in K$. Then

$$\begin{aligned} fT(\alpha x_1 + \beta x_2) &= f(T(\alpha x_1 + \beta x_2)) \\ &= f(\alpha T(x_1) + \beta T(x_2)) \\ &= \alpha f(T(x_1)) + \beta f(T(x_2)) \\ &= \alpha(fT)(x_1) + \beta(fT)(x_2). \end{aligned}$$

Thus fT is a linear functional. Also

$$\begin{aligned}|fT(x)| &= |f(T(x))| \\ &\leq \|f\|\|Tx\| \\ &\leq \|f\|\|T\|\|x\|, \forall x \in X.\end{aligned}$$

Therefore fT is bounded and hence $fT \in X^*$. Since $x_n \xrightarrow{w} x_0$ in X, for every $f \in Y^*$, as $fT \in X^*$, by Theorem 2.0.6 we have

$$\begin{aligned}&(fT)(x_n) \to (fT)(x_0) \\ \text{i.e.,}\qquad &f(T(x_)) \to f(T(x_0)).\end{aligned}$$

Thus $T(x_n) \xrightarrow{w} T(x_0)$ in Y. This proves the theorem. □

Definition 2.0.8. A sequence (f_n) in X^* is said to be weak* convergent if there is some $f_0 \in X^*$ such that

$$\lim_{n\to\infty} f_n(x) = f_0(x) \quad \forall x \in X.$$

In this case we write $f_n \xrightarrow{w*} f_0$.

Remark 2.0.9. The nomenclature 'weak* convergence' comes from the fact that the dual space of X is denoted by X^*.

Remark 2.0.10. Weak* convergence is just pointwise convergence of the operators f_n.

Remark 2.0.11. If we have a subsequence $(f_n)_{n\in\mathbb{N}}$ in X^*, then we can consider three types of convergence of f_n to f_0: strong, weak and weak*.

By definition, these are as follows:

i) $f_n \to f_0$ strongly if and only if $\|f_n - f_0\| \to 0$ as $n \to \infty$.

ii) $f_n \xrightarrow{w} f_0$ if and only if $\lim\limits_{n\to\infty} T(f_n) = T(f_0) \quad \forall T \in X^{**}$.

iii) $f_n \xrightarrow{w*} f_0$ if and only if $\lim\limits_{n\to\infty} f_n(x) = f_0(x) \quad \forall x \in X$.

Theorem 2.0.12. Weak* limits are unique.

Proof. Suppose that X is a normed linear space. If possible, let $f_n \xrightarrow{w*} f$ and $f_n \xrightarrow{w*} g$ in X^*. Then by definition we have

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

for all $x \in X$. This implies $f = g$. Hence weak* limits are unique. This proves the theorem. □

Theorem 2.0.13. In a dual space, strong convergence implies weak* convergence but no conversely.

Proof. Let X be a normed linear space and X^* be its dual space. Let (f_n) be a sequence in X^* and $f_n \to f$ strongly in X^*. Then $\|f_n - f\| \to 0$ as $n \to \infty$. Now for all $x \in X$,

$$\begin{aligned}|f_n(x) - f(x)| &= |(f_n - f)(x)| \\ &\leq \|f_n - f\|\|x\| \\ &\to 0 \text{ as } n \to \infty.\end{aligned}$$

This shows that $f_n \xrightarrow{w*} f$ in X^*. □

Weak* convergence not necessarily imply strong convergence as shown by the following example.

Example 2.0.14. Consider the Banach space $X = (C_0, \|.\|_\infty)$ so that $X^* = (l_1, \|.\|_1)$. Let $f_n = (0, 0, \cdots, 0, 1, 0, \cdots)$ where 1 is in the nth place be the nth coordinate functional defined on C_0. If $x = (x_n) \in C_0$, then $f_n(x) = x_n$. Therefore $f_n(x) \to 0$ as $n \to \infty$ for each $x \in X$ so that the sequence (f_n) is weak* convergent in X^*. But $\|f_n\| = 1, n \in \mathbb{N}$. Therefore (f_n) is not strongly convergent in X^*. This proves the theorem.

Definition 2.0.15. Let X be a normed linear space.

i) A sequence (x_n) in X is said to be a weak Cauchy sequence if $(f(x_n))$ is a Cauchy sequence for all $f \in X^*$.

ii) The space X is said to be weakly complete if every weak Cauchy sequence in X has a weak limit in X.

Theorem 2.0.16. Let X be a normed linear space. Then the following holds:

i) A weak Cauchy sequence in X is bounded.

ii) If $(x_n) \subset X$ converges weakly to $x \in X$, then $\|x\| \leq \liminf_{n\to\infty} \|x_n\|$.

iii) If X is strongly complete, it need not be weakly complete.

Proof. i) Let (x_n) be a weak Cauchy sequence in X. Then $(f(x_n))$ is a Cauchy sequence in K for all $f \in X^*$. Therefore $\lim_{n\to\infty} f(x_n)$ exists for each $f \in X^*$. This implies that $x_n \xrightarrow{w} x$. Hence in the view of (iii) of Theorem 2.0.2, the sequence $(\|x_n\|)$ is bounded.

ii) Since $x_n \xrightarrow{w} x$ in X, we have $\lim_{n\to\infty} f(x_n) = f(x), \forall f \in X^*$. Now,

$$\begin{aligned} |f(x)| &= \lim_{n\to\infty} |f(x_n)| \\ &\leq \|f\| \liminf_{n\to\infty} \|x_n\|. \end{aligned}$$

This shows that

$$\begin{aligned} \|x\| &= \sup\{|f(x)| : f \in X^*, \|f\| = 1\} \\ &\leq \liminf_{n\to\infty} \|x_n\|. \end{aligned}$$

iii) Consider the Banach space $X = (C_0, \|.\|_\infty)$. We show that X is not weakly complete. Let $x = (\xi_i) \in C_0$ and $y = (\eta_i) \in l_1$. Then $f(x) = \sum_{i=1}^{\infty} \xi_i \eta_i$ implies $f(e_k) = \eta_k$, (e_n) being the unit vectors in C_0. Therefore,

$$\lim_{k\to\infty} f(e_k) = \lim_{k\to\infty} \eta_k = 0.$$

Thus (e_k) is a weakly Cauchy sequence in C_0.

Let, if possible, $e_k \xrightarrow{w} x_0$ in C_0 for some $x_0 = (\xi_i^0) \in C_0$. Then $f(e_k - x_0) \to 0$ as $k \to \infty$ for all $f \in l_1$. Taking $f = e_k$, we see that

$$\begin{aligned} |1 - \xi_n^0| &= 0, \forall n \geq 1 \\ \Rightarrow \xi_n^0 &= 1, \forall n \geq 1. \end{aligned}$$

Thus $x_0 = (1, 1, 1, \cdots) \notin C_0$. Hence (e_k) does not converge weakly in C_0. This proves the theorem.

□

Definition 2.0.17. A subset M of a normed linear space X is said to be a fundamental (or total) set if the span M is dense in X, i.e., $\overline{\text{span } M} = X$.

Theorem 2.0.18. Let X be a Banach space and let $(f_n) \subset X^*$ be a sequence. Then (f_n) is weak* convergent if and only if

i) the sequence $(\|f_n\|)$ is bounded; and

ii) the sequence $(f_n(x))$ is Cauchy for each $x \in M$, where M is fundamental subset of X.

Proof. Let $f_n \xrightarrow{w*} f$ in X^*. Then

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

This shows that $(f_n(x))$ is bounded for all $x \in X$. But X being complete, Principle of Uniform Boundedness when applied to bounded linear functionals gives that $(\|f_n\|)$ is bounded. This proves (i).

Note that (ii) is trivial, since $(f_n(x))$ is a convergent sequence of scalars for each $x \in X$, in particular, for $x \in M$.

Conversely, suppose that (i) and (ii) hold. Since the sequence $(\|f_n\|)$ is bounded, there exist a constant c such that $\|f_n\| \leq c, \forall n \in \mathbb{N}$.

Let $\epsilon > 0$ be given. Since $\overline{\text{span } M} = X$, it follows that for each $x \in X$, there exist a $y \in \text{span } M$ such that $\|x - y\| < \dfrac{\epsilon}{3c}$.

For $y \in \text{span } M$, (ii) implies that the sequence $(f_n(y))$ is Cauchy. Hence there exist a positive integer N such that $|f_n(y) - f_m(y)| < \dfrac{\epsilon}{3}$ for all $n, m \geq N$.

Now for an arbitrary $x \in X$, we have

$$\begin{aligned} |f_n(x) - f_m(x)| &\leq |f_n(x) - f_n(y)| + |f_n(y) - f_m(y)| + |f_m(y) - f_m(x)| \\ &\leq \|f_n\|\|x - y\| + \frac{\epsilon}{3} + \|f_m\|\|x - y\| \\ &< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} \\ &= \epsilon, \forall n, m \geq N. \end{aligned}$$

This shows that $(f_n(x))$ is a Cauchy sequence in $\mathbb{R}$. But $\mathbb{R}$ being complete, $(f_n(x))$ converges to $f(x)$, say, in $\mathbb{R}$. Further, x is an arbitrary element of X. Therefore,

$$\lim_{n\to\infty} f_n(x) = f(x), \forall x \in X.$$

Thus $f_n \xrightarrow{w*} f$. This proves the theorem. $\square$

Unit 3

Reflexive Spaces

Course Structure

- Reflexive spaces: Definition of reflexive space, canonical mapping, relation between reflexivity and separability, some consequences of reflexivity.

In this chapter we shall assume throughout that the spaces considered are all normed linear spaces.

Definition 3.0.1. Let X be a normed linear space and consider the conjugate space X^*. We know that X^* is a Banach space with the norm

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

As X^* was constructed from X, we can form successively the spaces $(X^*)^* = X^{**}, (X^{**})^* = X^{***}$ and so on.

We shall mainly concentrate on the space X^{**}, which is known as the second conjugate space of X. Suppose $x \in X$ is fixed and $f \in X^*$ is variable. Then for different $f \in X^*$, we obtain different values of $f(x)$. Therefore, the expression $f(x)$ where x is fixed and f is variable, represents a certain functional F_x, say, over X^*. So we write

$$F_x(f) = f(x)$$

where x is fixed and f is variable. We show that F_x is a continuous linear functional defined on X^* and therefore $F_x \in X^{**}$.
Let $f_1, f_2 \in X^*$ and λ be a scalar. Then

$$\begin{aligned} F_x(f_1 + f_2) &= (f_1 + f_2)(x) \\ &= f_1(x) + f_2(x) \\ &= F_x(f_1) + F_x(f_2). \end{aligned}$$

Also, $F_x(\lambda f_1) = (\lambda f_1)(x) = \lambda f_1(x) = \lambda F_x(f_1)$. Further,

$$|F_x(f)| = |f(x)| \leq \|x\| \|f\|, \forall f \in X^*.$$

This shows that F_x is linear and bounded and hence $F_x \in X^{**}$. Thus for each $x \in X$ there corresponds a unique continuous linear functional $F_x \in X^{**}$ given by

$$F_x(f) = f(x) \quad \forall f \in X^*$$

This defines a mapping $C : X \to X^{**}$ by $C(x) = F_x$ if and only if $F_x(f) = f(x)$ for all $f \in X^*$. This mapping C is called the cannonical mapping or cannonical embedding of X into X^{**}.

Next we show that C is an isometric isomorphism between X and the range of C, which is a subset of X^{**}. If α, β be scalars, then

$$\begin{aligned} F_{\alpha x+\beta y}(f) &= f(\alpha x + \beta y) \\ &= \alpha f(x) + \beta f(y) \\ &= \alpha F_x(f) + \beta F_y(f) \\ &= (\alpha F_x + \beta F_y)(f). \end{aligned}$$

Since this is true for all $f \in X^*$, we have

$$\begin{aligned} &F_{\alpha x+\beta y} = \alpha F_x + \beta F_y \\ \text{i.e.,} \qquad &C(\alpha x + \beta y) = \alpha C(x) + \beta C(y). \end{aligned}$$

This shows that C is linear. Now

$$\begin{aligned} \|C(x)\| = \|F_x\| &= \sup\left\{\frac{|F_x(f)|}{\|f\|} : \|f\| \neq 0\right\} \\ &= \sup\left\{\frac{|f(x)|}{\|f\|} : \|f\| \neq 0\right\} \\ &= \|x\|. \end{aligned}$$

This shows that C preserves norm. Now,

$$\begin{aligned} F_{x-y}(f) = f(x-y) &= f(x) - f(y) \\ &= F_x(f) - F_y(f) \\ &= (F_x - F_y)(f). \end{aligned}$$

This gives $F_{x-y} = F_x - F_y$ and hence

$$\|C(x) - C(y)\| = \|F_x - F_y\| = \|F_{x-y}\| = \|x - y\|.$$

This shows that if $x \neq y$ then $C(x) \neq C(y)$. Thus C is one-one. Therefore C is an isometric isomorphism between X and the range of C which is a subset of X^{**}. If the mapping C is onto, that is if the range of C is the whole of X^{**}, i.e., if $X = X^{**}$, then the space X is called reflexive.

Example 3.0.2. i) The space $\mathbb{R}_n$ is reflexive.

ii) The spaces l_p^n are reflexive for $1 \leq p < \infty$.

iii) The spaces l_p for $1 < p < \infty$ are reflexive.

iv) The spaces l_1, l_∞ and $C[a, b]$ are not reflexive.

Theorem 3.0.3. Each reflexive space X is a Banach space but not conversely.

Proof. We note that X^{**} is always complete. Since X is reflexive, X and X^{**} are isometrically isomorphic. Hence X is also complete. Thus X is a Banach space. □

The following example shows that the converse of the theorem need not necessarily true.

Example 3.0.4. Let $X = C_0 = \{x = (x_n) : x_n \to 0 \text{ as } n \to \infty\}$. Then $X^* = C_0^* = l_1$. Again $X^{**} = l_1^* = l_\infty$. Thus $X^{**} = l_\infty \neq C_0 = X$.
This shows that C_0, though a Banach space, is not reflexive. This proves the theorem.

Theorem 3.0.5. The conjugate space of a normed linear space remains the same under any equivalent norm.

Proof. Let X be a normed linear space and $\|.\|_1$, $\|.\|_2$ be two equivalent norms on X. Then there exists constants $a, b > 0$ such that

$$a\|x\|_2 \leq \|x\|_1 \leq b\|x\|_2, \forall x \in X.$$

Let X_1^*(resp. X_2^*) be the conjugate space of X with respect to the norm $\|.\|_1$(resp. $\|.\|_2$). Let $f \in X_1^*$. Then f is linear and bounded with respect to the norm $\|.\|_1$. So there is a constant $M(>0)$ such that

$$|f(x)| \leq M\|x\|_1 \leq Mb\|x\|_2.$$

This shows that f is bounded with respect to the norm $\|.\|_2$. That is $f \in X_2^*$ and hence $X_1^* \subset X_2^*$. Similarly, it can be shown that $X_2^* \subset X_1^*$. Thus we have $X_1^* = X_2^*$. This proves the theorem. □

Corollary 3.0.6. If X is reflexive then it remains reflexive under any equivalent norm.

Proof. Since the conjugate space of X remains same under any equivalent norm, the corollary follows. □

Theorem 3.0.7. Every closed subspace of a reflexive space is reflexive.

Proof. Let Y be a closed subspace of a reflexive space X and $C_Y : Y \longrightarrow Y^{**}$ be the cannonical mapping. We have to show that C_Y is surjective. Let $y^{**} \in Y^{**}$ and we define a mapping $x^{**} : X^* \longrightarrow K$ by $x^{**}(x^*) = y^{**}(x_y^*)$, where $x^* \in X^*$ and x_y^* is the restriction of x^* in Y. It can be easily verified that x^{**} is linear and bounded and hence $x^{**} \in X^{**}$. Since X is reflexive, the cannonical mapping $C : X \longrightarrow X^{**}$ is surjective. So there is an element $x \in X$ such that $C(x) = x^{**}$. We assert that $x \in Y$. Suppose, if possible, $x \notin Y$. Then $d = \inf\limits_{y \in Y} \|y - x\| > 0$, because Y is closed. So by an application of Hahn Banach theorem, there is a continuous linear functional $x_0^* \in X^*$ such that

$$x_0^*(x) = 1 \text{ and } x_0^* = 0 \text{ on } Y.$$

This gives

$$1 = x_0^*(x) = x^{**}(x_0^*) = y^{**}(0) = 0,$$

a contradiction. Therefore $x \in Y$ and $C_Y(x) = y^{**}$. This proves the theorem. □

Theorem 3.0.8. If X is reflexive then X^* is also reflexive.

Proof. Let $C^* : X^* \longrightarrow X^{***}$ be the cannonical mapping of X^* into X^{***}. We have to show that C^* is surjective. Let $x^{***} \in X^{***}$ be arbitrary. We define a functional x^* on X as follows:

$$x^*(x) = x^{***}(C(x))$$

where $x \in X$ and C is the cannonical mapping of X onto X^{**}. We first show that $x^* \in X^*$. Let $x, y \in X$ and λ be a scalar. Then

$$\begin{aligned}
x^*(x+y) &= x^{***}(C(x+y)) \\
&= x^{***}(C(x) + C(y)) \\
&= x^{***}(C(x)) + x^{***}(C(y)) \\
&= x^*(x) + x^*(y). \\
x^*(\lambda x) &= x^{***}(C(\lambda x)) \\
&= x^{***}(\lambda C(x)) \\
&= \lambda x^{***}(C(x)) \\
&= \lambda x^*(x).
\end{aligned}$$

Further,

$$\begin{aligned}|x^*(x)| &= |x^{***}(C(x))| \\ &\leq \|x^{***}\|\|C(x)\| \\ &= \|x^{***}\|\|x\| \text{ [since } \|C(x)\| = \|x\|].\end{aligned}$$

This shows that x^* is a bounded linear functional defined on X and so $x^* \in X^*$.

To prove the theorem we have to show that $C^*(x^*) = x^{***}$. Since $x^{**} \in X^{**}$ and the cannonical mapping $C : X \longrightarrow X^{**}$ is surjective, there is an element $x \in X$ such that $C(x) = x^{**}$. Now

$$\begin{aligned}C^*(x^*)(x^{***}) &= x^{**}(x^*) \\ &= C(x)(x^*) \\ &= x^*(x) \\ &= x^{***}(C(x)) \\ &= x^{***}(x^{**}).\end{aligned}$$

Since $x^{**} \in X^{**}$ is arbitrary, it follows that $C^*(x^*) = x^{***}$. This proves the theorem. □

Remark 3.0.9. The converse of Theorem 3.0.8 is also true. Let X^* be reflexive. Then by Theorem 3.0.5, X^{**} is reflexive. Since $C(X)$ is a closed linear subspace of X^{**}, by Theorem 3.0.7, it follows that $C(X)$ is reflexive. Hence X is reflexive as C is an isometric isomorphism of X onto $C(X)$.

Theorem 3.0.10. If X is a reflexive space and if X is separable then X^* is also separable.

Proof. To prove the theorem we need a lemma which we state and prove first.

Lemma:

If the conjugate space X^* of a normed linear space X is separable, then X is also separable.

Proof of the Lemma

Let $S = \{f : f \in X^*, \|f\| = 1\}$. Since every subspace of a separable metric space is separable, S is separable. Therefore, S contains a countable dense subset $D = \{f_1, f_2, \cdots, f_n, \cdots\}$ where $\|f_n\| = 1, \forall n$.

Since $\|f_n\| = \sup\{|f_n(x)| : \|x\| = 1\}$ for all n, there must exist some vectors x_n with $\|x_n\| = 1$ and $|f_n(x_n)| > \dfrac{1}{2}$.

Let M be the closed linear subspace of X generated by (x_n). We shall show that $M = X$. If possible, let $M \neq X$ and $x_0 \in X \setminus M$. Then there exist $f_0 \in X^*$ such that $\|f_0\| = 1, f_0(x_0) \neq 0$ and $f_0(x) = 0, \forall x \in M$. Therefore $f_0 \in S$ and $f_0(x_n) = 0, \forall n$. Hence,

$$\begin{aligned}\frac{1}{2} < |f_n(x_n)| &= |f_n(x_n) - f_0(x_n)| \\ &= |(f_n - f_0)x_n| \leq \|f_n - f_0\|\|x_n\| \\ &= \|f_n - f_0\| \text{ [since } \|x_n\| = 1].\end{aligned}$$

This shows that the open sphere $\|f_n - f_0\| < \frac{1}{2}$ centered at $f_0 \in S$ does not contain any point of D, contradicts the fact that D is dense in S. Hence $M = X$.

If X is a real normed linear space, then the set of all finite linear combinations of x_n's with rational coefficient is dense in X. Hence X is separable.

If X is a complex normed linear space, then the set of all finite linear combinations of x_n's whose coefficients have real and imaginary part as rational is dense in X. Consequently X is separable. This proves the lemma.

Proof of Theorem

Since X is reflexive, X is isometrically isomorphic with X^{**} and so X^{**} is separable. So by the above lemma, X^* is separable. This proves the theorem. □

Theorem 3.0.11. If X is a reflexive space then every bounded sequence in X has a weakly convergent subsequence.

Proof. Let (x_n) be a bounded sequence in X and let Y be the closure of the subspace generated by $x_1, x_2, \cdots$. So by Theorem 3.0.7, Y is reflexive. Clearly Y is separable and so by Theorem 3.0.10, Y^* is separable. Let C_y be the cannonical mapping of Y into Y^{**}. Since (x_n) is bounded in Y and C_y is an isometry, $(C_y(x_n))$ is bounded in Y^{**}. So, by a known result, [Result: If X is a separable Banach space then every bounded sequence $(f_n), f_n \in X^*$, contains a weakly convergent subsequence.] we obtain a sequence $(C_y(x_{n_j}))$ which converges weakly to some $y_0^{**} \in Y^{**}$. Since Y is reflexive, there exist $y_0 \in Y$ such that $y_0^{**} = C_y(y_0)$.

We wish to show that the subsequence (x_{n_j}) converges weakly to y_0. Let $x^* \in X^*$ and let y^* be the restriction of x^* on Y. Then

$$\begin{aligned} x^*(x_{n_j}) &= y^*(x_{n_j}) = C_y(x_{n_j})(y^*) \\ &\to y_0^{**}(y^*) = C_y(y_0)(y^*) = y^*(y_0) \\ &= x^*(y_0). \end{aligned}$$

This shows that $x_{n_j} \xrightarrow{w} y_0$. This proves the theorem. □

Theorem 3.0.12. For a sequence (f_n) in X^* and $f \in X^*$, we have

i) $f_n \to f \Rightarrow f_n \xrightarrow{w} f \Rightarrow f_n \xrightarrow{w*} f$.

ii) If X is reflexive then $f_n \xrightarrow{w} f$ if and only if $f_n \xrightarrow{w*} f$.

Proof. i) Since strong convergence implies weak convergence, $f_n \to f$ implies $f_n \xrightarrow{w} f$.

Let $x \in X$. Let $C : X \longrightarrow X^{**}$ be the cannonical mapping. Then $C(x) = F_x$ provided that $F_x(f) = f(x) \;\; \forall f \in X^*$. Let $f_n \xrightarrow{w} f$. Then $F(f_n) \to F(f)$ for all $F \in X^{**}$. Hence $F_x(f_n) \to F_x(f)$. Since $F_x(f_n) = f_n(x)$ and $F_x(f) = f(x)$, we obtain $f_n(x) \to f(x)$ for every $x \in X$. Thus $f_n \xrightarrow{w*} f$.

ii) Let $C : X \longrightarrow X^{**}$ be the cannonical mapping of X into X^{**}. Since X is reflexive, C is onto. Let $f_n \xrightarrow{w*} f$. Then $f_n(x) \to f(x)$ for all $x \in X$. Since $F(f_n) = f_n(x)$ and $F(f) = f(x)$ for all $F \in X^{**}$, $f_n(x) \to f(x)$ implies $F(f_n) \to F(f) \;\; \forall F \in X^{**}$. This gives $f_n \xrightarrow{w} f$. From (i), $f_n \xrightarrow{w} f \implies f_n \xrightarrow{w*} f$. This proves the theorem.

□

Example 3.0.13. The following example shows that the weak* convergence does not necessarily imply weak convergence.

Solution. Consider the space C_0 of null sequences. We know that $C_0^* = l_1$. Let $f_n = (0, 0, \cdots, 0, 1, 0 \cdots)$ where 1 is in the nth place be the nth coordinate functional defined on C_0. If $x = (x_n) \in C_0$, then $f_n(x) = x_n$ and $x_n \to 0$ and $n \to \infty$. Therefore $f_n \xrightarrow{w*} 0$. We now show that (f_n) does not converge weakly to zero. For let $F = (1, 1, \cdots) \in l_1^* = l_\infty$. Now $F(f_n) = 1$ for all n and $F(0) = 0$. Hence (f_n) does not converge weakly to 0. This completes the solution. ■

3.0.1 Reflexivity of a Hilbert space

Let H be a Hilbert space and H^* be the conjugate space of H. We define a mapping $T : H \longrightarrow H^*$ by $T(y) = f$ where for $x \in H$, $f(x) = (x, y)$. So for all $x \in H$, we have

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

Now, if $y_1, y_2 \in H$, then

$$\begin{aligned}
(T(y_1 + y_2))(x) &= (x, y_1 + y_2) \\
&= (x, y_1) + (x, y_2) \\
&= (Ty_1)(x) + (Ty_2)(x) \\
&= (Ty_1 + Ty_2)(x).
\end{aligned}$$

Thus T is additive.

For any scalar α,

$$(T(\alpha y))(x) = (x, \alpha y) = \overline{\alpha}(x, y) = \overline{\alpha}(Ty)(x).$$

This shows that $T(\alpha y) = \overline{\alpha} T(y)$, that is T is conjugate linear.

Further, if $f \in H^*$, then Riesz Representation theorem provide us a unique $y \in H$ such that for all $x \in H$, $f(x) = (x, y)$. Moreover,

$$\|y\| = \|f\| = \|Ty\|.$$

It follows, therefore, that T is a one-one, onto, isometric and conjugate linear mapping from H into H^*.

Theorem 3.0.14. If H is a Hilbert space then H is reflexive.

Proof. Let $C : H \longrightarrow H^{**}$ be the cannonical mapping defined by $C(x) = h$ iff $h(f) = f(x)$ for all $f \in H^*$ where $x \in H$ and $h \in H^{**}$. We have to show that C is surjective.

Let $f_1 \in H^{**}$ and we should find an element $z \in H$ such that $C(z) = f_1$. Let $T : H \longrightarrow H^*$ be defined by $T(y) = f$ where for $x \in H$, $f(x) = (x, y)$ and so for all $x \in H$ we have

$$(Ty)(x) = f(x) = (x, y). \tag{3.0.1}$$

Let g be a functional defined on H by $g(x) = \overline{f_1(Tx)}$.

Using the conjugate linearity of T, it can be shown that g is linear. As T is isometric,

$$\begin{aligned}
|g(x)| &= |\overline{f_1(Tx)}| = |f_1(Tx)| \\
&\leq \|f_1\| \|Tx\| = \|f_1\| \|x\|.
\end{aligned}$$

Hence g is bounded and therefore $g \in H^*$.

So by Riesz Representation theorem, there exist $z \in H$ such that for all $x \in H, g(x) = (x, z)$. So,

$$\begin{aligned} \overline{f_1(Tx)} &= (x, z) \\ \text{i.e., } f_1(Tx) &= (z, x). \end{aligned}$$

Again by the definition of T we have

$$f_1(Tx) = (Tx)(z) \quad \text{[by (3.0.1)]} \tag{3.0.2}$$

for any $Tx \in H^*$. Since T is surjective, any element of H^* may be written in the form Tx, the relation (3.0.2) gives that $C(z) = f_1$ and C becomes surjective. Therefore H is reflexive. This proves the theorem. $\square$

Unit 4

Properties of Operators - I

Course Structure

- Bounded linear operator, uniqueness theorem, adjoint of an operator and its properties.

Definition 4.0.1. Let H be a Hilbert space and let $A : H \longrightarrow H$ be a continuous linear operator. For $y \in H$, define a functional f_y on H by
$$f_y(x) = (Ax, y). \tag{4.0.1}$$
Then
$$\begin{aligned} f_y(x_1 + x_2) &= (A(x_1 + x_2), y) \\ &= (Ax_1 + Ax_2, y) \\ &= (Ax_1, y) + (Ax_2, y) \\ &= f_y(x_1) + f_y(x_2). \end{aligned}$$
If λ is a scalar, then
$$f_y(\lambda x) = (A\lambda x, y) = \lambda(Ax, y) = \lambda f_y(x).$$
Moreover, for $x \in H$
$$|f_y(x)| = |(Ax, y)| \leq \|Ax\|\|y\| \leq \|A\|\|x\|\|y\|.$$
Therefore, f_y is a continuous linear functional defined on H and $\|f_y\| \leq \|A\|\|y\|$.

Hence by Riesz Representation theorem f_y has the form
$$f_y(x) = (x, y^*) \tag{4.0.2}$$
for all $x \in H$ and $y^* \in H$ is uniquely determined by f_y. If y is changed then f_y is changed and so y^* is also changed. Thus we obtain an operator $A^* : H \longrightarrow H$ such that $y^* = A^*y$.

This operator A^* is called the adjoint operator of A. From (4.0.1) and (4.0.2) we see that A and A^* are connected by the relation
$$(Ax, y) = (x, A^*y).$$
We note that
$$(A^*x, y) = \overline{(y, A^*x)} = \overline{(Ay, x)} = (x, Ay).$$

4.0.1 Some Properties of Adjoint operators

i) The definition of A^* is unique.

Proof. For all $x, y \in H$,

$$(Ax, y) = (x, A^*y) \quad \text{and} \quad (Ax, y) = (x, A_1^*y).$$

Then

$$\begin{aligned} (x, A^*y) - (x, A_1^*y) &= 0 \\ \text{i.e., } (x, A^*y - A_1^*y) &= 0. \end{aligned}$$

Therefore,

$$\begin{aligned} A^*y - A_1^*y &= 0 \text{ for all } y \in H \\ \text{i.e., } A^* &= A_1^*. \end{aligned}$$

□

ii) A^* is a continuous linear operator with

$$\|A^*\| \leq \|A\|.$$

Proof. For $x, y, z \in H$ we have

$$\begin{aligned} (x, A^*(y+z)) &= (Ax, y+z) \\ &= (Ax, y) + (Ax, z) \\ &= (x, A^*y) + (x, A^*z) \\ &= (x, A^*y + A^*z) \end{aligned}$$

So,

$$A^*(y+z) = A^*y + A^*z.$$

If λ is a scalar, then

$$\begin{aligned} (x, A^*\lambda y) &= (Ax, \lambda y) \\ &= \overline{\lambda}(Ax, y) \\ &= \overline{\lambda}(x, A^*y) \\ &= (x, \lambda A^*y). \end{aligned}$$

So,

$$A^*\lambda y = \lambda A^*y$$

This shows that A^* is linear.

Now by Cauchy Schwarz inequality we see that for all $y \in H$

$$\begin{aligned} \|A^*y\|^* &= (A^*y, A^*y) \\ &= (AA^*y, y) \\ &\leq \|AA^*y\|\|y\| \\ &\leq \|A\|\|A^*y\|\|y\| \\ \text{i.e., } \|A^*y\| &\leq \|A\|\|y\| \text{ for all } y \in H. \end{aligned}$$

Therefore A^* is a continuous linear operator with

$$\|A^*\| \leq \|A\|.$$

□

iii) $A^{**} = A$.

Proof. Since A^* is continuous linear operator, A^{**} is defined. In the relation

$$(Ax, y) = (x, A^*y)$$

we replace A by A^* and obtain

$$(A^*x, y) = (x, A^{**}y).$$

Interchanging x and y we get

$$(A^*y, x) = (y, A^{**}x).$$

Taking conjugate we get

$$(A^{**}x, y) = (x, A^*y) = (Ax, y).$$

Thus $A^{**}x = Ax, \forall x \in H$ and so $A^{**} = A$. □

iv) $\|A^*\| = \|A\|$.

Proof. For any continuous linear operator $T : H \longrightarrow H$ we have by property (ii) above,

$$\|T^*\| \leq \|T\|.$$

Putting $T = A^*$, we get

$$\begin{aligned} \|A^{**}\| &\leq \|A^*\| \\ \text{i.e., } \|A\| &\leq \|A^*\|. \end{aligned}$$

Hence $\|A^*\| = \|A\|$. □

v) If $A_1 : H \longrightarrow H$ and $A_2 : H \longrightarrow H$ are continuous linear operators, then $(A_1A_2)^* = A_2^*A_1^*$.

Proof. We note that if $A_1 : H \longrightarrow H$ and $A_2 : H \longrightarrow H$ are continuous linear operators, then $A_1A_2 : H \longrightarrow H$ is also a continuous linear operator. Now for $x, y \in H$ we have

$$\begin{aligned} (x, (A_1A_2)^*y) &= (A_1A_2x, y) \\ &= (A_2x, A_1^*y) \\ &= (x, A_2^*A_1^*y). \\ \text{So, } (A_1A_2)^*y &= A_2^*A_1^*y, \forall y \in H. \\ \text{Thus } (A_1A_2)^* &= A_2^*A_1^*. \end{aligned}$$

□

vi) $\|A^*A\| = \|A\|^2 = \|AA^*\|$.

Proof. We always have

$$\|A^*A\| < \|A^*\|\|A\| = \|A\|^2. \tag{4.0.3}$$

Now

$$\begin{aligned} \|Ax\|^2 &= (Ax, Ax) \\ &= (A^*Ax, x) \\ &\leq \|A^*Ax\|\|x\| \\ &\leq \|A^*A\|\|x\|^2. \end{aligned}$$

So if $\|x\| \leq 1$, then

$$\|Ax\|^2 \leq \|A^*A\|.$$

Therefore

$$\|A\|^2 = \sup\{\|Ax\|^2 : \|x\| \leq 1\} \leq \|A^*A\|. \tag{4.0.4}$$

From (4.0.3) and (4.0.4) we get

$$\|A^*A\| = \|A\|^2.$$

Again

$$\begin{aligned} \|AA^*\| &= \|A^{**}A^*\| \\ &= \|(A^*)^*A^*\| \\ &= \|A^*\|^2 \\ &= \|A\|^2. \end{aligned}$$

□

vii) $(A + B)^* = A^* + B^*$

Proof. For all $x, y \in H$ we have

$$\begin{aligned} ((A + B)^*x, y) &= (x, (A + B)y) \\ &= (x, Ay + By) \\ &= (x, Ay) + (x, By) \\ &= (A^*x, y) + (B^*x, y) \\ &= (A^*x + B^*x, y). \end{aligned}$$

Since this is true for all $y \in H$, we have

$$(A + B)^*x = A^*x + B^*x = (A^* + B^*)x, \forall x \in H.$$

Hence, $(A + B)^* = A^* + B^*$. □

viii) For any scalar λ, $(\lambda A)^* = \overline{\lambda} A^*$.

Proof. For all $x, y \in H$ we have

$$\begin{aligned} ((\lambda A)^* x, y) &= (x, (\lambda A) y) \\ &= (x, \lambda A y) \\ &= \overline{\lambda}(x, Ay) \\ &= \overline{\lambda}(A^* x, y) \\ &= (\overline{\lambda} A^* x, y). \end{aligned}$$

Since this is true for all $y \in H$ we have

$$\begin{aligned} (\lambda A)^* x &= \lambda A^* x, \forall x \in H \\ \text{and so } (\lambda A)^* &= \overline{\lambda} A^*. \end{aligned}$$

□

4.0.2 Self Adjoint Operator

Let H be a Hilbert space. A continuous linear operator $A : H \longrightarrow H$ is called self-adjoint if $A^* = A$.

Theorem 4.0.2. If

a) A is self-adjoint,

b) $(Ax, y) = (x, Ay), \forall x, y \in H$,

c) $(Ax, x) = (x, Ax), \forall x \in H$,

d) (Ax, x) is real $\forall x \in H$,

then (a)⇒(b)⇒(c)⇒(d).

Proof. Let A is self-adjoint. Then $A = A^*$. So for all $x, y \in H$

$$(Ax, y) = (x, A^* y) = (x, Ay).$$

Thus (a) ⇒ (b).

Since $(Ax, y) = (x, Ay)$ for all $x, y \in H$, taking $y = x$ we get

$$(Ax, x) = (x, Ax) \quad \forall x \in H.$$

Thus (b) ⇒ (c).

Again

$$\begin{aligned} (Ax, x) &= (x, Ax) \ \forall x \in H \\ \Rightarrow (Ax, x) &= \overline{(Ax, x)} \\ \Rightarrow (Ax, x) &\quad \text{is real} \quad . \end{aligned}$$

Thus (c) ⇒ (d). This proves the theorem. □

Theorem 4.0.3. i) If A and B are self-adjoint then so is also $A + B$.

ii) For any continuous linear operator A, the operators $A^*A, AA^*, A + A^*$ are self-adjoint.

iii) If A is self-adjoint and α is real constant then αA is self-adjoint.

iv) If A and B are self-adjoint, then AB is self-adjoint if and only if $AB = BA$.

Proof. Since A and B are self-adjoint, we have

$$A^* = A \text{ and } B^* = B.$$

i) Hence $(A + B)^* = A^* + B^* = A + B$. Thus, $A + B$ is self-adjoint.

ii)

$$\begin{aligned}
&(A^*A)^* = A^*A^{**} = A^*A,\\
&(AA^*)^* = A^{**}A^* = AA^*,\\
&(A + A^*)^* = A^* + A^{**} = A^* + A = A + A^*
\end{aligned}$$

Therefore, A^*A, AA^* and $A + A^*$ are self-adjoint.

iii)

$$(\alpha A)^* = \overline{\alpha}A^* = \alpha A^* \text{ (since } \alpha \text{ is real).}$$

Hence αA is self-adjoint.

iv)

$$(AB)^* = B^*A^* = BA$$

From this it follows that AB is self-adjoint if and only if $AB = BA$. This proves the theorem.

□

Theorem 4.0.4. The collection of all self-adjoint operators form a closed real linear subspace of the space of all continuous linear operators that map H into itself.

Proof. It is clear that $0^* = 0$ and $I^* = I$ where 0 and I denote respectively the zero and the identity operators. Let $\alpha_i (i = 1, 2, \cdots, n)$ are real and $A_i (i = 1, 2, \cdots, n)$ are self-adjoint operators. Then

$$\begin{aligned}
(\alpha_1 A_1 + \alpha_2 A_2 + \cdots + \alpha_n A_n)^* &= \overline{\alpha_1}A_1^* + \overline{\alpha_2}A_2^* + \cdots + \overline{\alpha_n}A_n^*\\
&= \alpha_1 A_1 + \alpha_2 A_2 + \cdots + \alpha_n A_n.
\end{aligned}$$

This shows that $\alpha_1 A_1 + \alpha_2 A_2 + \cdots + \alpha_n A_n$ is self-adjoint.

Now assume that (A_n) is a sequence of self-adjoint operators converging in norm to a continuous linear operator A. Then

$$\begin{aligned}
\|A^* - A\| &\leq \|A^* - A_n^*\| + \|A_n^* - A_n\| + \|A_n - A\|\\
&= \|(A - A_n)^*\| + \|A_n - A\|\\
&= 2\|A_n - A\| \to 0 \text{ as } n \to \infty.
\end{aligned}$$

Hence $A^* = A$. Thus A is self-adjoint. This proves the theorem. □

Theorem 4.0.5. Let $A : H \longrightarrow H$ be a continuous linear operator. Then $(Ax, y) = 0$ for all $x, y \in H$ if and only if $A = 0$, the zero operator.

Proof. Let $A = 0$, the zero operator. Then

$$(Ax, y) = (\theta, y) = 0, \forall x, y \in H.$$

Conversely, if $(Ax, y) = 0, \forall x, y \in H$, then choosing a fixed x, we see that

$$(Ax, y) = 0, \forall y \in H$$

and so $Ax = \theta$. This is true for any $x \in H$ and so $A = 0$, the zero operator. This proves the theorem. □

Theorem 4.0.6. Let $A : H \longrightarrow H$ be a continuous linear operator. If $(Ax, x) = 0$ for all $x \in H$, then $A = 0$, the zero operator.

Proof. It is sufficient to show that $(Ax, y) = 0.\forall x, y \in H$. For arbitrary scalars α and β we have

$$\begin{aligned} 0 &= (A(\alpha x + \beta y), \alpha x + \beta y) \\ &= (\alpha Ax + \beta Ay, \alpha x + \beta y) \\ &= (\alpha Ax, \alpha x) + (\alpha Ax, \beta y) + (\beta Ay, \alpha x) + (\beta Ay, \beta y) \\ &= |\alpha|^2(Ax, x) + \alpha\overline{\beta}(Ax, y) + \overline{\alpha}\beta(Ay, x) + |\beta|^2(Ay, y) \\ &= \alpha\overline{\beta}(Ax, y) + \overline{\alpha}\beta(Ay, x). \end{aligned} \tag{4.0.5}$$

Putting $\alpha = \beta = 1$ in (4.0.5) we get

$$(Ax, y) + (Ay, x) = 0. \tag{4.0.6}$$

Putting $\alpha = i$ and $\beta = 1$ in (4.0.5) we get

$$\begin{aligned} & i(Ax, y) - i(Ay, x) = 0 \\ \text{i.e.,} \quad & (Ax, y) - (Ay, x) = 0. \end{aligned} \tag{4.0.7}$$

From (4.0.6) and (4.0.7) we obtain

$$(Ax, y) = 0, \forall x, y \in H.$$

This proves the theorem. □

Theorem 4.0.7. A continuous linear operator $A : H \longrightarrow H$ is self-adjoint if and only if (Ax, x) is real for all $x \in H$.

Proof. First we suppose that A is self-adjoint. Then $A^* = A$. Now

$$\overline{(Ax, x)} = (x, Ax) = (A^*x, x) = (Ax, x).$$

Hence (Ax, x) is real for all x in H.

Next we assume that (Ax, x) is real for all x in H. Then,

$$\begin{aligned} & (Ax, x) = \overline{(Ax, x)} = \overline{(x, A^*x)} = (A^*x, x) \\ \text{i.e.,} \quad & (Ax - A^*x, x) = 0, \forall x \in H \\ \text{i.e.,} \quad & ((A - A^*)x, x) = 0, \forall x \in H \\ \text{i.e.,} \quad & A - A^* = 0, \text{ the zero operator.} \\ \text{i.e.,} \quad & A = A^*. \end{aligned}$$

Thus A is self-adjoint. This proves the theorem. □

Theorem 4.0.8. Suppose that $A : H \longrightarrow H$ is self-adjoint operator. Then

$$\|A\| = \sup\{|(Ax, x)| : \|x\| = 1\}.$$

Proof. We write $S_A = \sup\{|(Ax, x)| : \|x\| = 1\}$. If $\|x\| = 1$, then

$$|(Ax, x)| \leq \|Ax\|\|x\| \leq \|A\|\|x\|^2 = \|A\|$$

and hence

$$S_A \leq \|A\|. \tag{4.0.8}$$

On the other hand, if $\|y\| = 1$ then clearly

$$|(Ay, y)| \leq S_A.\|y\|^2.$$

If $\|y\| \neq 1$ and $y \neq \theta$, let $y' = \dfrac{y}{\|y\|}$. Then $\|y'\| = 1$ and

$$\begin{aligned} & |(Ay', y')| \leq S_A \\ \text{i.e.,}\qquad & |(Ay, y)| \leq S_A.\|y\|^2. \end{aligned} \tag{4.0.9}$$

The inequality (4.0.9) also holds if $y = \theta$. If $z \in H$ and $z \neq \theta$, we put

$$\lambda = \left(\frac{\|Az\|}{\|z\|}\right)^{\frac{1}{2}} \text{ and } u = \frac{1}{\lambda}Az.$$

Then

$$\begin{aligned} (A(\lambda z + u), \lambda z + u) &= (\lambda Az + Au, \lambda z + u) \\ &= |\lambda|^2(Az, z) + \lambda(Az, u) + \lambda(Au, z) + (Au, u) \\ (\text{since } \lambda \text{ is real}). & \\ \text{Also, } (A(\lambda z - u), \lambda z - u) &= (\lambda Az - Au, \lambda z - u) \\ &= |\lambda|^2(Az, z) - \lambda(Az, u) - \lambda(Au, z) + (Au, u). \\ \text{So, } (A(\lambda z + u), \lambda z + u) - (A(\lambda z - u), \lambda z - u) &= 2\lambda[(Az, u) + (Au, z)] \\ &= 2\lambda\left[\left(Az, \frac{1}{\lambda}Az\right) + \left(\frac{1}{\lambda}Az, Az\right)\right] \; [\text{since } A^* = A] \\ &= 4\|Az\|^2. \end{aligned}$$

Thus,

$$\begin{aligned} \|Az\|^2 &= \frac{1}{4}[(A(\lambda z + u), \lambda z + u) - (A(\lambda z - u), \lambda z - u)] \\ &\leq \frac{1}{4}S_A[\|\lambda z + u\|^2 + \|\lambda z - u\|^2] \\ &\quad [\text{ since } (Ay, y) \leq S_A\|y\|^2 \text{ and } -(Ay, y) \leq S_A\|y\|^2] \\ &= \frac{1}{2}S_A[\|\lambda z\|^2 + \|u\|^2], \text{ by Parallelogram law} \\ &= \frac{1}{2}S_A\left[\frac{\|Az\|}{\|z\|}\|z\|^2 + \frac{\|z\|}{\|Az\|}\|Az\|^2\right] \\ &= S_A\|Az\|\|z\| \\ \text{i.e., } \|Az\| &\leq S_A\|z\| \end{aligned}$$

and consequently

$$\|A\| \leq S_A. \tag{4.0.10}$$

From (4.0.8) and (4.0.10) we obtain

$$S_A = \sup\{|(Ax, x)| : \|x\| = 1\} = \|A\|.$$

This proves the theorem. □

Theorem 4.0.9. Let $A : H \longrightarrow H$ be a bounded linear operator. Then the following statements are equivalent.

i) $A^*A = I$, the identity operator.

ii) $(Ax, Ay) = (x, y), \forall x, y \in H$.

iii) $\|Ax\| = \|x\|, \forall x \in H$.

Proof. (i) ⇒ (ii)

$$(Ax, Ay) = (A^*Ax, y) = (Ix, y) = (x, y).$$

(ii) ⇒ (iii)
By (ii) we have

$$\begin{aligned} & (Ax, Ay) = (x, y), \forall x, y \in H \\ \text{i.e.,} \quad & (Ax, Ax) = (x, x), \forall x \in H \\ \text{i.e.,} \quad & \|Ax\|^2 = \|x\|^2, \forall x \in H \\ \text{i.e.,} \quad & \|Ax\| = \|x\|, \forall x \in H. \end{aligned}$$

(iii) ⇒ (i)

$$\begin{aligned} & \|Ax\| = \|x\|, \forall x \in H \\ \text{i.e.,} \quad & (Ax, Ax) = (x, x), \forall x \in H \\ \text{i.e.,} \quad & (A^*Ax, x) - (x, x) = 0, \forall x \in H \\ \text{i.e.,} \quad & ((A^*A - I)x, x) = 0, \forall x \in H. \end{aligned}$$

This implies that

$$\begin{aligned} & A^*A - I = 0, \ \text{ the zero operator} \\ \text{i.e.,} \quad & A^*A = I. \end{aligned}$$

This proves the theorem. □

Definition 4.0.10. Let T be an operator on a Hilbert space H. A scalar λ is called an eigen value of T if there exists a non-zero vector x in H such that $Tx = \lambda x$.

If λ is an eigen value of T, then any non-zero vector x in H that satisfies $Tx = \lambda x$ is called an eigen vector of T.

Note 4.0.11. i) Corresponding to a single eigen value of T, there may correspond more than one eigen vector.

ii) If x is an eigen vector of T, then x cannot correspond more than one eigen value of T.

Theorem 4.0.12. Let $T : H \longrightarrow H$ be a self-adjoint operator. Then

i) all eigen values of T, if exist, are real;

ii) eigen vectors corresponding to distinct eigen values of T are orthogonal.

Proof. i) Let λ be an eigen value and $x(\neq 0)$ be a corresponding eigen vector of T. Then $Tx = \lambda x$. Since T is self-adjoint, we have $T^* = T$. Now

$$\begin{aligned} & \lambda(x,x) = (\lambda x, x) = (Tx, x) = (x, Tx) = (x, \lambda x) = \overline{\lambda}(x,x) \\ \text{i.e.,}\quad & (\lambda - \overline{\lambda})(x,x) = 0 \\ \text{i.e.,}\quad & \lambda = \overline{\lambda} \text{ [since } (x,x) \neq 0]. \end{aligned}$$

Thus λ is real.

ii) Let x and y be two eigen vectors corresponding to distinct eigen values λ and μ respectively. Then $Tx = \lambda x$ and $Ty = \mu y$. Now,

$$\begin{aligned} & \lambda(x,y) = (\lambda x, y) = (Tx, y) = (x, Ty) = (x, \mu y) = \mu(x,y) \\ \text{i.e.,}\quad & (\lambda - \mu)(x,y) = 0 \\ \text{i.e.,}\quad & (x,y) = 0 \text{ [since } \lambda \neq \mu]. \end{aligned}$$

This proves the theorem.

□

Theorem 4.0.13. Let $T : H \longrightarrow H$ be any continuous linear operator. Then T can be expressed uniquely in the form $T = A + iB$ where A and B are self-adjoint operators.

Proof. Let $A = \frac{1}{2}(T + T^*), B = \frac{1}{2i}(T - T^*)$. Then

$$\begin{aligned} A^* &= \frac{1}{2}(T^* + T) = A \text{ and} \\ B^* &= -\frac{1}{2i}(T^* - T) = B \end{aligned}$$

So that A and B are self-adjoint and $A + iB = T$.

If $T = C + iD$ where C and D are self-adjoint, then

$$T^* = C^* - iD^* = C - iD.$$

Therefore $T + T^* = 2C$ and $T - T^* = 2iD$. Thus $C = A$ and $D = B$. This proves the theorem. □

Theorem 4.0.14. Let $T : H \longrightarrow H$ be a bounded linear operator such that $T^*T = TT^*$. Then $Tx = \lambda x$ if and only if $T^*x = \overline{\lambda}x$ for $x \in H$ and λ is a scalar.

Proof. Consider the operator $T - \lambda I$ where I is the identity operator. Then

$$\begin{aligned} (T - \lambda I)(T - \lambda I)^* &= (T - \lambda I)(T^* - \overline{\lambda}I) \\ &= TT^* - \overline{\lambda}T - \lambda T^* + |\lambda|^2 I \\ \text{and } (T - \lambda I)^*(T - \lambda I) &= (T^* - \overline{\lambda}I)(T - \lambda I) \\ &= T^*T - \lambda T^* - \overline{\lambda}T + |\lambda|^2 I. \end{aligned}$$

Since $T^*T = TT^*$, we have

$$\begin{aligned} & (T-\lambda I)(T-\lambda I)^* = (T-\lambda I)^*(T-\lambda I) \\ \text{i.e.,}\quad & SS^* = S^*S, \text{ where } S = T-\lambda I. \end{aligned}$$

Therefore,

$$\begin{aligned} & SS^*(x) = S^*S(x), \forall x \in H \\ \text{i.e.,}\quad & (SS^*(x), x) = (S^*S(x), x) \\ \text{i.e.,}\quad & (S^*(x), S^*(x)) = (S(x), S(x)) \\ \text{i.e.,}\quad & \|S^*(x)\|^2 = \|S(x)\|^2 \\ \text{i.e.,}\quad & \|S^*(x)\| = \|S(x)\| \\ \text{i.e.,}\quad & \|(T^* - \overline{\lambda} I)x\| = \|(T-\lambda I)(x)\| \\ \text{i.e.,}\quad & \|T^*x - \overline{\lambda}x\| = \|Tx - \lambda x\|. \end{aligned}$$

This shows that $Tx = \lambda x$ if and only if $T^*x = \overline{\lambda}x$. This proves the theorem. $\square$

Unit 5

Properties of Operators - II

Course Structure

- Self-adjoint, compact, normal, unitary and positive operators, norm of self -adjoint operator, group of unitary operator, square root of positive operator-characterization and basic properties

Definition 5.0.1 (Completely Continuous Operators)**.** A linear operator T mapping a Hilbert space H into a Hilbert space H_1(or a Banach space B into another such space B_1) is called a completely continuous operator if given any sequence (x_n) in H such that $(\|x_n\|)$ is bounded, the sequence (Tx_n) has a convergent subsequence.

It is clear that an operator T is completely continuous if and only if $\|x_n\| \leq 1$ implies that (Tx_n) has a convergent subsequence because if the sequence (y_n) be such that $\|y_n\| \leq M$, then we may take $x_n = \dfrac{y_n}{M}$ and in that case $\|x_n\| \leq 1$.

Note 5.0.2. A completely continuous operator is sometimes called a compact operator.

Note 5.0.3. The zero operator is completely continuous.

Theorem 5.0.4. A completely continuous operator is continuous.

Proof. Let T be a completely continuous operator. Then T is linear. We show that T is bounded. If possible, let T be not bounded. Then there is a sequence (x_n) such that

$$\|Tx_n\| > n\|x_n\| \quad \text{for all } n.$$

Let $y_n = \dfrac{x_n}{\|x_n\|}, \forall n$. Then $\|y_n\| = 1$. Now

$$T(y_n) = T\left(\frac{x_n}{\|x_n\|}\right) = \frac{1}{\|x_n\|}Tx_n > n, \forall n.$$

Therefore the sequence (Ty_n) cannot have any convergent subsequence which contradicts the fact that T is completely continuous. Hence T is bounded and so T is continuous. This proves the theorem. $\square$

Example 5.0.5. Let H be a Hilbert space. If $y, z \in H$ are fixed, then the operator $T : H \longrightarrow H$ defined by

$$T(x) = (x, y)z$$

is completely continuous.

Solution. We have

$$\begin{aligned} T(x_1 + x_2) &= (x_1 + x_2, y)z \\ &= (x_1, y)z + (x_2, y)z \\ &= Tx_1 + Tx_2 \\ \text{and} \quad T(\lambda x) &= (\lambda x, y)z = \lambda(x, y)z = \lambda Tx. \end{aligned}$$

Hence T is linear. Let (α_n) be a sequence of elements from H such that $\|\alpha_n\| \leq 1$. Then

$$|(\alpha_n, y)| \leq \|\alpha_n\|\|y\| \leq \|y\|$$

and so the sequence $((\alpha_n, y))$ has a convergent subsequence $((\alpha_{n_k}, y))$ that converges to α, say. But in that case

$$T(\alpha_{n_k}) = (\alpha_{n_k}, y)z \to \alpha z \quad \text{as } k \to \infty$$

and so T becomes completely continuous. ∎

Example 5.0.6. Let H be a Hilbert space. The identity operator I is not completely continuous although it is continuous.

Solution. Let (x_n) be a sequence of elements in H such that $x_n \to x$ as $n \to \infty$. Then

$$Ix_n = x_n \to x = Ix \quad \text{as } n \to \infty,$$

showing that the identity operator I is continuous.

Let (x_n) be an orthonormal sequence. Then

$$\begin{aligned} (x_n, x_m) &= 1, \ \text{if } n = m \\ &= 0, \ \text{if } n \neq m. \end{aligned}$$

so that $\|x_n\| = 1$ for all n. If $n \neq m$, then

$$\begin{aligned} \|Ix_n - Ix_m\|^2 &= \|x_n - x_m\|^2 \\ &= (x_n - x_m, x_n - x_m) \\ &= 2 \\ \text{i.e., } \|Ix_n - Ix_m\| &= \sqrt{2}, \quad \text{if } n \neq m. \end{aligned}$$

Therefore (Ix_n) cannot have any subsequence which is Cauchy and hence it cannot have any convergent subsequence. Therefore I is not completely continuous. ∎

Theorem 5.0.7. If T is completely continuous and λ is a scalar then λT is completely continuous.

Proof. Let (x_n) be a sequence of elements from H such that $\|x_n\| < 1$. Since T is completely continuous, there is a subsequence (x_{n_k}) of (x_n) such that (Tx_{n_k}) is convergent. Let

$$Tx_{n_k} \to u, \quad \text{as } k \to \infty.$$

In that case,

$$\lim_{k\to\infty} (\lambda T)x_{n_k} = \lambda \lim_{k\to\infty} Tx_{n_k} = \lambda u,$$

which shows that λT is completely continuous. This proves the theorem. □

Theorem 5.0.8. If S and T are completely continuous operators then $S + T$ is also completely continuous.

Proof. Let (x_n) be a sequence of elements from H such that $\|x_n\| < 1$. Since S is completely continuous, there is a subsequence (x_{n_k}) of (x_n) such that (Sx_{n_k}) is convergent. Let

$$\lim_{k\to\infty} Sx_{n_k} = u.$$

Again, since T is completely continuous, there is a subsequence $(x_{n_{k_l}})$ of (x_{n_k}) such that $(Tx_{n_{k_l}})$ is convergent. Let

$$\lim_{l\to\infty} Tx_{n_{k_l}} = v.$$

It is clear that

$$\lim_{l\to\infty} Sx_{n_{k_l}} = u.$$

Therefore,

$$\begin{aligned} \lim_{l\to\infty} (S+T)x_{n_{k_l}} &= \lim_{l\to\infty} S_{n_{k_l}} + \lim_{l\to\infty} T_{n_{k_l}} \\ &= u + v. \end{aligned}$$

This shows that $S + T$ is completely continuous. This proves the theorem. □

Theorem 5.0.9. If T is a completely continuous operator and S is a continuous linear operator then both TS and ST are completely continuous where S and T map H(or B) into itself.

Proof. Clearly both TS and ST are continuous linear operators. Let (x_n) be a sequence of elements from H such that $\|x_n\| < 1$. Since T is completely continuous, there is a subsequence (x_{n_p}) of (x_n) such that (Tx_{n_p}) is convergent. Let

$$Tx_{n_p} \to u \quad \text{as } p \to \infty.$$

Since S is continuous, we have

$$S(Tx_{n_p}) = STx_{n_p} \to Su \quad \text{as } p \to \infty,$$

showing that ST is completely continuous.

In order to show that TS is completely continuous, we first observe that

$$\|Sx_n\| \leq \|S\|\|x_n\| \leq \|S\|.$$

That means the sequence (Sx_n) is bounded in H(or B). Since T is completely continuous, there is a subsequence (Sx_{n_k}) of (Sx_n) such that (TSx_{n_k}) is convergent. This shows that TS is completely continuous. This proves the theorem. □

Theorem 5.0.10. If (T_n) is a sequence of completely continuous operators in a Hilbert space H(or a Banach space B) into a Hilbert space H_1(or a Banach space B_1) and if $T : H \longrightarrow H_1$(or $T : B \longrightarrow B_1$) is a bounded linear operator such that $\|T - T_n\| \to 0$ as $n \to \infty$, then T is completely continuous operator.

Proof. The proof of this theorem is beyond the scope of this study material. □

Theorem 5.0.11. Let H be a Hilbert space and $T : H \longrightarrow H$ be continuous linear operator. If T^*T is completely continuous then T is also completely continuous.

Proof. Let (x_n) be a sequence of elements in H such that $\|x_n\| \leq 1$. Since T^*T is completely continuous there is a subsequence (x_{n_k}) of (x_n) such that the sequence $(T^*Tx_{n_k})$ converges. Now

$$\begin{aligned}
\|Tx_{n_k} - Tx_{n_p}\|^2 &= (Tx_{n_k} - Tx_{n_p}, Tx_{n_k} - Tx_{n_p}) \\
&= (T(x_{n_k} - x_{n_p}), T(x_{n_k} - x_{n_p})) \\
&= (T^*T(x_{n_k} - x_{n_p}), x_{n_k} - x_{n_p}) \\
&\leq \|T^*T(x_{n_k} - x_{n_p})\|\|x_{n_k} - x_{n_p}\| \\
&\leq \|T^*Tx_{n_k} - T^*Tx_{n_p}\|\{\|x_{n_k}\| + \|x_{n_p}\|\} \\
&\leq 2\|T^*Tx_{n_k} - T^*Tx_{n_p}\| \\
&\to 0, \quad \text{as } k, p \to \infty.
\end{aligned}$$

This however means that the sequence (Tx_{n_k}) is a Cauchy sequence in H. So (Tx_{n_k}) is convergent and hence T is completely continuous. This proves the theorem. □

Theorem 5.0.12. If $T : H \longrightarrow H$ is completely continuous, then its adjoint T^* is also completely continuous.

Proof. Let (x_n) be a sequence of elements in H such that $\|x_n\| \leq 1$. Then,

$$\|T^*x_n\| \leq \|T^*\|\|x_n\| \leq \|T^*\| = \|T\|.$$

Therefore the sequence (T^*x_n) is bounded. Since T is completely continuous, the sequence (TT^*x_n) has a convergent susequence. This shows that the operator TT^* is completely continuous.

If $S = T^*$, then since $TT^* = (T^*)^*T^* = S^*S$, we have S^*S is completely continuous. So by Theorem 5.0.11, $S = T^*$ is completely continuous. This proves the theorem. □

Theorem 5.0.13. The range of a completely continuous operator is separable.

Proof. Let $T : E \longrightarrow E_1$ be a completely continuous operator where E and E_1 are normed linear spaces. Let $T(A_n) = G_n$ where $A_n = \{x \in E : \|x\| \leq n\}$. Since the set A_n is bounded in E and the operator T is completely continuous, any sequence of elements from G_n has a convergent subsequence. That means G_n is compact and hence G_n is separable. So there is a countable set $F_n \subset G_n$ which is everywhere dense in G_n.

Clearly $G = \bigcup_{n=1}^{\infty} G_n$ is the range of T and $F = \bigcup_{n=1}^{\infty} F_n$ is a countable subset everywhere dense in G. Hence G is separable. This proves the theorem. □

Definition 5.0.14 (Normal Operator)**.** Let H be a Hilbert space. A continuous linear operator $N : H \longrightarrow H$ is said to be normal if it commutes with its adjoint, that is if $NN^* = N^*N$.

Note 5.0.15. Since $N^{**} = N$, it follows that if an operator N is normal, then its adjoint N^* is also normal.

Note 5.0.16. If N is self-adjoint, then$N = N^*$. In that case $NN^* = N^*N$ and hence N is normal.

Theorem 5.0.17. If N_1 and N_2 are normal operators on a Hilbert space H such that either commutes with the adjoint of the other then the operators N_1N_2 and $N_1 + N_2$ are both normal.

Proof. Suppose that N_1 commutes with the adjoint of N_2 so that $N_1N_2^* = N_2^*N_1$. Then we see that

$$\begin{aligned} (N_1N_2^*)^* &= (N_2^*N_1)^* \\ \text{i.e., } N_2N_1^* &= N_1^*N_2. \end{aligned}$$

This means that N_2 commutes with the adjoint of N_1.

Now

$$\begin{aligned} (N_1N_2)(N_1N_2)^* &= N_1N_2N_2^*N_1^* \\ &= N_1N_2^*N_2N_1^* \\ &= N_2^*N_1N_1^*N_2 \\ &= N_2^*N_1^*N_1N_2 \\ &= (N_1N_2)^*(N_1N_2). \end{aligned}$$

Hence N_1N_2 is normal. Also

$$\begin{aligned} (N_1+N_2)(N_1+N_2)^* &= (N_1+N_2)(N_1^*+N_2^*) \\ &= N_1N_1^*+N_1N_2^*+N_2N_1^*+N_2N_2^* \\ \text{i.e., } (N_1+N_2)^*(N_1+N_2) &= (N_1^*+N_2^*)(N_1+N_2) \\ &= N_1^*N_1+N_1^*N_2+N_2^*N_1+N_2^*N_2 \\ &= N_1N_1^*+N_2N_1^*+N_1N_2^*+N_2N_2^*. \end{aligned}$$

That is $(N_1 + N_2)(N_1 + N_2)^* = (N_1 + N_2)^*(N_1 + N_2)$. This means that $N_1 + N_2$ is normal. This proves the theorem. □

Theorem 5.0.18. Let H be a Hilbert space. A continuous linear operator $T : H \longrightarrow H$ is normal if and only if

$$\|Tx\| = \|T^*x\|, \forall x \in H.$$

Proof. We have

$$\begin{aligned} & \|Tx\| = \|T^*x\| \\ \text{if and only if} \quad & \|Tx\|^2 = \|T^*x\|^2 \\ \text{i.e., if and only if} \quad & (T^*x, T^*x) = (Tx, Tx) \\ \text{i.e., if and only if} \quad & (TT^*x, x) = (T^*Tx, x) \\ \text{i.e., if and only if} \quad & ((TT^* - T^*T)x, x) = 0. \end{aligned}$$

Note that the above relation holds for all $x \in H$. Hence T is normal if and only if if

$$\|Tx\| = \|T^*x\|, \forall x \in H.$$

This proves the theorem. □

Theorem 5.0.19. If N is a normal operator, then

$$\|N^2\| = \|N\|^2$$

Proof. We first note that if A and B are continuous linear operators mapping H into itself with the property $\|Ax\| = \|Bx\|$ for all x in H, then $\|A\| = \|B\|$. Now by theorem 5.0.18, we have

$$\|N^2x\| = \|N(Nx)\| = \|N^*(Nx)\| = \|N^*Nx\|.$$

So by our previous note, we get

$$\|N^2\| = \|N^*N\|.$$

Also for any continuous linear operator $A : H \longrightarrow H$,

$$\|A^*A\| = \|A\|^2.$$

So,

$$\begin{aligned} & \|N^*N\| = \|N\|^2 \\ \text{i.e.,} \quad & \|N^2\|\| = \|N\|^2. \end{aligned}$$

This proves the theorem. ☐

Theorem 5.0.20. A continuous linear operator $T : H \longrightarrow H$ is normal if and only if its real and imaginary part commute.

Proof. Let A_1 and A_2 be the real and imaginary parts of T. Then $T = A_1 + iA_2$ where A_1, A_2 are self-adjoint operators [see Theorem 4.0.13].

Therefore,

$$\begin{aligned} T^* &= (A_1 + iA_2)^* \\ &= A_1^* + \bar{i}A_2^* \\ &= A_1 - iA_2. \\ \text{So, } TT^* &= (A_1 + iA_2)(A_1 - iA_2) \\ &= A_1^2 - iA_1A_2 + iA_2A_1 + A_2^2 \\ \text{and } \ T^*T &= (A_1 - iA_2)(A_1 + iA_2) \\ &= A_1^2 + iA_1A_2 - iA_2A_1 + A_2^2. \end{aligned}$$

From above it is clear that

$$T^*T = TT^*, \ \text{ if } A_1A_2 = A_2A_1.$$

Hence T is normal if A_1 and A_2 commute.

Conversely, if T is normal, then

$$TT^* = T^*T$$

and in that case

$$\begin{aligned} & -A_1A_2 + A_2A_1 = A_1A_2 - A_2A_1 \\ \text{i.e.,} \quad & 2A_1A_2 = 2A_2A_1 \\ \text{i.e.,} \quad & A_1A_2 = A_2A_1 \\ \text{i.e.,} \quad & A_1 \text{ and } A_2 \text{ commute.} \end{aligned}$$

This proves the theorem. ☐

Theorem 5.0.21. If T is normal then $Tx = \lambda x$ if and only if $T^*x = \overline{\lambda}x$, for $x \in H$ and where λ is a scalar.

Proof. The proof of the theorem follows from Theorem 4.0.14. □

Theorem 5.0.22. The set of all normal operators on a Hilbert space H is a closed subspace of the set of all continuous linear operators that map H into itself which contains the set of all self-adjoint operators and is closed under scalar multiplication.

Proof. i) If N is self-adjoint, then $N = N^*$ and so $NN^* = N^*N$. Therefore every self-adjoint operator is normal.

ii) If N is normal and λ is scalar, then

$$\begin{aligned}(\lambda N)(\lambda N)^* &= \lambda N(\overline{\lambda}N^*) = \lambda\overline{\lambda}NN^* = \lambda\overline{\lambda}N^*N \\ &= (\overline{\lambda}N^*)(\lambda N) = (\lambda N)^*(\lambda N).\end{aligned}$$

So λN is normal.

iii) Let (N_k) be a sequence of normal operators that converges in norm to the continuous linear operator N so that

$$\|N_k - N\| \to 0 \quad \text{as } k \to \infty.$$

Now

$$\|N_k^* - N^*\| = \|(N_k - N)^*\| = \|N_k - N\|.$$

This shows that the sequence (N_k^*) converges to N^*.

Therefore

$$\begin{aligned}\|N_kN_k^* - NN^*\| &= \|N_kN_k^* - NN_k^* + NN_k^* - NN^*\| \\ &= \|(N_k - N)N_k^* + N(N_k^* - N^*)\| \\ &\leq \|N_k - N\|\|N_k^*\| + \|N\|\|N_k^* - N^*\| \\ &\to 0, \text{ as } k \to \infty.\end{aligned}$$

Similarly,

$$\|N_k^*N_k - N^*N\| \to 0 \quad \text{as } k \to \infty.$$

So,

$$\begin{aligned}\|NN^* - N^*N\| &= \|NN^* - N_kN_k^* + N_k^*N_k - N^*N\| \\ &\leq \|NN^* - N_kN_k^*\| + \|N_k^*N_k - N^*N\| \\ &\to 0 \quad \text{as } k \to \infty.\end{aligned}$$

This shows that $NN^* = N^*N$ and hence N is normal. This proves the theorem.

□

Definition 5.0.23 (Unitary Operators)**.** A continuous linear operator T that maps a Hilbert space H into itself is said to be unitary if it satisfies the condition

$$TT^* = T^*T = I$$

where I is the identity mapping.

Remark 5.0.24. If T is unitary, then T is injective. For, if $Tx_1 = Tx_2$ then operating both the sides by the operator T^*, we get $T^*Tx_1 = T^*Tx_2$ and $T^*T = I$ implies that $x_1 = x_2$. Also if T is unitary, then T is surjective mapping. Because if $y \in H$ then

$$T(T^*(y)) = TT^*y = Iy = y.$$

It follows, therefore that the unitary operators on H are precisely those operators whose inverses are equal to their adjoints.

Remark 5.0.25. Every unitary operator is normal. The following example shows that a normal operator need not be self-adjoint or unitary.

Let $I : H \longrightarrow H$ be the identity operator. Let $T = 2iI$. Then $T^* = (2iI) = -2iI$. Also $T^{-1} = -\dfrac{1}{2}iI$. So, $TT^* = T^*T = 4I, T^* \neq T$. Also, $T^* \neq T^{-1}$. This shows that T is normal which is neither self-adjoint nor unitary.

Theorem 5.0.26. A continuous linear operator $T : H \longrightarrow H$ is unitary if and only if T is an isomorphism of H onto itself.

Proof. If T is unitary, then T is bijective and also $T^*T = I$. So by Theorem 4.0.9, $\|Tx\| = \|x\|$. So T is an isomorphism of H onto itself.

Conversely, if T is an isomorphism of H onto itself, then T^{-1} exists and $\|Tx\| = \|x\|$. So by Theorem 4.0.9, $T^*T = I$. It now follows that

$$\begin{aligned} & (T^*T)T^{-1} = IT^{-1} \\ \text{i.e.,}\quad & T^* = T^{-1} \\ \text{i.e.,}\quad & TT^* = I. \end{aligned}$$

So, $TT^* = T^*T = I$ and T is unitary. This proves the theorem. □

Definition 5.0.27 (Positive Operators)**.** Let H be a Hilbert space and $A : H \longrightarrow H$ be a self-adjoint operator so that (Ax, x) is real for all x in H. The operator A is called positive, $A \geq 0$, if $(Ax, x) \geq 0$ for all x in H.

Remark 5.0.28. If both the operators A and B are self-adjoint and if

$$\begin{aligned} & (Ax, x) \geq (Bx, x),\ \forall x \in H \\ \text{i.e., if}\quad & ((A - B)x, x) \geq 0\ \ \forall x \in H \\ \text{i.e., if}\quad & A - B \geq 0, \end{aligned}$$

then A is said to be greater than B or B is said to be less than A. In notation, we write $A \geq B$ or $B \geq A$.

Remark 5.0.29. Let A be a self-adjoint operator. Since $(A^2x, x) = (Ax, A^*x) = (Ax, Ax) = \|Ax\|^2 \geq 0$, it follows that the square of a self-adjoint operator is positive.

Remark 5.0.30. We note that for any any continuous linear operator $A : H \longrightarrow H$, the operator AA^* and A^*A are self-adjoint. We also have

$$\begin{aligned} & (AA^*x, x) = (A^*x, A^*x) = \|A^*x\|^2 \geq 0 \\ \text{and}\quad & (A^*Ax, x) = (Ax, Ax) = \|Ax\|^2 \geq 0. \end{aligned}$$

Hence the operators AA^* and A^*A are always positive. It is also clear that the sum of two positive operators is positive.

Theorem 5.0.31. If A and B are positive self-adjoint operators such that $AB = BA$, then the operator AB is positive.

Proof. If $A = 0$, the zero operator, then the result is clear. We therefore suppose that $A \neq 0$. By Theorem 4.0.3(iv), AB is self-adjoint. We put

$$\begin{aligned}&A_1 = \frac{1}{\|A\|}A, A_2 = A_1 - A_1^2, A_3 = A_2 - A_2^2, \cdots,\\ &A_{n+1} = A_n - A_n^2, \cdots\end{aligned}$$

and we show that

$$0 \le A_n \le I \tag{5.0.1}$$

for $n = 1, 2, 3, \cdots$ and I is the identity operator. Since A is self-adjoint, it is clear that each A_n is self-adjoint. If $n = 1$, then

$$(A_1x, x) = \frac{1}{\|A\|}(Ax, x) \ge 0, \text{ and so } A_1 \ge 0.$$

Also, because

$$\begin{aligned}(A_1x, x) &= |(A_1x, x)| \le \|A_1\|\|x\|^2 = \|x\|^2\\ &= (x, x) = (Ix, x),\end{aligned}$$

we obtain

$$\begin{aligned}&((I - A_1)x, x) \ge 0\\ \text{i.e.,}\quad &A_1 \le I.\end{aligned}$$

Thus the relation (5.0.1) is true for $n = 1$.

We now suppose that (5.0.1) is true for $n = k$. Then

$$\begin{aligned}(A_k^2(I - A_k)x, x) &= (A_k(I - A_k)x, A_k^*x)\\ &= (A_k(I - A_k)x, A_kx)\\ &= ((I - A_k)A_kx, A_kx) \ge 0\end{aligned}$$

because $I - A_k$ is positive. Therefore,

$$A_k^2(I - A_k) \ge 0.$$

Similarly, it can be shown that $A_k(I - A_k)^2 \ge 0$. As the sum of two positive operators is positive, it follows that

$$A_{k+1} = A_k^2(I - A_k) + A_k(I - A_k)^2 \ge 0.$$

Further,

$$I - A_{k+1} = (I - A_k) + A_k^2 \ge 0.$$

This shows that the relation (5.0.1) is true for $n = k + 1$. Hence (5.0.1) is true for all n.

Moreover,

$$\begin{aligned}A_1 &= A_1^2 + A_2 = A_1^2 + A_2^2 + A_3 = \cdots\\ &= A_1^2 + A_2^2 + \cdots + A_n^2 + A_{n+1}\end{aligned}$$

This implies that for any positive integer n

$$\sum_{k=1}^{n} A_k^2 = A_1 - A_{n+1} \le A_1. \tag{5.0.2}$$

So,

$$A_1 - \sum_{k=1}^{n} A_k^2 \geq 0$$

$$\text{i.e.,} \qquad \left((A_1 - \sum_{k=1}^{n} A_k^2)x, x \right) \geq 0$$

$$\text{i.e.,} \qquad (A_1 x, x) - \sum_{k=1}^{n} (A_k^2 x, x) \geq 0$$

$$\text{i.e.,} \qquad (A_1 x, x) - \sum_{k=1}^{n} \|A_k x\|^2 \geq 0$$

$$\text{i.e.,} \qquad \sum_{k=1}^{n} \|A_k x\|^2 = \sum_{k=1}^{n} (A_k x, A_k x) \leq (A_1 x, x).$$

This however means that the infinite series

$$\sum_{k=1}^{\infty} \|A_k x\|^2 = \sum_{k=1}^{\infty} (A_k x, A_k x)$$

is convergent and so $\lim_{n\to\infty} \|A_n x\| = 0$ i.e., $\lim_{n\to\infty} A_n x = 0$.

So from (5.0.2) we obtain

$$\begin{aligned} \left(\sum_{k=1}^{n} A_k^2 \right) x &= (A_1 - A_{n+1})x = A_1 x - A_{n+1} x \\ &\to A_1 x \quad \text{as} \quad n \to \infty. \end{aligned}$$

Since B is continuous, we have

$$B \left(\sum_{k=1}^{n} A_K^2 \right) x \to BA_1 x \quad \text{as } n \to \infty$$

$$\text{i.e.,} \qquad \sum_{k=1}^{n} BA_K^2 x \to BA_1 x \quad \text{as } n \to \infty$$

$$\text{and so} \qquad \left(\sum_{k=1}^{n} BA_K^2 x, x \right) \to (BA_1 x, x) \quad \text{as } n \to \infty.$$

Now since B commutes with A, it commutes with each A_k and hence

$$\begin{aligned} (ABx, x) &= \|A\|(A_1 Bx, x) = \|A\|(BA_1 x, x) \\ &= \|A\| \lim_{n\to\infty} \sum_{k=1}^{n} (BA_k^2 x, x) \\ &= \|A\| \lim_{n\to\infty} \sum_{k=1}^{n} (A_k^2 Bx, x) \\ &= \|A\| \lim_{n\to\infty} \sum_{k=1}^{n} (A_k Bx, A_k^* x) \\ &= \|A\| \lim_{n\to\infty} \sum_{k=1}^{n} (BA_k x, A_k x) \geq 0, \end{aligned}$$

because B is positive. This shows that $AB \geq 0$ and the theorem is proved.

$\square$

Definition 5.0.32. A sequence (A_n) of self-adjoint operators in a Hilbert space H is called increasing(decreasing) if $A_n \leq A_{n+1}(A_n \geq A_{n+1})$ for all n.

Theorem 5.0.33. Let (A_n) be a sequence of self-adjoint operators in a Hilbert space H such that

$$A_1 \leq A_2 \leq \cdots \leq A_n \leq \cdots \leq B$$

where B is a self-adjoint operator on H. Suppose further that any A_j permutes with B and with every A_m. Then (A_n) is strongly convergent and the limit operator A is linear, bounded and self-adjoint and satisfies $A \leq B$.

An analogous result holds for monotone decreasing sequence.

Proof. The proof of the theorem is beyond the scope of this study material. $\square$

Definition 5.0.34. Let H be a Hilbert space and $A : H \longrightarrow H$ be a positive operator. A self-adjoint operator B defined on H is called a square root of A if $B^2 = A$. If, in addition $B \geq 0$, then B is called a positive square root of A and is denoted by $B = A^{\frac{1}{2}}$.

Theorem 5.0.35. Every positive self-adjoint operator A has a unique positive square root B. The operator B is permutable with any operator that permutes with A.

Proof. If $A = 0$, then we take $B = 0$. So, we assume that $A \neq 0$. We can further assume that $A \leq I$ where I is the identity operator. Because, if not, let

$$A_1 = \frac{1}{\|A\|}A$$

so that $\|A_1\| = 1$. By Schwarz inequality

$$(A_1x, x) \leq \|A_1x\|\|x\| \leq \|A_1\|\|x\|^2 = \|x\|^2 = (x, x)$$

and hence

$$\begin{aligned} ((I - A_1)x, x) &= (Ix, x) - (A_1x, x) \\ &= (x, x) - (A_1x, x) \geq 0 \\ \text{i.e., } A_1 &\leq I. \end{aligned}$$

We now construct a sequence of operators by

$$\begin{aligned} &B_0 = 0 \\ &B_1 = B_0 + \frac{1}{2}(A - B_0^2) = \frac{1}{2}A \\ &B_2 = B_1 + \frac{1}{2}(A - B_1^2) \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &B_{n+1} = B_n + \frac{1}{2}(A - B_n^2) \end{aligned} \tag{5.0.3}$$

and so on. Since A is self-adjoint and the sequence of a self-adjoint operator is self-adjoint, it follows that all B_n are self-adjoint.

We now show that each B_n is permutable with any operator that is permutable with A. In fact, if B' permutes with A, then

$$B'A = AB'.$$

In that case,

$$B'B_1 = B'\left(\frac{1}{2}A\right) = \frac{1}{2}(B'A) = \frac{1}{2}AB' = B_1B',$$

$$\begin{aligned}
B'B_2 &= B'\left[B_1 + \frac{1}{2}(A - B_1^2)\right] \\
&= \left[B'B_1 + \frac{1}{2}(B'A - B'B_1B_1)\right] \\
&= \left[B_1B' + \frac{1}{2}(AB' - B_1^2B')\right] \\
&= \left[B_1 + \frac{1}{2}(A - B_1^2)\right]B' \\
&= B_2B'.
\end{aligned}$$

So, in general, $B'B_n = B_nB'$ for all n. In particular, $AB_n = B_nA$ and $AB_m = B_mA$, $\forall m, n$ and so by our preceding remark

$$B_nB_m = B_mB_n \ \forall\, m \text{ and } n.$$

This shows that the sequence (B_n) constructed above is mutually permutable. Now

$$\begin{aligned}
& \frac{1}{2}(I - B_n)^2 + \frac{1}{2}(I - A) \\
&= \frac{1}{2}(I - 2B_n + B_n^2) + \frac{1}{2}(I - A) \\
&= I - B_n + \frac{1}{2}B_n^2 - \frac{1}{2}A \\
&= I - \left[B_n + \frac{1}{2}(A - B_n^2)\right] \\
&= I - B_{n+1} \\
\text{i.e., } I - B_{n+1} &= \frac{1}{2}(I - B_n^2) + \frac{1}{2}(I - A).
\end{aligned}$$

Since $A \le I$, we obtain $B_n \le I$ for all n.

Again, by (5.0.3) we have

$$\begin{aligned}
B_{n+1} - B_n &= \left[B_n + \frac{1}{2}(A - B_n^2)\right] - \left[B_{n-1} + \frac{1}{2}(A - B_{n-1}^2)\right] \\
&= (B_n - B_{n-1}) - \frac{1}{2}(B_n^2 - B_{n-1}^2) \\
&= \left[I - \frac{1}{2}(B_n + B_{n-1})\right](B_n - B_{n-1}) \\
&= \left[\frac{1}{2}(I - B_n) + \frac{1}{2}(I - B_{n-1})\right](B_n - B_{n-1}). \qquad (5.0.4)
\end{aligned}$$

Since $B_n \le I$ for all n, equality (5.0.4) shows that $B_{n+1} \ge B_n$ provided $B_n \ge B_{n-1}$ for each n. But

$$B_1 = \frac{1}{2}A \ge B_0$$

and so $B_{n+1} \geq B_n, \forall n$.

We therefore see that the sequence (B_n) constructed above is such that

$$B_0 \leq B_1 \leq B_2 \leq \cdots \leq B_n \leq B_{n+1} \leq \cdots \leq I.$$

Further this sequence is mutually permutable and commutes with I. So by Theorem 5.0.33, the sequence (B_n) converges strongly to a self-adjoint operator B which satisfies the relation $B \leq I$. We now verify that $B \geq 0$ and $B^2 = A$.

Since $B_1 \geq 0$ and the sequence (B_n) is increasing, it follows that each B_n is positive and hence

$$(B_n x, x) \geq 0 \;\; \forall\, n.$$

Proceeding to the limit as $n \to \infty$ we obtain

$$\begin{aligned} & (Bx, x) \geq 0 \\ \text{i.e.} \qquad & B \geq 0. \end{aligned}$$

Letting $n \to \infty$ in (5.0.3) we get

$$B = B + \frac{1}{2}(A - B^2) \;\; \text{i.e., } B^2 = A.$$

So, the existence of a positive square root B of the operator A is obtained.

Now, B_n is permutable with any operator that permutes with A. So if the operator C permutes with A then C permutes with B_n. That means

$$B_n C = C B_n$$

and so $B_n C x = C B_n x$ for all x in H.

Taking limit,

$$\begin{aligned} BCx &= CBx \;\; \forall x \in H \\ \text{i.e., } BC &= CB. \end{aligned}$$

So B permutes with C, i.e., B is permutable with any operator which permutes with A.

We now prove the uniqueness. Let $\overline{B}$ be another positive square root of A. Since $\overline{B}$ permutes with A, by the preceeding remark we have

$$\overline{B}B = B\overline{B}.$$

Let $x \in H$ and $y = (B - \overline{B})x$. Then

$$\begin{aligned} (By, y) + (\overline{B}y, y) &= ((B + \overline{B})y, y) \\ &= ((B^2 - \overline{B}^2)x, y) \\ &= ((A - A)x, y) = 0. \end{aligned}$$

Since both B and $\overline{B}$ are positive, it follows that

$$(By, y) = 0 = (\overline{B}y, y).$$

Because B is positive, by what we have already proved, there is a self-adjoint operator C such that $B = C^2$. So

$$\begin{aligned} \|Cy\| &= (Cy, Cy) = (y, C^*Cy) = (y, C^2 y) \\ &= (y, By) = 0 \\ \text{i.e., } Cy &= \theta. \end{aligned}$$

So,
$$By = C^2y = C(Cy) = \theta.$$
Similarly, $\overline{B}y = \theta$. Thus for $x \in H$,
$$\begin{aligned}
\|Bx - \overline{B}x\|^2 &= \|(B - \overline{B})x\|^2 \\
&= ((B - \overline{B})x, (B - \overline{B})x) \\
&= ((B - \overline{B})^*(B - \overline{B})x, x) \\
&= ((B - \overline{B})y, x) \\
&= (By, x) - (\overline{B}y, x) \\
&= 0 \\
\text{i.e., } Bx &= \overline{B}x \ \ \forall\, x \in H.
\end{aligned}$$
Since $x \in H$ is arbitrary, it follows that
$$B = \overline{B}.$$
This proves the uniqueness and hence the theorem. $\square$

Unit 6

Course Structure

- Projection operator and their sum, product & permutability, invariant subspaces, closed linear transformation, closed graph theorem and open mapping theorem.

Let L be a closed subspace of a Hilbert space H and $x \in H$. Then there exist a unique decomposition $x = y + z$, where $y \in L$ and $z \in L$.

The element y is called the projection of the element x in L.

6.0.1 Projection Operator:-

We can define an operator P by the rule $P(x) = y$ because this association depends on the subspace L. We sometimes write P_L instead of P to indicate the subspace L. This operator P_L whose domain is H and range is L is called a projection operator. We say that P is a projection on the closed subspace L.

Theorem 6.0.1. P_L is a self-adjoint operator with $\|P_L\| \leq 1$ and $P_L = P_L$.

Proof. Clearly P_L is a linear operator. If $x = y + z$ where $y \in L$ and $z \perp L$, then $P_L(x) = y$ and since $y \perp z$, we have

$$\|x\|^2 = \|y + z\|^2 = \|y\|^2 + \|z\|^2 \geq \|y\|^2$$

So,

$$\|P_L(x)\| = \|y\| \leq \|x\|, \ \ \forall x \in H$$

so that, $\|P_L\| \leq 1$.

But if $x \in L$, then $P_L(x) = x$ and then $\|P_L(x)\| = \|x\|$, i.e., $\|P_L\| = 1$.

Let $x_1, x_2 \in H$ and y_1, y_2 be their projections on L, i.e. $x_1 = y_1 + z_1$, $x_2 = y_2 + z_2$ where $y_1, y_2 \in L$ and $z_1, z_2 \perp L$ then,

$$\begin{aligned}
(P_L x_1, x_2) = (y_1, x_2) &= (y_1, y_2 + z_2) \\
&= (y_1, y_2) + (y_1, z_2) \\
&= (y_1, y_2)
\end{aligned}$$

and $(x_1, P_L x_2) = (x_1, y_2) = (y_1, y_2)$ so that, $(x_1, P_L x_2) = (P_L x_1, x_2) = (x_1, P_L^* x_2)$ for every $x_1, x_2 \in H$.

Hence, $P_L = P_L^*$ and so P_L is self-adjoint.

Now for all $x \in H$, $P_L x \in L$ and for $x' \in L$, $P_L x' = x'$. So,

$$P_L^2 x = P_L(P_L x) = P_L x; \;\; \forall x \in H$$

So, $P_L^2 = P_L$.

This proves the theorem. $\square$

Theorem 6.0.2. Every self-adjoint operator with $P^2 = P$ is a projection operator on some closed subspace.

Proof. Let, L denote the set of all elements $y \in H$ of the form $y = Px$ for all $x \in H$. Since, P is linear, it can be verified that L is a subspace. We will show that L is closed.

Suppose that $y_n \to y$ where $y_n \in L$. So, we can assume that $y_n = P(x_n)$ with $x_n \in H$. So, $Py_n = P(Px_n) = P^2 x_n = Px_n = y_n$ and the continuity of P implies that $Py_n \to Py$, i,e, $y_n \to Py$.

So, $y = Py$ and $y \in L$.

Now, for $x, x' \in H$

$$\begin{aligned}
(x - Px, Px') &= (P^*(x - Px), x') \\
&= (P(x - Px), x') \;\; [\text{ since } P^* = P] \\
&= (Px - P^2, x') \\
&= (Px - Px, x') \\
&= (\theta, x') = 0
\end{aligned}$$

So, $x - Px \perp Px'$.

Since, x' is an arbitrary element in H, we have $x - Px \perp L$.

Now, $x = Px + (x - Px)$, where $Px \in L$ and $(x - Px) \perp L$. So, P is a projection operator on L. This proves the theorem. $\square$

Definition 6.0.3. The projection operators P_1 and P_2 are called orthogonal if $P_1 P_2 = 0$, zero operator.

Theorem 6.0.4. Two projection operators P_1 and P_2 are orthogonal iff their corresponding subspaces L_1 and L_2 are orthogonal to each other, i.e. $L_1 \perp L_2$.

Proof. If $P_1 P_2 = 0$, then if $x_1 \in L_1$ and $x_2 \in L_2$, we have,

$$\begin{aligned}
(x_1, x_2) = (P_1 x_1, P_2 x_2) &= (x_1, P_1^* P_2 x_2) \\
&= (x_1, P_1 P_2 x_2) \\
&= (x_1, \theta) = 0
\end{aligned}$$

so that $L_1 \perp L_2$.

If $L_1 \perp L_2$, then because for all $x \in H$ $P_2 x \in L_2$, it follows that $P_2 x \perp L_1$.

So, $P_2 x = \theta + P_2 x$, where $\theta \in L_1$ and $P_2 x \perp L_1$.

Therefore,

$$\begin{aligned}
& P_1 P_2 x = \theta \quad \text{for } x \in H \\
\text{i.e.} \qquad & P_1 P_2 = 0
\end{aligned}$$

This proves the theorem. $\square$

Theorem 6.0.5. The sum of two projection operators P_{L_1} and P_{L_2} is a projection operator iff these operators are orthogonal. If P_{L_1} is orthogonal to P_{L_2} then $P_{L_1} + P_{L_2} = P_{L_1 \oplus L_2}$.

Proof. Necessity: Let $P_{L_1} + P_{L_2}$ be a projection operator.

$$\begin{aligned} \text{So} \quad & (P_{L_1} + P_{L_2})^2 = P_{L_1} + P_{L_2} \\ \text{i.e.} \quad & P_{L_1}^2 + P_{L_2}^2 + P_{L_1}P_{L_2} + P_{L_1}P_{L_2} = P_{L_1} + P_{L_2} \\ \text{i.e.} \quad & P_{L_1}P_{L_2} + P_{L_2}P_{L_1} = 0 \end{aligned}$$

We operate by P_{L_1} on the left and obtain that

$$\begin{aligned} & P_{L_1}^2 P_{L_2} + P_{L_1}P_{L_2}P_{L_1} = 0 \\ \text{i.e.} \quad & P_{L_1}P_{L_2} + P_{L_1}P_{L_2}P_{L_1} = 0 \end{aligned} \tag{6.0.1}$$

If we now operate by P_{L_1} on the right then we obtain that

$$\begin{aligned} & P_{L_1}P_{L_2}P_{L_1} + P_{L_1}P_{L_2}P_{L_1}^2 = 0 \\ \text{i.e.} \quad & P_{L_1}P_{L_2}P_{L_1} = 0 \end{aligned} \tag{6.0.2}$$

So, from (6.0.1) and (6.0.2), $P_{L_1}P_{L_2} = 0$. This proves the necessity part of the theorem.

Sufficiency: Let $P_{L_1}P_{L_2} = P_{L_2}P_{L_1} = 0$ then

$$\begin{aligned} (P_{L_1} + P_{L_2})^2 &= P_{L_1}^2 + 2P_{L_1}P_{L_2} + P_{L_2}^2 \\ &= P_{L_1} + P_{L_2} \\ \text{and, } (P_{L_1} + P_{L_2})^* &= P_{L_1}^* + P_{L_2}^* = P_{L_1} + P_{L_2} \end{aligned}$$

Therefore, $(P_{L_1} + P_{L_2})$ is a projection operator. Suppose, now that $P_{L_1}P_{L_2} = 0$, so that by earlier theorem $L_1 \perp L_2$. If $P = P_{L_1} + P_{L_2}$ and $x \in H$, then,

$$Px = (P_{L_1} + P_{L_2})x = P_{L_1}x + P_{L_2}x \in P_{L_1 \oplus L_2}$$

Also,

$$\begin{aligned} (x - Px, Px) &= (P^*(x - Px), x) \\ &= (Px - P^2x, x) = 0 \\ \text{i.e. } \; x - Px &\perp Px \end{aligned}$$

So,

$$\begin{aligned} & x = Px + (x - Px), \quad \text{where } Px \in P_{L_1 \oplus L_2} \\ \text{and} \quad & x - Px \perp (L_1 \oplus L_2) \end{aligned}$$

So, P is a projection operator on $L_1 \oplus L_2$. This proves the theorem. □

Theorem 6.0.6. The product of the projection operators P_{L_1} and P_{L_2} is a projection operator iff P_{L_1} and P_{L_2} are permutable. If this condition is satisfied then $P_{L_1}P_{L_2} = P_{L_1 \cap L_2}$.

Proof. Suppose that $P_{L_1}P_{L_2}$ is a projection operator. Then

$$(P_{L_1}P_{L_2}) = (P_{L_1}P_{L_2})^* = P_{L_2}^* P_{L_1}^* = P_{L_2}P_{L_2}$$

and the permutability is obtained.

Sufficiency: Suppose that $P_{L_1}P_{L_2} = P_{L_2}P_{L_1}$ then,

$$(P_{L_1}P_{L_2})^* = P^*_{L_2}P^*_{L_1} = P_{L_2}P_{L_1} = P_{L_1}P_{L_2},$$

so that, $P_{L_1}P_{L_2}$ is self-adjoint.

Also,

$$\begin{aligned}(P_{L_1}P_{L_2})^2 = P_{L_1}P_{L_2}P_{L_1}P_{L_2} &= P_{L_1}P_{L_1}P_{L_2}P_{L_2} \\ &= P^2_{L_1}P^2_{L_2} = P_{L_1}P_{L_2}.\end{aligned}$$

So, $P_{L_1}P_{L_2}$ is a projection operator.

Third part: Now suppose that $P_{L_1}P_{L_2} = P_{L_2}P_{L_1}$ and let $x \in H$ be arbitrary.

If $P = P_{L_1}P_{L_2}$ then,

$$Px = P_{L_1}P_{L_2}x = P_{L_2}P_{L_1}x$$

lies both in L_1 and L_2 and so lies in $L_1 \cap L_2$.

If $y \in L_1 \cap L_2$, then,

$$Py = P_{L_1}(P_{L_2}y) = P_{L_1}y = y$$

If now $x \in H$ and $y \in L_1 \cap L_2$ then

$$\begin{aligned}(x - Px, y) = (x - Px, Py) &= (P^*(x - Px), y) \\ &= (Px - P^2x, y) \\ &= 0,\end{aligned}$$

so that, $x - Px \perp L - 1 \cap L_2$.

Therefore, any $x \in H$ has a representation $x = Px + (x - Px)$, where $Px \in L_1 \cap L_2$ and $x - Px \in L_1 \cap L_2$. So, P is a projection operator on $L_1 \cap L_2$. □

6.0.2 Some results for $P_1 \sim P_2$

Proof. Necessity: Suppose that $P_1 - P_2$ is a projection operator. Then,

$$\begin{aligned}P_1 - P_2 = (P_1 - P_2)^2 &= P_1^2 - P_1P_2 - P_2P_1 + P_2^2 \\ &= P_1 - P_1P_2 - P_2P_1 + P_2,\end{aligned}$$

so that,

$$P_2P_1 + P_1P_2 = 2P_2 \tag{6.0.3}$$

Operating by P_1 from the left and from the right, we get that

$$P_1P_2P_1 + P_1P_2 = 2P_1P_2,$$

since $P_1^2 = P_1$, $P_1P_2P_1 = P_1P_2$.

Again, operating from the right, we get

$$\begin{aligned}&P_2P_1 + P_1P_2P_1 = 2P_2P_1 \\ \Rightarrow\ &P_1P_2P_1 = P_2P_1\end{aligned}$$

Hence, $P_1P_2P_1 = P_1P_2$, $P_1P_2P_1 = P_2P_1$ and by (6.0.3) $P_1P_2 = P_2P_1 = P_2$.

Sufficiency: Let $P_1P_2 = P_2$ i.e. $(P_1P_2)^* = P_2^*$ then $P_2P_1 = P_2$.
If $P = P_1 - P_2$ then

$$\begin{aligned} P^2 = (P_1 - P_2)^2 &= P_1 - P_1P_2 - P_2P_1 + P_2 \\ &= P_1 - P_2P_1 = P_1 - P_2 = P \end{aligned}$$

and $P^* = (P_1 - P_2)^* = P_1 - P_2 = P$ i.e. $P = P_1 - P_2$ is a projection operator. This proves the theorem. □

Definition 6.0.7. Let, H be a Hilbert space and A be continuous linear operator such that $A : H \longrightarrow H$. If $X \subset H$, let $A(X) = \{A(x) : x \in X\}$. A closed subspace M of H is said to be invariant under A if $A(M) \subset M$.

If both M and $M^\perp$ are invariant under A, then we say that M reduces A or that A reduced by M.

Theorem 6.0.8. A closed subspace M of H is invariant under A iff $M^\perp$ is invariant under A^*.

Proof. Suppose that, M is invariant under A. If $y \in M^\perp$ then $Ax \perp y$ for all $x \in M$, i.e. $(Ax, y) = 0$. But $(Ax, y) = (x, A^*y)$. So, $(x, A^*y) = 0$, i.e. $A^*y \perp x$ for $y \in M^\perp$. So, $A^*y \in M^\perp$, $\forall y \in M^\perp$.

This however implies that $A^*(M^\perp) \subset M^\perp$. So, $M^\perp$ is invariant under A^*.

Converse part: Let, $M^\perp$ be invariant under A^*. Then what we have just shown $(M^\perp)^\perp$ will be invariant under $(A^*)^*$.

We know that for a closed subspace M of a Hilbert space H, $(M^\perp)^\perp = M$.

Also, we have shown earlier $(A^*)^* = A$.

Therefore we may conclude that M is invariant under A. This proves the theorem. □

Theorem 6.0.9. A closed subspace M of H reduces A iff M is invariant under both A and A^*.

Proof. Suppose that, M reduces A, then by definition both M and $M^\perp$ are invariant under A. Now by earlier theorem $(M^\perp)^\perp$ is invariant under A^*, i.e. M is invariant under A^*.

Converse part: Suppose now that M is invariant under both A and A^*. Then by an earlier theorem $M^\perp$ is invariant under A^{**}. But $A^{**} = A$. So, M and $M^\perp$ are both invariant under A. So, M reduces A. This proves the theorem. □

Problem 6.1. If P is a projection operator on a closed subspace M of H then M is invariant under a continuous linear operator T iff $TP = PTP$.

Solution. If M is invariant under T and $x \in H$ then TPx is in M and so,

$$PTPx = TPx$$

So, $PTP = TP; \forall x \in H$.

Conversely if $TP = PTP$ and $x \in M$, then $Tx = TPx = PTPx \in M$. So, M is invariant under T. ■

Problem 6.2. If P is a projection on a closed subspace M of H then M reduces to a continuous linear operator T iff $TP = PT$.

Solution. By earlier theorems, M reduces T iff M is invariant under T amd T^* i.e. if and only if $TP = PTP$ and $T^*P = PT^*P$.

But $T^*P = PT^*P$ is equivalent to

$$\begin{aligned} (PT)^* &= (PTP)^* \\ \text{i.e. } PT &= PTP. \end{aligned}$$

So, M reduces T iff $TP = PTP$ and $PT = PTP$.

Suppose that, M reduces T, then from above $PT = TP$.

Conversely if, $PT = TP$, then

$$\begin{aligned} PTP &= TP^2 \\ \Rightarrow PTP &= TP \quad [\text{ since } P^2 = P] \\ \text{and } \; P^2T &= PTP \\ \text{i.e. } \; PT &= PTP. \end{aligned}$$

So, M reduces T. ■

Definition 6.2.1 (Closed Linear Transformation)**.** Let X and Y be normed linear spaces and M is a subspace of X. Then a linear transformation $T : M \longrightarrow Y$ is said to be closed if, $x_n \to x$, where $x_n \in M$ and $Tx_n \to y$. Then $x \in M$ and $y = Tx$.

6.2.1 Open Mapping Theorem

Some Definitions

Definition 6.2.2 (Continuous)**.** Let (X, d) and (Y, ρ) be metric spaces. We say that a function $f : X \longrightarrow Y$ is continuous on X if for every open set U in Y, the inverse image $f^{-1}(U)$ is open in X, i.e. the inverse image of an open set is open. Equivalently, the inverse image of a closed set is closed.

Definition 6.2.3 (Open map)**.** Let (X, d) and (Y, ρ) be metric spaces. We say that a function $f : X \longrightarrow Y$ is open if for every open set G in X, the image $f(G)$ is open in Y, i.e. the image of an open set is open.

Definition 6.2.4 (Closed map)**.** Let (X, d) and (Y, ρ) be metric spaces. We say that a function $f : X \longrightarrow Y$ is closed if for every closed set F in X, the image $f(F)$ is closed in Y, i.e. the image of a closed set is closed.

Notations

1. $B_X(x_0, r) = \{x \in X : \|x - x_0\| < r\}$ is an open ball centered at x_0 with radius r in X.
2. $B_X(x_0, r) + z = \{x + z : x \in B_X(x_0, r)\}$ where $z \in X$.
   It is easy to verify that $B_X(0, r) + x_0 = B_X(x_0, r)$.
3. $cB_X(x_0, r) = \{cx : x \in B_X(x_0.r)\}$ where c is scalar. It is easy to verify that $B_X(0, r) = rB_X(0, 1)$.

Lemma 6.2.5 (Open unit ball)**.** Suppose T is a bounded linear operator from a Banach space X onto a Banach space Y. Then $B_Y(0, r) \subset T(B_X(0, 1))$ for some $r > 0$.

Proof. Claim-1: $B_Y(y_0, \delta) \subset \overline{T\left(B_X\left(0, \frac{1}{2}\right)\right)}$ for some $\delta > 0$. We can write

$$X = \cup_{k=1}^{\infty} B_X\left(0, \frac{k}{2}\right) = \cup_{k=1}^{\infty} kB_X\left(0, \frac{1}{2}\right), \quad \text{since } x \in X, \|x\| \leq \frac{k}{2}, \text{ for some } k$$

Thus,

$$\begin{aligned} T(X) = \cup_{k=1}^{\infty} T\left(B_X\left(0, \frac{k}{2}\right)\right) &= \cup_{k=1}^{\infty} kT\left(B_X\left(0, \frac{1}{2}\right)\right), \quad \text{since } T \text{ is linear} \\ \Rightarrow \qquad Y = \cup_{k=1}^{\infty} kT\left(B_X\left(0, \frac{1}{2}\right)\right) &= \overline{\cup_{k=1}^{\infty} kT\left(B_X\left(0, \frac{1}{2}\right)\right)}, \quad [\text{ since } T \text{ is onto}]. \end{aligned}$$

Since Y is a Banach space and using Baire's category theorem, we get the interior of $\overline{kT\left(B_X\left(0,\frac{1}{2}\right)\right)}$ is non-empty for some k. Therefore, the interior of $\overline{T\left(B_X\left(0,\frac{1}{2}\right)\right)}$ is non-empty, $B_Y(y_0,\delta)\subset\overline{T\left(B_X\left(0,\frac{1}{2}\right)\right)}$.

Claim-2: $B_Y(y_0,\frac{\delta}{2^n})\subset\overline{T\left(B_X\left(0,\frac{1}{2^n}\right)\right)}, \forall n\geq 0$.

It is enough to show that for $n=0$, $B_Y(0,\delta)\subset\overline{T(B_X(0,1))}$.

Let $y\in B_Y(0,\delta)$. Then $y+y_0\in B_Y(y_0,\delta)\subset\overline{T\left(B_X\left(0,\frac{1}{2}\right)\right)}$. By definition of closure of a set, $\exists u_n\in T\left(B_X\left(0,\frac{1}{2}\right)\right)$, $w_n\in B_X\left(0,\frac{1}{2}\right)$ such that $T(w_n)=u_n\to y+y_0$ and $\exists v_n\in\overline{T\left(B_X\left(0,\frac{1}{2}\right)\right)}$, $z_n\in B_X\left(0,\frac{1}{2}\right)$ such that $T(z_n)=v_n\to y_0$. From this, we get $u_n-v_n=T(w_n-z_n)\to y$. Notice that $\|w_n-z_n\|<1$, so we get $y\in\overline{T(B_X(0,1))}$. Therefore, we get $B_Y(0,\delta)\subset\overline{T(B_X(0,1))}$.

Claim-3: $B_Y\left(0,\frac{\delta}{2}\right)\subset T(B_X(0,1))$.

Let, $y\in B_Y\left(0,\frac{\delta}{2}\right)$. Then by the above Claim-2, $y\in T\left(B_X\left(0,\frac{1}{2}\right)\right)$. So there exists $x_1\in B_X\left(0,\frac{1}{2}\right)$ such that $\|y-Tx_1\|<\frac{\delta}{4}$.

Now $y-Tx_1\in B_Y\left(0,\frac{\delta}{4}\right)$. Again by the above Claim 2, $y-Tx_1\in T\left(B_X\left(0,\frac{1}{4}\right)\right)$. So, there exists $x_2\in B_X\left(0,\frac{1}{4}\right)$ such that

$$\|y-Tx_1-Tx_2\|<\frac{\delta}{2^3}.$$

By repeating this procedure and using induction, we get a sequence $x_n\in B_X\left(0,\frac{1}{2^n}\right)$ such that

$$\left\|y-\sum_{k=1}^{n}Tx_k\right\|<\frac{\delta}{2^{n+1}}. \tag{6.2.1}$$

Define $z_n=\sum_{k=1}^{n}x_k$. Then

$$\|z_n-z_m\|\leq\sum_{k=m+1}^{n}\frac{1}{2^k}$$

is a Cauchy sequence in X. Since X is Banach space, $\{z_n\}$ converges to a element $x\in X$ and $\|x\|<1$. From the equation (6.2.1), we have $Tz_n\to y$. Since T is continuous, we get $Tx=y$. Therefore, $y\in T(B_X(0,1))$. □

Theorem 6.2.6 (Open Mapping Theorem)**.** Suppose T is a bounded linear operator from a Banach space X onto a Banach space Y. Then T is an open map.

Proof. Let, G be an open subset of X. We have to show that $T(G)$ is open in Y. Let, $y\in T(G)$. Then we have a $x\in G$ such that $Tx=y$. Since G is open, there exists a $\epsilon>0$ such that $B_X(x,\epsilon)\subset G$. Thus, $B_X(0,\epsilon)\subset G\setminus\{x\}$.

By the above lemma, $\exists$ a $\delta > 0$ such that

$$\begin{array}{rcl} B_Y(0,\delta) & \subset & T(B_X(0,1)) \\ \Rightarrow \epsilon B_Y(0,\delta) & \subset & \epsilon T(B_X(0,1)) = T(B_X(0,1)) = T(B_X(0,\epsilon)) \\ & \subset & T(G \setminus \{x\}) = T(G) \setminus T(\{x\}) = T(G) \setminus y \ \text{[since } T \text{ is linear]} \end{array}$$

So, we get $B_Y(0, \epsilon\delta) + y \subset T(G)$. Therefore, y is an interior point of $T(G)$. Hence, $T(G)$ is open in Y. This proves the theorem. $\square$

6.2.2 Closed Graph Theorem

In this section, we introduce closed linear operators which appear more frequently in the application. In particular, most of the practical applications we encounter unbounded operators are closed linear operators.

Definition 6.2.7. Let X and Y be normed linear spaces. Then a linear operator $T : X \longrightarrow Y$ is said to be closed operator if for every sequence $\{x_n\}$ in X such that,

$$x_n \to x \quad \text{and}\ Tx_n \to y \ \Rightarrow Tx = y.$$

Definition 6.2.8 (Equivalent Definition)**.** Define a normed space $X \times Y$, where the two algebraic operations are defined as,

$$\begin{array}{rcl} (x_1, y_1) + (x_2, y_2) & = & (x_1 + x_2, y_1 + y_2) \\ \alpha(x, y) & = & (\alpha x, \alpha y) \end{array}$$

and the norm on $X \times Y$ is defined by

$$\|(x, y)\| = \|x\| + \|y\|.$$

Then a linear operator $T : X \longrightarrow Y$ is closed operator if the graph of T, $G(T) = \{(x, Tx) : x \in X\}$ is closed in $X \times Y$.

Example 6.2.9. Consider the differential operator $T : f \longrightarrow f'$ from $(C^{\perp}[a, b], \|.\|_\infty)$ to $(C[a, b], \|.\|_\infty)$. We know that, the operator is not continuous. Now we show that the operator is closed using uniform convergence property. Let $\{(f_n, f_n')\}$ be a sequence in $G(T)$ such that (f_n) converges to f and f_n' converges to g in sup-norm. We have to show that $g = f'$. Using fundamental theorem of integral calculus, we write

$$f_n(x) = f_n(a) + \int_a^x f_n'(t)dt$$
$$f(x) = f(a) + \int_a^x g(t)dt \quad (\text{as } n \to \infty)$$

The result follows by fundamental theorem of integral calculus.

Remark 6.2.10. Continuous linear operator $\Rightarrow$ Closed linear operator.

The converse is not true(see the above example). Under certain conditions, the converse is true which is stated as,

Theorem 6.2.11 (Closed Graph Theorem)**.** Statement: If X and Y are Banach spaces and $T : X \longrightarrow Y$ is linear operator, then

$$T \text{ is continuous } \Rightarrow T \text{ is closed.}$$

Proof. If T is continuous, then T is closed.

Conversely, suppose T is closed operator. Then the graph of T, $G(T)$ is closed in $X \times Y$. Moreover, it is a subspace and so it is a complete space.

Define $P : G(T) \longrightarrow X$ by $P(x, Tx) = x$. It is easy to verify that P is continuous, injective and surjective. By Bounded inverse theorem (6.2.12), $P^{-1} : X \longrightarrow G(T)$ is continuous, i.e., $\|P^{-1}(x)\| \leq c\|x\|, \forall x \in X$ for some $c > 0$. Hence T is bounded because of

$$\begin{aligned}
\|Tx\| \leq \|Tx\| + \|x\| &= \|(x, Tx)\| \\
&= \|P^{-1}(x)\| \\
&\leq c\|x\|, \forall x \in X.
\end{aligned}$$

This proves the theorem. □

Theorem 6.2.12 (Bounded inverse theorem)**.** If X and Y are Banach spaces and $T \in B[X, Y]$ is injective and surjective, then $T^{-1} \in B[Y, X]$.

Exercise 6.2.13.

1. Prove that, an operator T is a projection iff $T = T^*T$.
2. If P and Q are non-zero projections and $PQ = 0$, then show that $\|P + Q\| < \|P\| + \|Q\|$.
3. Show that, the null space $\mathcal{N}(T)$ of a closed linear operator $T : X \longrightarrow Y$ is closed subspace of X.
4. Let $(X, \|.\|_X)$ and $(Y, \|.\|_Y)$ be Banach spaces and $T : X \longrightarrow Y$ be a surjective linear operator from X onto Y such that $\exists c > 0 \forall x \in X : \|Tx\|_Y \leq c\|x\|_X$. Then T is bounded.

Unit 7

Course Structure

- Unbounded operator: Basic properties, Cayley transform, change of measure principle, spectral theorem.

7.0.1 Basic Properties

Definition 7.0.1. Let D be a subspace of a Hilbert space H. In this chapter D will almost never be closed. An unbounded operator T in H with domain D is a linear mapping from D into H. We will write $D(T)$ for the domain of T. T is densely defined if $D(T)$ is dense in H.

For an example, let $H = L^2[0,1]$, let $D = C^1[0,1]$ and let $Tf = f'$. Note that T is not a bounded operator. For another example, let $D = \{f \in C^2 : f(0) = f(1) = 0\}$ and $Uf = f''$. Then one can show that $\{-n^2\pi^2\}$ are eigenvalues.

Recall that $G(T)$, the graph of T, is the set $\{(x, Tx) : x \in D(T)\}$. If U is an extension of T, that means that $D(T) \subset D(U)$ and $Ux = Tx$ if $x \in D(T)$. Note that U will be an extension of T iff $G(T) \subset G(U)$. One often writes $T \subset$ U to mean that U is an extension of T.

A closed operator in H is one whose graph is a closed subspace of $H \times H$. This is equivalent to saying that whenever $x_n \to x$ and $Tx_n \to y$, then $x \in D(T)$ and $y = Tx$.

Proposition 7.0.2. If $D(T) = H$ and T is closed, then T is a bounded operator.

Proof. Recall the closed graph theorem, which says that if M is a closed linear map from a Banach space to itself, then M is bounded. The proposition follows immediately from this. ☐

Given a densely defined operator T, we want to define its adjoint T^*. First we define $D(T^*)$ to be the set of $y \in H$ such that the linear functional $l(x) = \langle Tx, y\rangle$ is continuous(i.e. bounded) on $D(T)$. If $y \in D(T^*)$, the Hahn-Banach theorem allows us to extend l to a bounded linear functional on H. By the Riesz representation theorem for Hilbert spaces, there exists $z_y \in H$ such that

$$l(x) = \langle x, z_y\rangle, \;\; x \in D(T)$$

Of course z_y depends on y. We then define $T^*y = z_y$.

Since T is densely defined, it is routine to check that T^* is well-defined and also that T^* is an operator in H, that is $D(T^*)$ is a subspace of H and T^* is linear.

For example, let $H = L^2[0,1]$, $D(T) = \{f \in C^1 : f(0) = f(1) = 0\}$, and $Tf = f'$. If $f \in D(T)$ and $g \in C^1$, then

$$\begin{aligned}
\langle Tf, g\rangle &= \int_0^1 f'(x)\overline{g}(x)dx \\
&= f(1)\overline{g}(1) - f(0)\overline{g}(0) - \int_0^1 f(x)\overline{g'}(x)dx \\
&= \langle f, \overline{g}\rangle
\end{aligned}$$

by applying integration by parts. Thus $|\langle Tf, g\rangle| \leq \|f\|\|g'\|$ is a bounded linear functional, and we see that $C^1 \subset D(T^*)$ and $T^*g = -g'$ if $g \in C^1$.

Some care is needed for the sum and composition of unbounded operators. We define

$$\begin{aligned}
D(S+T) &= D(S) \cap D(T) \\
\text{and } D(ST) &= \{x \in D(T) : Tx \in D(S)\}
\end{aligned}$$

Proposition 7.0.3. If S, T and ST are densely defined operators in H, then $T^*S^* \subset (ST)^*$. If in addition S is bounded, then

$$T^*S^* = (ST)^*.$$

Proof. Suppose $x \in D(ST)$ and $y \in D(T^*S^*)$. Since $x \in D(T)$ and $S^*y \in D(T^*)$, then

$$\langle Tx, S^*y\rangle = \langle x, T^*S^*y\rangle.$$

Since $Tx \in D(S)$ and $y \in D(S^*)$, then

$$\langle STx.y\rangle = \langle Tx, S^*y\rangle.$$

Assume now that S is bounded and $y \in D((ST)^*)$. Then S^* is also bounded and $D(S^*)$ is therefore equal to H. Hence,

$$\langle Tx, S^*y\rangle = \langle STx, y\rangle = \langle x, (ST)^*y\rangle$$

for every $x \in D(ST)$. Thus $S^*y \in D(T^*)$, and so $y \in D(T^*S^*)$. □

An operator T in H is symmetric if $\langle Tx, y\rangle = \langle x, Ty\rangle$ whenever x, y are both in $D(T)$. Thus a densely defined symmetric operator T is one such that $T \subset T^*$. If $T = T^*$, we say T is self-adjoint. Note that the domains of T and T^* are crucial here. This is not an issue with bounded operators because every symmetric bounded operator is self-adjoint.

Let us look at some examples. These will all be the same operator, but with different domains. Let $H = L^2[0,1]$. Let $D(S)$ be the set of absolutely continuous function f on $[0,1]$ such that $f' \in L^2$. Let $D(T)$ be the set of $f \in D(S)$ such that in addition $f(0) = f(1)$, and let $D(U)$ be the set of functions in $D(S)$ such that $f(0) = f(1) = 0$. Note that if $f' \in L^2$, then

$$|f(t) - f(s)| = \left|\int_s^t f'(x)dx\right| \leq \|f'\|_{L^2}|t-s|^{\frac{1}{2}}$$

by Cauchy-Schwarz inequality, so functions in any of these domains can be well-defined at points.

The operator will be the same in case:

$Sf = if'$, and the same for Tf and Uf provided f is in the appropriate domain. We see that U $\subset T \subset S$. We will show that T is self-adjoint, U is symmetric but not self-adjoint, and S is not symmetric.

By integration by parts,

$$\begin{aligned}\langle Tf, g\rangle &= \int_0^1 (if')\overline{g} \\ &= if(1)\overline{g}(1) - if(0)\overline{g}(0) - \int_0^1 if(\overline{g})' \\ &= if(1)\overline{g}(1) - if(0)\overline{g}(0) - \int_0^1 if(\overline{ig'}).\end{aligned}$$

Thus if $f, g \in D(T)$, we have $\langle Tf, g\rangle = \langle f, Tg\rangle$, since $f(1) = f(0)$ and $g(1) = g(0)$ for $f, g \in D(T)$.

Then some calculation with T replaced by S shows that S is not symmetric. The calculation with T replaced by U shows that U is symmetric. $\langle Tf, g\rangle = \displaystyle\int_0^1 (if')\overline{g}$ shows that $U \subset S^*$.

Suppose $g \in D(T^*)$ and $\phi = T^*g$. Let $\phi(x) = \displaystyle\int_0^x \phi(y)dy$. If $f \in D(T)$, then

$$\begin{aligned}\int_0^1 if'\overline{g} = \langle Tf, g\rangle &= \langle f, \phi\rangle \\ &= f(1)\overline{\phi}(1) - \int_0^1 f'\phi\end{aligned}$$

the last equality by 'integration by parts'. Since $D(T)$ contains non-zero constants, take f identically equal to 1 to conclude that $\phi(1) = 0$. Therefore we have $\displaystyle\int_0^1 f'\overline{G} = 0$ whenever $f \in D(T)$ and $G = ig - \phi$.

Taking the complex conjugate and replacing f by $\overline{f}$, $\displaystyle\int_0^1 f'\overline{G} = 0$ if $f \in D(T)$.

We claim that G is constant (a.e). Suppose $a < b$ is such that $[a, a+h]$, $[b, b+h]$ are both subsets of $[0, 1]$ and take f such that

$$f' = \frac{1}{h}\chi[a, a+h] - \frac{1}{h}\chi[b, b+h]$$

Then $f \in D(T)$ and so

$$\frac{1}{h}\int_a^{a+h} G(x)dx - \frac{1}{h}\int_b^{b+h} G(x)dx = 0$$

There is a set N of Lebesgue measure 0 such that if $y \in N$, then

$$\frac{1}{h}\int_y^{y+h} G(x)dx \to G(y)$$

So if $ab \notin N$, taking the limit shows $G(a) = G(b)$. Since we are on L^2, we can modify G on a set of Lebesgue measure 0 and take G constant.

This implies that $g = -i\phi + c$ is absolutely continuous and $g' = -i\phi \in L^2$. Also, $g(0) = -i\phi(0) + c = -i\phi(1) + c$, hence $g \in D(T)$. Thus $T^* \subset T$.

In the case of U: If $g \in D(U^*)$ and $f \in D(U)$, then $f(1) = 0$ and so

$$\int_0^1 if'\overline{g} = f(1)\overline{\phi}(1) - \int_0^1 f'\overline{\phi} = -\int_0^1 f'\overline{\phi}$$

If $G = ig - \phi$, then $\displaystyle\int_0^1 f'G = 0$. As before G is constant, so $g = -\phi + c$, but now we no longer know that $\phi(1) = 0$. So $g(1)$ might not be equal to $g(0)$. Therefore $U^* \subset S$.

If $g \in D(S)$ and $f \in D(V)$, we have

$$\begin{aligned} \langle Uf, g\rangle &= if(1)\overline{g}(1) - if(0)\overline{g}(0) + \int_0^1 f(\overline{ig'}) \\ &= \langle f, Ug\rangle. \end{aligned}$$

Hence $g \in D(U^*)$. Thus $S \subset U^*$, and with the above $U^* = S$. Hence U is not self-adjoint.

Proposition 7.0.4. Let H be a Hilbert space over $\mathbb{C}$, A is self-adjoint. Then A is closed.

Proof. A is closed: If $x_n \to x$ and $Ax_n \to u$, then $\langle Ax_n, y\rangle = \langle x_n, Ay\rangle \to \langle x, Ay\rangle = \langle Ax, y\rangle$.
Also $\langle Ax_n, y\rangle \to \langle u, y\rangle$. This is true for all y, so $Ax = u$. □

If A is defined on all of H and is self-adjoint, we conclude that A is bounded.
We say z is in the resolvent set of A if $A - zI$ maps D one-to-one onto H.

Proposition 7.0.5. If z is not real, then z is in the resolvent set. Equivalently $\sigma(A) \subset \mathbb{R}$.

Proof.

1. $R =$Range$(A - zI)$ is a closed subspace.

   R is equal to the set of all vectors u of the form $Av - zv = u$ for some $v \in D$. Then $\langle Av, v\rangle - z\langle v, v\rangle = \langle u, v\rangle$.

   A is self-adjoint, so $\langle Av, v\rangle = \langle v, Av\rangle = \overline{\langle Av, v\rangle}$ is real. Looking at the imaginary parts,

   $$-Im(z\|v\|^2) = Im\langle u, v\rangle$$

   So,

   $$\begin{aligned} |Imz|\|v\|^2 &\le \|u\|\|v\|, \\ \text{or, } \|v\| &\le \frac{1}{|Imz|}\|u\| \end{aligned}$$

   If $u_n \in R$ and $u_n \to u$, then $\|v_n - v_m\| \le \left(\frac{1}{|Imz|}\right)\|u_n - u_m\|$, so v_n is a Cauchy sequence, and hence converges to some point v.

   Since $Av_n - zv_n = u_n \to u$ and zv_n converges to zv, then Av_n converges to $u + zv$. Since A is self-adjoint, it is closed, and so $v \in D(A)$. Since $\langle Av_n, w\rangle = \langle v_n, Aw\rangle$ for $w \in D$, then $\langle u + zv, w\rangle = \langle v, Aw\rangle$, which implies and $Av = u + zv$, or $u = (A - z)v \in R$.

2. $R = H$. If not, there exists $x \neq 0$ such that x is orthogonal to R, and then

   $$\langle Av - zv, x\rangle = \langle Av, x\rangle - \langle v, \overline{z}x\rangle = 0$$

   for all $v \in D$. Then $\langle Av, x\rangle = \langle v, \overline{z}x\rangle$, so $x \in D$ and $Ax = \overline{z}x$. But then $\langle x, Ax\rangle = z\langle x, x\rangle$ is not real, a contradiction.

3. $A - zI$ is one-to-one. If not, there exists $x \in D$ such that $(A - zI)x = 0$.

   But then $\|x\| \le \left(\frac{1}{|Imz|}\right)\|0\| = 0$, or $x = 0$.

□

If we set $R(z) = (A - zI)^{-1}$ the resolvent, we have

$$\|R(z)\| \leq \frac{1}{|Imz|}.$$

If $u, w \in H$ and $v = R(z)u$, then $(A - z)v = u$, and

$$\begin{aligned}
\langle u, R(\overline{z}w\rangle &= \langle (A - z)v, R(\overline{z})w \\
&= \langle v, (A - \overline{z}R(\overline{z})w\rangle \\
&= \langle v, w\rangle \\
&= \langle R(z)u, w\rangle.
\end{aligned}$$

So the adjoint of $R(z)$ is $R(\overline{z})$.

Theorem 7.0.6. Let A be a symmetric operator. A is self-adjoint if and only if $\sigma(A) \subset \mathbb{R}$.

Proof. That A is self-adjoint implies that all non-real z are in the resolvent set has already been proved. We thus have to show that if A is symmetric and $\sigma(A) \subset \mathbb{R}$, then A is self-adjoint.

If $x, y \in D(A)$,

$$\langle (A - z)x, y\rangle = \langle x, (A - \overline{z})y\rangle.$$

If z is not real, then $z \notin \sigma(A)$, so $z - A$ is invertible and $A - z$ and $A - \overline{z}$ map $D(A)$ one-to-one and onto H. For $f, g \in H$, we can define $x = (A - z)^{-1}f$ and $y \in (A - \overline{z})g$, and we note that x and y are both in $D(A)$.

We then have

$$\langle f, (A - \overline{z})^{-1}g\rangle = \langle (A - z)^{-1}f, g\rangle$$

for all $f, g \in H$.

Now we show that A is self-adjoint. Take z as non-real and suppose $v \in D(A^*)$. Set $w = A^*v \in H$. We have,

$$\langle Ax, v\rangle = \langle x, A^*v\rangle$$

for all $x \in D(A)$. Subtract $z\langle x, v\rangle$ from both sides:

$$\langle (A - z)x, v\rangle = \langle x, (A^* - \overline{z})v\rangle$$

Let $g = (A^* - \overline{z})v$ and $f = (A - z)x$. Then

$$\begin{aligned}
\langle f, v\rangle &= \langle (A - z)x, v\rangle = \langle x, (A^* - \overline{z})v\rangle \\
&= \langle (A - z)^{-1}f, g\rangle = \langle f, (A - \overline{z})^{-1}g\rangle.
\end{aligned}$$

The set of f of the form $(A - z)x$ for $x \in D(A)$ is all of H, hence $v = (A - \overline{z})^{-1}$, which is in $D(A)$. In particular $D(A^*) \subset D(A)$. We have $(A - \overline{z})v = g = (A^* - \overline{z})v$, so $A^*v = Av$. $\square$

7.0.2 Cayley Transform

Definition 7.0.7. The mapping

$$t \to \frac{t - i}{t + i} \tag{7.0.1}$$

sets up a one-to-one correspondence between the real line and the unit circle. This shows that every self-adjoint $T \in \mathcal{B}(H)$ gives rise to a unitary operator

$$U = (T - iI)(T + iI)^{-1} \tag{7.0.2}$$

and that every unitary U whose spectrum does not contain the point 1 is obtained in this way.

This relation $T \leftrightarrow U$ will now be extended to a one-to-one correspondence between symmetric operators, on the one hand, and isometries, on the other.

Let T be a symmetric operator in H.

Then

$$\begin{aligned} \|Tx+ix\|^2 &= \|x\|^2 + \|Tx\|^2 \\ &= \|Tx-ix\|^2 \;\; [x \in \mathcal{D}(T)]. \end{aligned} \tag{7.0.3}$$

Hence there is an isometry U, with

$$\mathcal{D}(U) = R(T+iI); \;\; R(U) = R(T-iI) \tag{7.0.4}$$

defined by

$$U(Tx+ix) = Tx - ix \quad [x \in \mathcal{D}(T)]. \tag{7.0.5}$$

Since $(T+iI)^{-1}$ maps $\mathcal{D}(U)$ onto $\mathcal{D}(T)$, U can also be written in the form

$$U = (T-iI)(T+iI)^{-1}. \tag{7.0.6}$$

This operator U is called the Cayley transform of T.

Alternative Definition

Define $U = (A-i)(A+i)^{-1}$.

This is the image of the operator A under the function $F(z) = \dfrac{z-i}{z+i}$,

which maps the real line to $\delta B(0,1) \setminus \{1\}$, and is called the Cayley transform of A.

Proposition 7.0.8. U is a unitary operator.

Proof. $A+i$ and $A-i$ each map $D(A)$ one-to-one onto H, so U maps H onto itself.

U is norm preserving: Let $u \in H$, $v = (A+i)^{-1}u$, $w = Uu$. So $(A+i)v = u$, $(A-i)v = w$. We need to show $\|u\| = \|w\|$. We have,

$$\begin{aligned} \|u\|^2 &= \langle (A+i)v, (A+i)v \rangle \\ &= \|Av\|^2 + \|v\|^2 + i\langle v, Av\rangle - i\langle Av, v\langle \\ &= \|Av, v\|^2 + \|v\|^2 \end{aligned}$$

and similarly

$$\|w\| = \langle (A-i)v, (A-i)v = \|Av\|^2 + \|v\|^2.$$

□

Proposition 7.0.9. Given A and U as above and E the spectral resolution for U, $E(\{1\}) = 0$.

Proof. Write E_1 for $E(\{1\}) = 0$. If $E_1 \neq 0$, there exists $z \neq 0$ in the range of E_1, so $z = E_1 w$. Then

$$\mathrm{U}z = \int_{\sigma(\mathrm{U})} \lambda E(d\lambda) z = \int_{\sigma(\mathrm{U})} \lambda (E - E_1) d(\lambda) z + \int_{\{1\}} \lambda E_1(d\lambda) z$$

The first integral is zero since $(E - E_1)(A)$ and E_1 are orthogonal for all A.

The second integral is equal to

$$E_1 z = E_1 E_1 w = E_1 w = z$$

since E_1 is a projection.

We conclude z is an eigenvector for U with eigenvalue 1. So

$$(A - iI)(A + iI)^{-1} z = z$$

Let $v = (A + iI)^{-1} z$, or $z = (A + iI)v$.

Then,

$$z = (A - iI)(A + iI)^{-1} z = (A - iI)v,$$

and then $iv = -iv$, so $v = 0$ and hence $z = 0$, a contradiction. □

Lemma 7.0.10. Suppose U is an operator in H which is an isometry: $\|\mathrm{U}x\| = \|x\|$ for every $x \in \mathcal{D}(\mathrm{U})$.

a) If $x \in \mathcal{D}(\mathrm{U})$ and $y \in \mathcal{D}(\mathrm{U})$, then $(Ux, \mathrm{U}y) = (x, y)$.

b) If $\mathcal{R}(I - \mathrm{U})$ is dense in H, then $I - U$ is one-to-one.

c) If any one of the three spaces $\mathcal{D}(\mathrm{U})$, $\mathcal{R}(\mathrm{U})$ and $\zeta(\mathrm{U})$ is closed, so are the other two.

Theorem 7.0.11. Suppose U is the Cayley transform of a symmetric operator T in H. Then the following statements are true:

a) U is closed if and only if T is closed.

b) $\mathcal{R}(I - \mathrm{U}) = \mathcal{T}$, $I - \mathrm{U}$ is one-to-one, and T can be reconstructed from U by the formula

$$T = i(I + \mathrm{U})(I - \mathrm{U})^{-1}$$

(The Cayley transforms of distinct symmetric operators are therefore distinct).

c) U is unitary if and only if T is self-adjoint.

Conversely, if V is an operator in H which is an isometry, and if $I - V$ is one-to-one, then V is the Cayley transform of a symmetric operator in H.

Proof. T is closed if and only if $\mathcal{R}(T + iI)$ is closed. By the above lemma, U is closed iff $\mathcal{D}(\mathrm{U})$ is closed. Since $\mathcal{D}(\mathrm{U}) = \mathcal{R}(T + iI)$, by the definition of the Cayley transform, (a) is proved.

The one-to-one correspondence $x \leftrightarrow z$ between $\mathcal{D}(\mathrm{T})$ to $\mathcal{D}(\mathrm{U}) = \mathcal{R}(T + iI)$, given by

$$z = Tx + ix, \;\; \mathrm{U} = Tx - ix \tag{7.0.7}$$

can be rewritten in the form

$$(I - \mathrm{U})z = 2ix, \;\; (I + U)z = 2Tx \tag{7.0.8}$$

This shows that $I - \mathrm{U}$ is one-to-one, that $\mathcal{R}(T + iI) = \mathcal{D}(T)$, so that $(I - \mathrm{U})^{-1}$ maps $\mathcal{D}(T)$ onto $\mathcal{D}(\mathrm{U})$, and that

$$2Tx = (I + \mathrm{U})z = (I + \mathrm{U})(I - \mathrm{U})^{-1}(2ix) \quad [x \in \mathcal{D}(T)] \tag{7.0.9}$$

This proves (b).

Assume now that T is self-adjoint. Then

$$\mathcal{R}(I + T^2) = H \tag{7.0.10}$$

Since,
$$(T+iI)(T-iI) = I + T^2 = (T-iI)(T+iI) \tag{7.0.11}$$
[the three operators (7.0.11) have domain $\mathcal{D}(T^2)$]
it follows from (7.0.10) that

$$\mathcal{D}(\mathrm{U}) = \mathcal{R}(T+iI) = H \tag{7.0.12}$$
$$\text{and} \qquad \mathcal{R}(\mathrm{U}) = \mathcal{R}(T-iI) = H. \tag{7.0.13}$$

Since U is an isometry, (7.0.12) and (7.0.13) imply that U is unitary.
To complete the proof of (c), assume that U is unitary. Then

$$[\mathcal{R}(I-U)]^{\perp} = \mathcal{N}(I-U) = \{0\},$$

by (b) and the normality of $I-$ U, so that $\mathcal{D}(T) = \mathcal{R}(I-\mathrm{U})$ dense in H. Thus, T^* is defined, and $T \subset T^*$.
Fix $y \in \mathcal{D}(T^*)$. Since $\mathcal{R}(T+iI) = \mathcal{D}(\mathrm{U}) = H$, there exists $y_0 \in \mathcal{D}(T)$ such that

$$(T^*+iI)y = (T+iI)y_0 = (T^*+iI)y_0. \tag{7.0.14}$$

The last equality holds because $T \subset T^*$. If $y_1 = y - y_0$, then $y_1 \in \mathcal{D}(T^*)$ and for every $x \in \mathcal{D}(T)$.

$$((T-iI)x, y_1) = (x, (T^*+iI)y_1) = (x, 0) = 0 \tag{7.0.15}$$

Thus, $y \perp \mathcal{R}(T-iI) = \mathcal{R}(\mathrm{U}) = H$, and so $y_1 = 0$, and $y = y_0 \in \mathcal{D}(T)$.
Hence $T^* \subset T$, and (c) is proved.
Finally, let V be as in the statement of the converse. Then there is a one-to-one correspondence $z \leftrightarrow x$ between $\mathcal{D}(V)$ and $\mathcal{R}(I-V)$, given by
$$x = z - Vz. \tag{7.0.16}$$

Define S on $\mathcal{D}(S) = \mathcal{R}(I-V)$ by

$$Sx = i(z+Vz) \quad \text{if } x = z - Vz. \tag{7.0.17}$$

If $x \in \mathcal{D}(S)$ and $y \in \mathcal{D}(S)$, then $x = z - Vz$ and $y = u - Vu$ for some $z \in \mathcal{D}(V)$ and $u \in \mathcal{D}(V)$. Since V is an isometry, it now follows from (a) of the lemma, that

$$\begin{aligned}(Sx, y) &= i(z+Vz, u-Vu) = i(Vz, u) - i(z, Vu) \\ &= (z - Vz, iu + iVu) = (x, Sy)\end{aligned}$$

Hence S is symmetric. Since (7.0.17) can be written in the form

$$2iVz = Sx - ix, \quad 2iz = Sx + ix \quad [z \in \mathcal{D}(V)] \tag{7.0.18}$$

We see that
$$V(Sx+ix) = Sx - ix \quad [x \in \mathcal{D}(S)] \tag{7.0.19}$$

and that $\mathcal{D}(V) = \mathcal{R}(S+iI)$. Therefore V is the Cayley transform of S. □

7.0.3 Change of measure principle

Suppose

(a) $\mathcal{R}$ and $\mathcal{R}'$ are σ-algebras in sets Ω and Ω'

(b) $E : \mathcal{R} \longrightarrow \mathcal{B}(H)$ is a resolution of the identity, and

(c) $\phi : \Omega \longrightarrow \Omega'$ has the property that $\phi^{-1}(w') \in \mathcal{R}$ for every $w' \in \mathcal{R}'$.

If $E'(w') = E(\phi^{-1}(w'))$, then $E' : \mathcal{R}' \longrightarrow \mathcal{B}(H)$ is also a resolution of the identity, and

$$\int_{\Omega'} f dE'_{x,y} = \int_{\Omega} (f \circ \phi) dE_{x,y} \tag{7.0.20}$$

for every $\mathcal{R}'$-measurable $f : \Omega' \longrightarrow \mathbb{C}$ for which either of these integral exists.

7.0.4 Resolution of the Identity

Notation $\mathcal{R}$ will be a σ-algebra in a set Ω, H will be a Hilbert space, and $E : \mathcal{R} \longrightarrow \mathcal{B}(H)$ will be a resolution of the identity.

7.0.5 Spectral Theorem

Proposition 7.0.12. Let M be a bounded operator and f a measurable function. Let

$$D_f = \left\{ x : \int_{\sigma(M)} |f(x)|^2 \mu_{x,x}(d\lambda) < \infty \right\}$$

Then

(1) D_f is a dense subspace of H.

(2) If $x, y \in H$,

$$\int_{\sigma(M)} |f(x)|^2 \mu_{x,x}(d\lambda) \leq \|y\| \left(\int_{\sigma(M)} |f(\lambda)|^2 \times \mu_{x,x}(d\lambda) \right)^{\frac{1}{2}}$$

(3) If f is bounded and $v = f(M)z$, then

$$\mu_{x.v}(d\lambda) = \overline{f}(\lambda)\mu_{x,z}(d\lambda), \quad x, z \in H.$$

Proof. (1) Let $S \subset \sigma(M)$ and $z = x + y$.

$$\begin{aligned} \|E(S)z\|^2 &\leq (\|E(S)x\| + \|E(S)y\|^2) \\ &\leq 2\|E(S)x\|^2 + 2\|E(S)y\|^2 \end{aligned}$$

So,

$$\mu_{z,z}(S) \leq 2\mu_{x,x}(S) + 2\mu_{y,y}(S)$$

This is true for all S, so

$$\mu_{z,z}(d\lambda) \leq 2\mu_{x,x}(d\lambda) + 2\mu_{y,y}(d\lambda)$$

This proves that D_f is a subspace. Let, $S_n = \{\lambda \in \sigma(M) : |f(\lambda)| < n\}$. Then if $x = E(S_n)z$,

$$\begin{aligned} E(S)x &= E(S)E(S_n)E(S_n)z \\ &= E(S\cap S_n)E(S_n)z = E(S\cap S_n)x, \end{aligned}$$

so

$$\mu_{x,x}(S) = \mu_{x,x}(S\cap S_n).$$

Then

$$\begin{aligned} \int_{\sigma(M)} |f(x)|^2\mu_{x,x}(d\lambda) &= \int_{S_n} |f(x)|^2\mu_{x,x}(d\lambda) \\ &\leq n^2\|x\|^2 < \infty. \end{aligned}$$

To see this last line, we know it holds when $|f|^2$ is replaced by g and g is the characteristic function of a set. It holds for g simply by linearity, and then it holds for $g = |f|^2$ by monotone convergence. Therefore the range of $E(S_n) \subset D(f)$. $\sigma(M) = U_nS_n$, so

$$\begin{aligned} \|E(S_n)y - y\|^2 &= \|E(S_n)(y) - E(\sigma(M))(y)\|^2 \\ &= \int |\chi\sigma(M)\setminus S_n(\lambda)|^2\mu_{y,y}(d\lambda) \longrightarrow 0 \end{aligned}$$

by dominated convergence. Hence y is in the closure of D_f.

(2) If $x, y \in H$, f is bounded,

$$f(\lambda)\mu_{x,y}(d\lambda) << |f(\lambda)||\mu_{x,y}|(d\lambda),$$

so there exists u with $|u| = 1$ such that

$$u(\lambda)f(\lambda)\mu_{x,y}(d\lambda) = |f(\lambda)||\mu_{x,y}|(d\lambda).$$

Hence

$$\begin{aligned} \int_{\sigma(M)} |f(x)|\mu_{x,x}(d\lambda) &= (\mu f(M)x, y) \\ &\leq \|uf(M)x\|\|y\|. \end{aligned}$$

But

$$\|uf(M)x\|^2 = \int |uf|^2 d\mu_{x,x} = \int |f|^2 d\mu_{x,x}.$$

So, (2) holds for bounded f. Now take a limit and use monotone convergence.

(3) Let g be continuous.

$$\begin{aligned} \int_{\sigma(M)} g d\mu_{x,v} &= (g(M)x, v) = (g(M)x, f(M)z) \\ &= ((\overline{f}g)(M)(x), z) = \int g\overline{f} d\mu_{x,z} \end{aligned}$$

this is true for all g continuous, so $d\mu_{x,x} = \overline{f}d\mu_{x,z}$.

$\square$

Theorem 7.0.13. Let E be a resolution of the identity.

(1) Suppose $f : \sigma(M) \longrightarrow \mathbb{C}$ is measurable. There exists a densely defined operator $f(M)$ with domain D_f and

$$\begin{aligned} \langle f(M)x, y\rangle &= \int_{\sigma(M)} f(\lambda)\mu_{x,y}(d\lambda) \\ \|f(M)x\| &= \int_{\sigma(M)} |f(\lambda)|^2\mu_{x,x}(d\lambda) \end{aligned} \tag{7.0.21}$$

(2) If $D_{fg} \subset D_g$, then $f(M)g(M) = (fg)(M)$.

(3) $f(M)^* = \overline{f}(M)$ and $f(M)f(M)^* = f(M)^*f(M) = |f|^2M$.

Proof. (1) If $x \in D_f$, then $l(y) = \int_{\sigma(M)} f d\mu_{x,y}$ is a bounded linear functional with norm at most $(\int |f|^2 d\mu_{x,x})^{\frac{1}{2}}$ by (2) of the preceding proposition. Choose $f(M)x \in H$ to satisfy (1) for all y.

Let $f_n = f\chi(|f| \le n)$. Then $D_f - f_n = D_f$ since $\int |f - f_n|^2 d\mu_{x,x}$ is finite if and only if $\int |f|^2 d\mu_{x,x}$ is finite, using that f_n is bounded. By the Dominated convergence theorem,

$$\|f(M)x - f_n(M)x\|^2 \le \int_{\sigma(M)} |f - f_n|^2 d\mu_{x,x} \to 0.$$

Since f_n is bounded, (7.0.21) holds with f_n.

Now let $n \to \infty$.

(2) : Define $g_m = g\chi(|g| \le m)$. Since f_n and g_m are bounded, (2) follows for f_n, g_m. Now let $m \to \infty$ and then $n \to \infty$.

(3) We know this holds for f_n since f_n is bounded. Now let $n \to \infty$.

□

Theorem 7.0.14 (Spectral Theorem)**.** Let A be a self-adjoint operator on a Hilbert space over the complex numbers. There exists a resolution of the identity E such that, $A = \int_{\sigma(A) zE(dz)}$.

Proof. Start with the unbounded operator A. Let $\mathrm{U} = (A - iI)(A + iI)^{-1}$. Then U is unitary with a spectrum on $\delta B(0, 1) \setminus \{1\}$. Let the resolution of the identity for U be given by $\tilde{E}$.

Let us define $\phi = F^{-1}$, which is a map taking $\delta B(0, 1) \setminus \{1\}$ to $\mathbb{R}$. Thus

$$\phi(z) = \frac{i(1 + z)}{1 - z}.$$

We check that $A = \phi(\mathrm{U})$. Since the range of ϕ is $\mathbb{R}$, then $\phi(\mathrm{U})$ is self-adjoint. Since $\phi(z)(1 - z) = i(1 + z)$, the above theorem implies that,

$$\phi(\mathrm{U})(I - \mathrm{U}) = i(I + \mathrm{U}).$$

In particular, the range of $I - \mathrm{U}$ is contained in the domain of $\phi(\mathrm{U})$. From the definition of the Cayley transform, we have

$$A(I - \mathrm{U}) = i(I + \mathrm{U})$$

and the domain of A is equal to the range of $I - \mathrm{U}$. Thus $A \subset \phi(\mathrm{U})$. Since both A and $\phi(\mathrm{U})$ are self-adjoint,

$$\phi(\mathrm{U}) = \phi(\mathrm{U})^* \subset A^* = A \subset \phi(\mathrm{U}),$$

and hence $A = \phi(\mathrm{U})$.

Let $E(S) = \tilde{E}(\phi^{-1}(S))$. We have

$$\begin{aligned} \langle Ax, y \rangle &= \langle \phi(\mathrm{U}x, y \rangle \\ &= \int_{\sigma(\mathrm{U})} \phi(z) \langle \tilde{E}(dz)x, y \rangle. \end{aligned}$$

By the change of measure principle, this is equal to $\displaystyle\int_{\sigma(A)} z\langle E(dz)x, y\rangle$. □

Exercise 7.0.15. 1. The associative law $(T_1T_2)T_3 = T_1(T_2T_3)$ has been used freely throughout this chapter. Prove it. Prove also that $T_1 \subset T_2$ implies $ST_1 \subset ST_2$ and $T_1S \subset T_2S$.

2. Suppose T is densely defined, closed operator in H, and $T^*T \subset TT^*$. Does it follow that T is normal?

3. Suppose T is densely defined operator in H, and $(Tx, x) = 0$ for every $x \in \mathcal{D}(T)$. Does it follow that $Tx = 0$ for every $x \in \mathcal{D}(T)$?

4. Let H^2 be the space of all holomorphic functions $f(z) = \sum c_n z^n$ in the open unit disc that satisfy,

$$\|f\|^2 = \sum_{n=0}^{\infty} |c_n|^2 < \infty$$

Define V now by, $(Vf)(z) = zf(z^2)$. Show that V is an isometry which is the Cayley transform of a closed symmetric operator T in H^2, whose deficiency indices are 0 and ∞.

5. Suppose T is a closed operator in H, $\mathcal{D}(T) = \mathcal{D}(T^*)$, and $\|Tx\| = \|T^*x\|$ for every $x \in \mathcal{D}(T)$. Prove that T is normal.

[Hint: Begin by proving that $(Tx, Ty) = (T^*x, T^*y), x \in \mathcal{D}(T), y \in \mathcal{D}(T)$]

Unit 8

Course Structure

- Compact map: Basic properties, compact symmetric operator, Rayleigh principle, Fisher's principle, Courant's principle, Mercer's theorem, positive compact operator.

8.0.1 Basic Properties

A subset S is precompact if $\overline{S}$ is compact. Recall that if A is a subset of a metric space, A is precompact iff every sequence in A has subsequene which converges in $\overline{A}$. Also, A is compact iff A is complete and totally bounded.

Write B_1 for the unit ball in X.

A map K from a Banach space X to a Banach space U is compact if $K(B_1)$ is precompact in U.

One example is if K is degenerate, so that R_K is finite dimensional. The identity on l^2 is not compact.

The following facts are easy:

1) If C_1, C_2 are precompact subsets of a Banach space, then $C_1 + C_2$ is precompact.

2) If C is precompact, so is the convex hull of C.

3) If $M : X \longrightarrow$ U and C is precompact in X, then $M(C)$ is precompact in U.

Proposition 8.0.1. (i) If K_1 and K_2 are compact maps, so is $KK_1 + K_2$.

(ii) If $X \xrightarrow{L}$ U $\xrightarrow{M}$, where M is bounded and L is compact, then ML is compact.

(iii) In the same situation as (ii), if L is bounded and M is compact

(iii) In the same situation as (ii), if L is bounded and M is compact, then ML is compact.

(iv) If K_n are compact maps and $lim\|K_n - K\| = 0$, then K is compact.

Proof. (i) For the sum, $(K_1 + K_2)(B_1) \subset K_1(B_1) + K_2(B_2)$ and the multiplication by K is similar.

(ii) $ML(B_1)$ will be compact because $L(B_1)$ is compact and M is continuous.

(iii) $L(B_1)$ will be contained in some ball, so $ML(B_1)$ is precompact.

(iv) Let $\epsilon > 0$. Choose n such that $\|K_n - K\| < \epsilon$. $K_n(B_1)$ can be covered by finitely many balls of radius ϵ, so $K(B_1)$ is covered by the set of balls with the same centres and radius 2ϵ. Therefore $K(B_1)$ is totally bounded.

$\square$

We can use (iv) to give a more complicated example of a compact operator.
Let, $X = U = l^2$ and define

$$K(a_1, a_2, \cdots) = \left(\frac{a_1}{2}, \frac{a_2}{2^2}, \frac{a_3}{2^3}, \cdots\right).$$

It is the limit in norm of K_n, where

$$K_n(a_1, a_2, \cdots) = \left(\frac{a_1}{2}, \frac{a_2}{2^2}, \frac{a_3}{2^3}, \cdots, \frac{a_n}{2^n}, 0, \cdots\right).$$

Note that any bounded operator K on l^2 maps B_1 into a set of the form $[-M, M]^{\mathbb{N}}$. By Tychonoff, this is compact in the product topology. However it is not necessarily compact in the topology of the space l^2.

Proposition 8.0.2. If X and Y are Banach spaces and $K : X \longrightarrow Y$ is compact and Z is a closed subspace of X, then the map $K|_Z$ is compact.

Let A be a bounded linear operator on a Banach space. If z is a complex number and I is the identity operator on H which might or might not be invertible. We define the spectrum of A by

$$\sigma(A) = \{z \in \mathbb{C} : I - A \text{ is not invertible}\}.$$

We sometimes write $z - A$ for $zI - A$. The resolvent set for A is the set of complex numbers z such that $z - A$ is invertible. A non-zero element z is an eigenvector for A with corresponding eigenvalue λ if $Az = \lambda z$.

Compact Symmetric Operator

If A is bounded operator on H, a Hilbert space over the complex numbers, the adjoint of A, denoted by A^*, is the operator on H such that $\langle Ax, y\rangle = \langle x, A^*y\rangle, \forall x$ and y.

It follows from the definition that the adjoint of cA is $\bar{c}A^*$ and the adjoint of A^n is $(A^*)^n$.

If $P(x) = \sum_{j=0}^{n} a_j A^j$ will be $\overline{P}(A^*) = \sum_{j=0}^{n} \overline{a_j} P(A^*)$.

The adjoint operator always exists.

Proposition 8.0.3. If A is a bounded operator on H, there exist a unique operator A^* such that $\langle Ax, y\rangle = \langle x, A^*y\rangle, \forall x$ and y.

Proof. Fix y for the moment. The function $f(x) = \langle Ax, y\rangle$ is a linear functional on H. By the Riesz representation theorem for Hilbert spaces, there exists z_y such that $\langle Ax, y\rangle = \langle x, z_y\rangle, \forall x$. Since,

$$\begin{aligned} \langle x, z_{y_1+y_2}\rangle &= \langle Ax, y_1 + y_2\rangle \\ &= \langle Ax, y_1\rangle + \langle Ax, y_2\rangle \\ &= \langle x, z_{y_1}\rangle + \langle x, z_{y_2}\rangle \end{aligned}$$

for all x, then $z_{y_1+y_2} = z_{y_1} + z_{y_2}$ and similarly $z_{cy} = cz_y$. If we define $A^*y = z_y$, this will be the operator we seek.

If A_1 and A_2 are two operators such that $\langle x, A_1y\rangle = \langle Ax, y\rangle = \langle x, A_2y\rangle, \forall x$ and y, then $A_1y = A_2y, \forall y$, so $A_1 = A_2$. Thus the uniqueness assertion is proved. $\square$

A bounded linear operator A mapping H into H is called symmetric if

$$\langle Ax.y\rangle = \langle x, Ay\rangle \tag{8.0.1}$$

for all x and y in H. Other names for symmetric are Hermitian or self-adjoint. When A is symmetric, then $A^* = A$, which explains the name "self-adjoint".

Example 8.0.4. For an example of a symmetric bounded linear operator, let (X, A, μ) be a measure space with μ and σ-finite measure, let $H = L^2(X)$, and let $F(x, y)$ be a jointly measurable function from $X \times X$ into $\mathbb{C}$ such that $F(y, x) = \overline{F(x, y)}$ and

$$\int\int F(x, y)^2 \mu(dx)\mu(dy) < \infty. \tag{8.0.2}$$

Define $A : H \longrightarrow H$ by

$$Af(x) = \int F(x, y)f(y)\mu(dy). \tag{8.0.3}$$

You can check that A is a bounded symmetric operator.

Here is an example of a compact symmetric operator.

Example 8.0.5. Let $H = L^2([0, 1])$ and let $F : [0, 1]^2 \longrightarrow \mathbb{R}$ be a continuous function with $F(x, y) = \overline{F(y, x)}$ for all x and y. Define $K : H \longrightarrow H$ by

$$Kf(x) = \int_0^1 F(x, y)f(y)dy.$$

We discussed in previous example that K is a bounded symmetric operator. Let us show that it is compact.

If $f \in L^2([0, 1])$ with $\|f\| \leq 1$, then

$$\begin{aligned} |Kf(x) - Kf(x')| &= \left|\int_0^1 [F(x, y) - F(x'y)]f(y)dy\right| \\ &\leq \left(\int_0^1 |F(x, y) - F(x'y)|^2 dy\right)^{\frac{1}{2}} \|f\|, \end{aligned}$$

using the Cauchy-Schwarz inequality. Since F is continuous on $[0, 1]^2$, which is a compact set, then it is uniformly continuous there.

Let $\epsilon > 0$. There exists δ such that

$$\sup_{|x-x'|<\delta} \sup_y |F(x, y) - F(x', y)| < \epsilon.$$

Hence if $|x-x'| < \delta$, then $|Kf(x)-Kf(x')| < \epsilon$ for every f with $\|f\| \leq 1$. In other words, $\{Kf : \|f\| \leq 1\}$ is an equicontinuous family.

Since F is continuous, it is bounded, say by N, and therefore

$$|Kf(x)| \leq \int_0^1 N|f(y)|dy \leq N\|f\|,$$

again using the Cauchy-Schwarz inequality.

If Kf_n is a sequence in $K(B_1)$, then $\{Kf_n\}$ is a bounded equicontinuous family of functions on [0,1], and by the Ascoli-Arzela theorem, there is a subsequence which converges uniformly on [0,1]. It follows that this subsequence also converges with respect to the L^2 norm. Since every sequence in $K(B_1)$ has a subsequence which converges, the closure of $K(B_1)$ is compact. Thus K is a compact operator.

We have the following proposition.

Proposition 8.0.6. Suppose A is a bounded symmetric operator.

(1) (Ax, x) is real for all $x \in H$.

(2) The function $x \longrightarrow \langle Ax, x\rangle$ is not identically 0 unless $A = 0$.

(3) $\|A\| = \sup_{\|x\|=1} |\langle Ax, x\rangle$.

Proof. (1) This one is easy since

$$\langle Ax, x\rangle = \langle x, Ax\rangle = \overline{\langle Ax, x\rangle}$$

where we use $\overline{z}$ for the complex conjugate of z.

(2) If $\langle Ax, x\rangle = 0$ for all x, then

$$\begin{aligned} 0 = \langle A(x+y), x+y\rangle &= \langle Ax, x\rangle + \langle Ay, y\rangle \\ &= \langle Ax, y\rangle + \langle y, Ax\rangle \\ &= \langle Ax, y\rangle + \overline{\langle Ax, y\rangle}. \end{aligned}$$

Hence, $\mathrm{Re}\langle Ax, y\rangle = 0$. Replacing x by ix and using linearity,

$$\begin{aligned} Im(\langle Ax, y\rangle) &= -Re(i\langle Ax.y\rangle) \\ &= -Re(\langle A(ix), y\rangle) = 0. \end{aligned}$$

Therefore, $\langle Ax, y\rangle = 0$ for all x and y. We conclude that $Ax = 0$ for all x and thus $A = 0$.

(3) Let $\beta = \sup\limits_{\|x\|=1} |\langle Ax, x\rangle$. By the Cauchy Schwarz inequality,

$$|\langle Ax, x\rangle| \leq \|Ax\|\|x\| \leq \|A\|\|x\|^2$$

so, $\beta \leq \|A\|$.

To get the other direction, let $\|x\| = 1$ and let $y \in H$ such that $\|y\| = 1$ and $\langle y, Ax\rangle$ is real. Then

$$\langle y, Ax\rangle = \frac{1}{4}(\langle x+y, A(x+y)\rangle - \langle x-y, A(x-y)\rangle).$$

We used that

$$\langle y, Ax\rangle = \langle Ay, x\rangle = \langle Ax, y\rangle = \langle x, Ay\rangle,$$

since $\langle y, Ax\rangle$ is real and A is symmetric.

Then

$$\begin{aligned} 16|\langle y, Ax\rangle|^2 &\leq \beta^2(\|x+y\|^2 + \|x-y\|^2)^2 \\ &= 4\beta^2(\|x\|^2 + \|y\|^2)^2 \\ &= 16\beta^2. \end{aligned}$$

We used the parallelogram law $[\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2]$ in the 1st equality. We conclude $|\langle y, Ax\rangle| \leq \beta$.

If $\|y\| = 1$ but $\langle y, Ax\rangle = re^{i\theta}$ is not real. Let $y' = e^{-i\theta}y$ and apply the above with y' instead of y. We then have

$$|\langle y, Ax\rangle| = |\langle y', Ax\rangle| \leq \beta.$$

Setting $y = \frac{Ax}{\|Ax\|}$, we have $\|Ax\| \leq \beta$. Taking the supremum over x with $\|x\| = 1$ we conclude $\|A\| \leq \beta$.

$\square$

If $(Ax, x) \geq 0, \forall x$, we say A is positive, and write $A \geq 0$. Writing $A \leq B$ means $B - A \geq 0$. For matrices, one uses the words "positive definite".

Now suppose A is compact.

Proposition 8.0.7. If $x_n \xrightarrow{w}$, then $Ax_n \xrightarrow{s}$.

Proof. If $x_n \xrightarrow{w} x$, then $Ax_n \xrightarrow{w} Ax$, since $\langle Ax_n, y\rangle = \langle x_n, Ay\rangle \to \langle x, Ay\rangle = \langle Ax, y\rangle$.

If x_n converges weakly, then $\|x_n\|$ is bounded so Ax_n lies in a precompact set.

Any subsequence of Ax_n has a further subsequence which converges strongly. The limit must be Ax. $\square$

Lemma 8.0.8. If K is a compact operator and $\{x_n\}$ is a sequence with $\|x_n\| \leq 1$ for each n, then $\{Kx_n\}$ has a convergent subsequence.

Proof. Since $\|x_n\| \leq 1$, then $\left\{\frac{1}{2}x_n\right\} \subset B_1$. Hence $\left\{\frac{1}{2}Kx_n\right\} = \left\{K\left(\frac{1}{2}x_n\right)\right\}$ is a sequence contained in $K(B_1)$, a compact set and therefore has a convergent subsequence. $\square$

We now prove the spectral theorem for compact symmetric operators.

Theorem 8.0.9. Suppose H is a separable Hilbert space over the complex numbers and K is a compact symmetric linear operator. There exist a sequence $\{z_n\}$ in H and a sequence $\{\lambda_n\}$ in $\mathbb{R}$ such that

(1) $\{z_n\}$ is an orthonormal basis for H.

(2) each z_n is an eigenvector with eigenvalue λ_n, i.e. $Kz_n = \lambda_n z_n$.

(3) for each $\lambda_n \neq 0$, the dimension of the linear space $\{x \in H : Kx = \lambda_n x\}$ is finite.

(4) the only limit point, if any, of $\{\lambda_n\}$ is 0; if there are infinitely many distinct eigenvalues, then 0 is a limit point of $\{\lambda_n\}$.

Note that part of the assertion of the theorem is that the eigenvalues are real. (3) is usually phrased as saying the non-zero eigenvalues have finite multiplicity.

Proof. If $K = 0$, any orthonormal basis will do for $\{z_n\}$ and all the λ_n are zero, so we suppose $K \neq 0$. We first show that the eigenvalues are real, that eigenvectors corresponding to distinct eigenvalues are orthogonal, the multiplicity of non-zero eigenvalues is finite, and that 0 is the only limit point of the set of eigenvalues. We then show how to sequentially construct a set of eigenvectors and that this construction yields a basis.

If λ_n is an eigenvalue corresponding to a eigenvector $z_n \neq 0$, we see that,

$$\begin{aligned}\lambda_n\langle z_n, z_n\rangle &= \langle\lambda_n z_n, z_n\rangle = \langle Kz_n, z_n\rangle = \langle z_n, Kz_n\rangle \\ &= \langle z_n, \lambda_n z_n\rangle = \overline{\lambda_n}\langle z_n, z_n\rangle,\end{aligned}$$

which proves that λ_n is real.

If $\lambda_m \neq \lambda_n$ are two distinct eigenvalues corresponding to the eigenvectors z_n and z_m, we observe that

$$\begin{aligned}\lambda_n\langle z_n, z_m\rangle &= \langle\lambda_n z_n, z_m\rangle = \langle Kz_n, z_m\rangle = \langle z_n, Kz_m\rangle \\ &= \langle z_n, \lambda_m z_m\rangle = \lambda_m\langle z_n, z_m\rangle\end{aligned}$$

using that λ_m is real. Since $\lambda_n \neq \lambda_m$, we conclude $\langle z_n, z_m \rangle = 0$.

Since $\lambda_n \neq 0$ and that there are infinitely many orthonormal vectors x_k such that $Kx_k = \lambda_n x_k$. Then

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

if $j \neq k$. But then no subsequence of $\lambda_n x_k = kx_k$ can converge, a contradiction to the above lemma. Therefore, the multiplicity of λ_n is finite.

Suppose we have a sequence of distinct non-zero eigenvalues converging to a real number $\lambda \neq 0$ and a corresponding sequence of eigenvectors each with norm one. Since K is compact, there is a subsequence $\{n_j\}$ such that Kz_{n_j} converges to a point in H, say w. Then

$$z_{n_j} = \frac{1}{\lambda_{n_j}} K z_{n_j} \to \frac{1}{\lambda} w.$$

or $\{z_{n_j}\}$ is an orthonormal sequence of vectors converging to $\lambda^{-1}w$. But as in the preceeding paragraph, we cannot have such a sequence.

Since $\{\lambda_n\} \subset \overline{B(0, r(K))}$, a bounded subset of the complex plane, if the set $\{\lambda_n\}$ is infinite, there will be a subsequence which converges. By the preceeding paragraph, 0 must be a limit point of the subsequence.

We now turn to constructing eigenvectors. We know that $\|K\| = \sup_{\|x\|=1} |\langle Kx, x \rangle|$.

We claim the maximum is attained. If $\sup_{\|x\|=1} |\langle Kx, x \rangle| = \|K\|$, let $\lambda = \|K\|$; otherwise let $\lambda = -\|K\|$. Choose x_n with $\|x_n\| = 1$ such that $\langle Kx_n, x_n \rangle$ converges to λ. There exists a subsequence $\{n_j\}$ such that Kx_{n_j} converges, say to z. Since $\lambda \neq 0$, then $z \neq 0$, for otherwise

$$\lambda = \lim_{j \to \infty} \langle Kx_{n_j}, x_{n_j} \rangle = 0$$

Now,

$$\begin{aligned} \|(K - \lambda I)z\|^2 &= \lim_{j \to \infty} \|(K - \lambda I)Kx_{n_j}\|^2 \\ &\leq \|K\|^2 \lim_{j \to \infty} \|(K - \lambda I)x_{n_j}\|^2 \end{aligned}$$

and,

$$\begin{aligned} \|(K - \lambda I)z\|^2 &= \|Kx_{n_j}\|^2 + \lambda^2 \|x_{n_j}\|^2 - 2\lambda \langle x_{n_j}, Kx_{n_j} \rangle \\ &\leq \|K\|^2 + \lambda^2 - 2\lambda \langle x_{n_j}, Kx_{n_j} \rangle \\ &\to \lambda^2 + \lambda^2 - 2\lambda^2 = 0. \end{aligned}$$

Therefore, $(K - \lambda I)z = 0$, or z is an eigenvector for K with corresponding eigenvalue λ.

Suppose, we have found eigenvalues $z_1, z_2, \cdots, z_n$. Let X_n be the linear subspace spanned by $\{z_1, z_2, \cdots, z_n\}$ and let $Y = X_n^{\perp}$ be the orthogonal complement of X_n, that is, the set of all vectors orthogonal to every vector in X_n. If $x \in Y$ and $K \leq n$, then

$$\langle Kx, z_k \rangle = \langle x, Kz_k \rangle = \overline{\lambda_k}(x, z_k) = 0,$$

or $Kx \in Y$. Hence K maps Y into Y. It is an exercise to show that $K|_Y$ is a compact symmetric operator. If Y is non-zero, we can then look at $K|_Y$, and find a new eigenvector z_{n+1}.

It remains to prove that the set of eigenvectors forms a basis. Suppose y is orthogonal to every eigenvector. Then

$$\langle Ky, z_k \rangle = \langle y, Kz_k \rangle = \langle y, \lambda_k z_k \rangle = 0.$$

If z_k is an eigenvector with eigenvalue λ_k, so Ky is also orthogonal to every eigenvector. Suppose X is the closure of the linear subspace spanned by $\{z_k\}$, $Y = X^{\perp}$, and $Y \neq \{0\}$. If $y \in Y$, then $\langle Ky, z_k\rangle = 0$ for each eigenvector z_k, hence $\langle Ky, z\rangle = 0$ for every $z \in X$, or $K : Y \longrightarrow Y$. Thus $K|_Y$ is a compact symmetric operator, and by the argument already given, there exists an eigenvector for $K|_Y$. This is a contradiction since Y is orthogonal to every eigenvector. □

Remark 8.0.10. If $\{z_n\}$ is an orthonormal basis of eigenvectors for K with corresponding eigenvalues λ, let E_n be the projection onto the subspace spanned by z_n, i.e., $E_n x = \langle x, z_n\rangle z_n$. A vector x can be written as $\sum_n \langle x, z_n\rangle z_n$, thus $Kx = \sum_n \lambda_n \langle x, z_n\rangle z_n$. We can then write, $K = \sum_n \lambda_n E_n$.

For general bounded symmetric operators there is a related expansion where the sum gets replaced by an integral.

Remark 8.0.11. If z_n is eigenvector for K with corresponding eigenvalue λ_n, then $Kz_n = \lambda_n z_n$, so

$$K^2 z_n = K(Kz_n) = K(\lambda_n z_n) = \lambda_n K z_n = (\lambda_n)^2 z_n$$

More generally, $K^j z_n = (\lambda_n)^j z_n$. Using the notation of the above Remark, we can write

$$K^j = \sum_n (\lambda)^j E_n.$$

If Q is any polynomial, we can then use linearity to write, $Q(K) = \sum_n Q(\lambda_n)E_n$.

It is a small step from here to make the definition $f(K) = \sum_n f(\lambda_n)E_n$ for any bounded and Borel measurable function f.

If $\alpha_1 \geq \alpha_2 \geq \cdots > 0$ and $Az_n = \alpha_n z_n$, then our construction shows that

$$\alpha_N = \max_{x + z_1, z_2, \cdots, z_{N-1}} \frac{\langle Ax, x\rangle}{\|x\|^2}.$$

This is known as the Rayleigh principle.

Let,

$$R_A(x) = \frac{\langle Ax, x\rangle}{\|x\|^2}.$$

Proposition 8.0.12. Let A be compact and symmetric and let α_k be the non-negative eigenvalues with $\alpha_1 \geq \alpha_2 \geq \cdots$. Then

(1) (Fisher's principle)

$$\alpha_N = \max_{S_N} \min_{x \in S_N} R_A(x)$$

where the maximum is over all linear subspaces S_N of dimension N.

(2) (Courant's principle)

$$\alpha_N = \min_{S_{N-1}} \max_{x \perp S_{N-1}} R_A(x)$$

where the minimum is over all linear subspaces of dimension $N - 1$.

Proof. Let $z_1, z_2, \cdots, z_N$ be eigenvectors with corresponding eigenvalues $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_N$. Let T_N be the linear subspace spanned by $\{z_1, z_2, \cdots, z_N\}$. If $y \in T_N$, we have $y = \sum_{j=1}^{N} c_j z_j$ for some complex numbers c_j and then

$$\begin{aligned}
\langle Ay, y\rangle &= \sum_{i=1}^{N}\sum_{j=1}^{N} c_i\overline{c_j}\langle Az_i, z_j\rangle = \sum_i\sum_j c_i\overline{c_j}\alpha_i\langle z_i, z_j\rangle \\
&= \sum_i |c_i|^2\alpha_i \geq \sum_i |c_i|^2\alpha_N \\
&= \langle y, y\rangle.
\end{aligned}$$

Using the fact that the z_i's are orthogonal by our construction.

(1) Let z_k be the eigenvectors. Let S_N be a subspace of dimension N. There exists $y \in S_N$ such that $\langle y, z_k\rangle = 0$ for $k = 1, \cdots, N-1$. Since

$$\alpha_N = \max_{x\perp z_1,\cdots,z_{N-1}} R_A(x)$$

then y is one of the vectors over which the max is being taken, so $R_A(y) \leq \alpha_N$ for this y. So, $\min_{x\in S_N} R_A(x) \leq \alpha_N$. This is true for all spaces of dimension N. So, the right hand side is less than or equal to α_N.

Now we show the right hand side is greater than or equal to α_N. Let S_N be the linear span of $\{z_1, \cdots, z_N\}$. By the first paragraph of the proof, $R_A(x) \geq \alpha_N$ for every $x \in S_N$, and $R_A(x) = \alpha_N$ when $x = z_N$. So, $\min_{x\in S_N} R_A(x) = \alpha_N$. The maximum over all subspaces of dimension N will be larger than the value for this particular subspace, so the right hand side is atleast as large as α_N.

(2) Let S_{N-1} be a subspace of dimension $N-1$ and let T_N be the span of $\{z_1, \cdots, z_N\}$. Since the dimension of T_N is larger than that of S_{N-1}, there must be a vector $y \in T_N$ perpendicular to S_{N-1}. Since $y \in T_N$, then $R_A(y) \geq \alpha_N$ by the first paragraph of this proof, so

$$\max_{x\perp S_{N-1}} R_A(x) \geq R_A(y) \geq \alpha_N.$$

Taking the minimum over all spaces S_{N-1} shows that right hand side is greater than or equal to α_N.

If $x \perp T_{N-1}$, then $x = \sum_{j=N+1}^{\infty} c_j z_j$, and then

$$\begin{aligned}
\langle Ax, x\rangle &= \sum_{j=N}^{\infty}\sum_{k=N}^{\infty} c_j\overline{c_k}\alpha_j\langle z_j, z_k\rangle \\
&= \sum_{j=N}^{\infty}\alpha_j|c_j|^2 \leq \alpha_N\sum_{j=N}^{\infty}|c_j|^2 \\
&= \alpha_N\langle x, x\rangle.
\end{aligned}$$

Therefore $R_A(x) \leq \alpha_N$. This leads to

$$\min_{S_{N-1}}\max_{x\perp S_{N-1}} R_A(x) \leq \max_{x\perp T_{N-1}} R_A(x) \leq \alpha_N.$$

Since T_{N-1} is a particular subspace of dimension $N-1$.

□

Proposition 8.0.13. Suppose $A \le B$ with eigenvalues α_k, β_k, respectively, ordered to be decreasing. Then $\alpha_k \le \beta_k, \forall k$.

Proof. $A \le B$ implies $\langle Ax, x\rangle \le \langle Bx, x\rangle$, so $R_A(x) \le R_B(x)$. Now use either Fisher's or Courant's principle. □

8.0.2 Mercer's Theorem

We will need to use Dini's theorem from analysis.

Proposition 8.0.14. Suppose g_n are continuous functions on [0,1] with $g_n(x) \le g_{n+1}(x)$ for each n and x and $g_\infty = \lim_{n\to\infty} g_n(x)$ is continuous. Then g_n converges to g uniformly.

Proof. Let $f_n = g_\infty - g_n$, so the f_n's are continuous and decrease to 0. Let $\epsilon > 0$. If $G_n(x) = \{x \in [0,1] : f_n(x) < \epsilon\}$, then G_n is an open set(with respect to the relative topology on [0,1]), since f_n is continuous. Since $f_n(x) \to 0$, each x will be in some G_n. Thus $\{G_n\}$ is an open cover for $[0,1]$. Let $G_{n_1}, G_{n_2}, \cdots, G_{n_m}$ be a finite subcover. If $n \ge \max(n_1, n_2, \cdots, n_m)$ and $x \in [0,1]$, then x is in some G_{n_j} and $f_n(x) \le f_{n_j}(x) < \epsilon$. Thus the convergence is uniform.

Define $K : L^2[0,1] \longrightarrow L^2[0,1]$ by

$$Ku(x) = \int_0^1 K(x,y)u(y)dy$$

K^* has kernel $\overline{K(y,x)}$.

Suppose, K is continuous, symmetric and real-valued. Then K is compact, as we showed before. Therefore there exist a complete orthonormal system $\{e_j\}$ of eigenvectors. Let K_j be the eigenvalue corresponding to e_j. $K : L^2 \longrightarrow C[0,1]$, so $e_j = K_j^{-1}Ke_j$ is continuous if $K_j \ne 0$. □

Theorem 8.0.15. Suppose K is real-valued, symmetric and continuous. Suppose K is positive $\langle Ku, u\rangle \ge 0$ for all $u \in H$. Then

$$K(x,y) = \sum_j K_j e_j(x)\overline{e_j(y)}$$

and the series converges uniformly and absolutely.

An example is to let $K = P_t$, the transition density of absorbing or reflecting Brownian motion.

Proof. First we observe that K_j are non-negative. To see this, let $u = e_j$, and we have $0 \le \langle e_j, Ke_j\rangle = K_j\langle e_j, e_j\rangle$.

$K \ge 0$ on the diagonal: Suppose $K(r,r) < 0$ for some r. Then $K(x,y) < 0$ if $|x-r|, |y-r| < \delta$ for some δ. Take $u = \chi\left[r - \dfrac{\delta}{2}, r + \dfrac{\delta}{2}\right]$. Then

$$\langle Ku, u\rangle = \int\int K(x,y)u(y)x(s)dsdt < 0,$$

a contradiction.

Let $K_N(x,y) = \sum_{j=1}^{N} K_j e_j(x)\overline{e_j(y)}$. If $f = \sum_{k=1}^{\infty} \langle f, e_k\rangle e_k$, we have

$$\begin{aligned} K_N f(x) &= \int_0^1 \sum_{j=1}^{N} K_j e_j(x)\overline{e_j(y)} \sum_{k=1}^{\infty} \langle f, e_k\rangle e_k(y)dy \\ &= \sum_{j=1}^{N} K_j \langle f, e_j\rangle e_j(x). \end{aligned}$$

We have

$$Kf(x) = \sum_{j=1}^{\infty} \langle f, e_{jj}(x) = \sum_{j=1}^{\infty} \langle f, e_j\rangle K_j e_j(x).$$

We conclude that $K - K_N$ is a positive operator, since

$$\begin{aligned} \langle f, (K - K_N)f\rangle &= \sum_{k=1}^{\infty} \sum_{j=1}^{N} K_j |\langle f, e_j\rangle|^2 \langle e_k, e_j\rangle \\ &= \sum_{j=1}^{N} |\langle f, e_j\rangle|^2 \geq 0 \end{aligned}$$

As above, $K - K_N$ is non-negative on the diagonal, which implies that

$$\sum_{j=1}^{N} K_j |e_j(x)|^2 \leq K(x,x).$$

Each term is non-negative, so the sum converges for each x. Let $J(x)$ be the limit. Let $M = \sup_{x,y\in[0,1]} |K(x,y)|$. By Cauchy-Schwarz inequality

$$\begin{aligned} |K_N(x,y)| &\leq \left(\sum_{j=1}^{N} K_j |e_j(x)|^2\right)^{\frac{1}{2}} \left(\sum_{j=1}^{N} K_j |e_j(y)|^2\right)^{\frac{1}{2}} \\ &= (K_N(x,x))^{\frac{1}{2}} (K_N(y,y))^{\frac{1}{2}}. \end{aligned}$$

Fix x. By the same argument,

$$\begin{aligned} \left|\sum_{j=m}^{n} K_j e_j(x)\overline{e_j(y)}\right| &\leq \left(\sum_{j=m}^{n} K_j |e_j(x)|^2\right)^{\frac{1}{2}} \left(\sum_{j=m}^{n} K_j |e_j(y)|^2\right)^{\frac{1}{2}} \\ &\leq \left(\sum_{j=m}^{n} K_j |e_j(x)|^2\right)^{\frac{1}{2}} .M^{\frac{1}{2}}. \end{aligned}$$

The last line goes to 0 as $m, n \to \infty$ since $K_N(x,x) \to J(x) \leq M$. Therefore for each x, the functions $K_N(x,.)$ converge uniformly. Let's call the limit $L(x,y)$. Then $L(x,y)$ will be continuous in y for each x.

Given f, let

$$f_N(x) = \sum_{j=1}^{N} \langle f, e_j\rangle e_j(x).$$

Note

$$\begin{aligned} Kf_N(x) &= \sum_{j=1}^{N} \langle f, e_j \rangle K e_j(x) \\ &= \sum_{j=1}^{N} \langle f, e_j \rangle K_j e_j(x) \\ &= K_N f(x). \end{aligned}$$

We have,

$$\|f - f_N\|^2 = \sum_{j=N+1}^{\infty} |\langle f, e_j \rangle|^2 \longrightarrow 0$$

as $N \to \infty$ by Bessel's inequality, so

$$\begin{aligned} |Kf(x) - Kf_N(x)| &\leq \int_0^1 |K(x,y)||f(y) - f_N(y)|dy \\ &\leq M\|f - f_N\| \end{aligned}$$

by Cauchy-Schwarz inequality. Therefore $K_N f(x) \to Kf(x)$ as $N \to \infty$.

By Dominated convergence theorem,

$$K_N f(x) = \int_0^1 K_N(x,y) f(y) dy \to \int_0^1 L(x,y) f(y) dy.$$

We therefore have,

$$\int_0^1 L(x,y) f(y) dy = Kf(x)$$

for all $f \in L^2[0,1]$. This implies that(x is still fixed) $K(x,y) = L(x,y)$ for almost every y. With x fixed, both sides are continuous functions of y, hence they are equal for every y.

This is true for each x, and $K(x,y)$ is continuous, hence L is continuous. We now can apply Dini's theorem to conclude that $K_N(x,x)$ converges to $L(x,x) = J(x)$ uniformly. Finally, again by Cauchy-Schwarz inequality,

$$\sum_{j=m}^{n} K_j |e_j(x)||\overline{e_j(y)}| \leq \left(\sum_{j=m}^{n} K_j |e_j(x)|^2 \right)^{\frac{1}{2}} \left(\sum_{j=m}^{n} K_j |e_j(y)|^2 \right)^{\frac{1}{2}}$$

and this proves that $K_N(x,y)$ converges to K uniformly and absolutely. □

8.0.3 Positive Compact Operators

We will do the Krein-Rutman theorem, which is a generalization of the Perron-Frobenius theorem for matrices.

Theorem 8.0.16. Suppose Q is compact and Hausdorff and $X = C(Q)$, the complex-valued continuous functions on Q. Suppose $K : C(Q) \longrightarrow C(Q)$ and K is compact. Suppose further that K maps real-valued functions to real-valued functions. Finally, suppose that whenever $f \geq 0$ and f is not identically zero, then Kf is strictly positive. Then Kf has a positive eigen value σ of multiplicity one, the assigned eigen function is positive, and all the other eigenvalues of K are strictly smaller in absolute value than σ.

Examples include matrices with all positive entries, the semigroup P_t when $t = 1$ for reflecting Brownian motion on a bounded interval, and

$$Kf(x) = \int K(x, y) f(y) \mu(dy)$$

where K is jointly continuous, positive and μ is a finite measure. We have seen that the operator K is compact.

Proof. If $f \le g$ and $f \not\equiv g$, then $g - f \ge 0$, so $K(g - f) > 0$, or $Kf < Kg$.

Step-1: We show there exists a non-zero eigenvalue. Let f be the identically one function. Since Kf is continuous and everywhere positive, there exists a positive number b such that $Kf \ge b = bf$.

If f and b are any pair such that $f \ge 0$, and $Kf \ge bf$, then

$$b^2 f \le bKf = K(bf) \le K(Kf) = K^2 f$$

and continuing, $b^n f \le K^n f$.

Since $f \ge 0$,

$$b^n \|f\| \le \|K^n f\| \le \|K^n\| \|f\|.$$

So,

$$r(K) = \lim \|K^n\|^{\frac{1}{n}} \ge b.$$

Therefore $r(K)$ is strictly positive. Since K is compact, the set of eigenvalues of K is nonempty. We have shown that there exists a non-zero eigenvalue for K. Moreover, any b that satisfies $Kf \ge bf$ for some $f \ge 0$ is less than or equal to $r(K)$.

Step-2: K is compact, so there exists an eigenvalue λ and an eigen function g such that $Kg = \lambda g$, $|\lambda| = r(K)$. Let λ and g be any pair with $|\lambda| = r(K)$.

(a) We claim: If $f = |g|$ and $\sigma = |\lambda|$, then $\sigma f \le Kf$.

Proof. Let $x \in Q$. Multiply g by $\alpha \in \mathbb{C}$ such that $|\alpha| = 1$ and $\alpha\lambda g(x)$ is real and non-negative. Of course α depends on x. Write $g = u + iv$. Then

$$Ku(x) + iKv(x) = Kg(x) = \lambda g(x).$$

Looking at the real part,

$$\lambda g(x) = (Ku)(x)$$

Next, $u \le |g| = f$, and

$$|\lambda| f(x) = |\lambda g(x)| = Ku(x) \le (Kf)(x).$$

Then

$$\sigma f(x) \le Kf(x). \tag{8.0.4}$$

Although g depends on α, which depends on x, neither σ nor f depend on x. Since x was arbitrary, the above inequality (8.0.4) holds for all x. □

(b) We claim, $\sigma f = Kf$.

Proof. If not, there exists x such that $\sigma f(x) < Kf(x)$. By continuity, there exists a neighbourhood N about x such that

$$\sigma f(s) + \epsilon \le Kf(s), s \in N$$

Let $h > 0$ in N, 0 outside of N and so $Kh > 0$.

We will find c, $\epsilon > 0$ and set $F = f + \epsilon h$, $k = \sigma + c\epsilon$ and get $kF \le KF$. This will be a contradiction to step-1: if $bf \le Kf$, then we know $b \le r(K)$; use this with b replaced by k and f replacing F.

(i) Now $Kh > 0$, so there exists $c \leq 1$ such that $cf \leq Kh$. If $s \in N$,

$$\begin{aligned} KF(s) &= Kf(s) + \epsilon Kh(s) \geq Kf(s) + \epsilon cf(s) \\ &\geq \sigma f(s) + \delta + \epsilon cf(s). \end{aligned}$$

Then

$$\begin{aligned} KF(s) &= (\sigma + c\epsilon)(f + \epsilon h)(s) \\ &= \sigma f(s) + \epsilon cf(s) + \sigma\epsilon h(s) + c\epsilon^2 h(s) \\ &\leq KF(s) - \delta + \epsilon cf(s) + \sigma\epsilon h(s) + c\epsilon^2 h(s). \end{aligned}$$

Since h is bounded above, we can take ϵ small enough so that the last line is less than or equal to $KF(s)$.

(ii) If $s \notin N$, then $h(s) = 0$ and

$$\begin{aligned} KF(s) &= Kf(s) = (\sigma + c\epsilon)f(s) = \sigma f(s) + \epsilon cf(s) \\ &\leq Kf(s) + \epsilon Kh(s) = KF(s) \end{aligned}$$

using that $cf \leq Kh$.

□

Step-3: We next show that any other eigenvalue that has absolute value σ is in fact equal to σ. Let G be any eigen function corresponding to λ with $|\lambda| = \sigma$. Fix $x \in Q$. As before, we may assume $\lambda G(x) \geq 0$. As before, write $G = u + iv$ and then $\lambda G(x) = Ku(x)$. We have $u \leq |G| = f$. Suppose $u < f$ at some point $y \in Q$. Then $u \leq f$ and $u < f$ at one point means that we have $Ku < Kf$ at every point, and so

$$|\lambda| f(x) = |\lambda G(x)| = \lambda G(x) = Ku(x) < Kf(x).$$

So, $\sigma f(x) < Kf(x)$. But we showed $\sigma f = Kf$. Therefore u is identically equal to f. This implies that G is real and positive, and then it follows that λ is real and positive. Since $G = \sigma^{-1} KG$, G is strictly positive.

Step-4: Finally, we show σ has multiplicity 1. If not, there exists distinct real eigen functions f_1, f_2. But some linear combination H of f_1, f_2 will be real, take the value 0, but not be identically zero. As before $|H|$ will be an eigen function that is non-negative, and must also take the value 0. Moreover, the corresponding eigenvalue is σ. But then $0 < K|H| = \sigma|H|$, a contradiction to $|H|$ taking the value 0.

□

Exercise 8.0.17.

1. Show that if T is a compact operator on a Hilbert space H, then its adjoint is again compact.

2. Suppose that $T \in K(H)$. Let $(e_n)_{n=1}^{\infty}$ be an orthonormal basis of H and P_n a projection onto linear span of $\{e_1, e_2, \cdots, e_n\}$. Prove that $\|P_n T - T\| \to 0$ as $n \to \infty$.

3. Show that for a positive compact operator T on a Hilbert space H there is a positive compact operator S such that $S^2 = T$.

Unit 9

Course Structure

- Strongly continuous semigroup: Strongly continuous semigroup of operator and contraction, infinitesimal generator

---

9.0.1 Semigroup

A semigroup is a set S coupled with a binary operator $*(* : S \times S \to S)$ which is associative. That is, for all $x, y, z \in S$, $(x * y) * z = x * (y * z)$. Associativity can also be realized as $F(F(x, y), z) = F(x, F(y, z))$, where $F(x, y)$ serves as the mapping $S \times S$ to S.

A semigroup, unlike a group, need not have an identity element e such that $x * e = x, \forall x \in S$. Further, a semigroup need not have an inverse. Therefore, many problems which can be solved with semigroups can only be solved in the forward direction.

Example 9.0.1. Some of the simplest examples of semigroups are:

$$\begin{aligned} 2S &= \mathbb{R} & * &= \text{ addition} \\ S &= M_{2\times 2}(\mathbb{R}) & * &= \text{ matrix multiplication} \end{aligned}$$

where $M_{2\times 2}(\mathbb{R})$ = the set of 2×2 matrices with real entries.

9.0.2 Strongly Continuous Semigroups

Let X be a Banach space over the complex numbers, $T(t) = T_t$ linear bounded operators for $t \geq 0$. T is a semigroup if $T_{t+s} = T_t T_s, T_0 = I$.

These come up in the study material and in probability. For example, if one wants to solve the equation,

$$\frac{\partial u}{\partial t}(t, x) = \frac{\partial^2 u}{\partial x^2}(t, x), \;\; u(0, x) = f(x)$$

where f is a given function(i.e., the heat equation on $\mathbb{R}$), the solution is given by $u(t, x) = T_t f(x)$ for a certain semigroup T_t.

If X_t is a Markov process, then $T_t f(x) = \mathbb{E}^x f(X_t)$ will be a semigroup, where $\mathbb{E}^x$ means expectation starting at x.

Here is an example: If X is a Hilbert space and $\{\phi_n\}$ is an orthonormal basis and λ_j a sequence of real numbers increasing to infinity, let

$$T_t f = \sum_{j=1}^{\infty} e^{-\lambda_j t}\langle f, \phi_j\rangle \phi_j$$

Another example is to let

$$T_t f(x) = \int f(y)\frac{1}{\sqrt{2\pi t}}e^{-(x-y)^2/2t}dy \tag{9.0.1}$$

where X is the set of continuous functions on $\mathbb{R}$ vanishing at infinity.

A third example is given by the next proposition.

Proposition 9.0.2. Let $A : X \longrightarrow X$ be bounded. Then $T_t = e^{tA}$(defined as $e^{tA} = \sum t^n \frac{A^n}{n!}$) is a semigroup that is continuous in the norm topology.

Proof. This follows easily from the functional calculus for operators. □

We say T_t is strongly continuous at $t = 0$ if $\|T_t x - x\| \to 0$ as $t \to 0$ for all $x \in X$.

Proposition 9.0.3. Suppose T_t is strongly continuous semigroup at 0.

(1) There exists b and K such that $\|T_t\| \le be^{Kt}$.

(2) $T_t x$ is strongly continuous in t for all $x \in X$.

Proof. (1) We claim $\|T_t\|$ is bounded near 0. If not, there exists $t_j \to 0$ such that $\|T_{t_j}\| \to \infty$. By the uniform boundedness principle, $T_{t_j}x$ cannot converge to x for all x, a contradiction to strong continuity. So there exists a, b such that $\|T_{t_j}\| \le b$ for $t \le a$.

Write $t = na + r$, $T_t = T_a^n T_r$, so

$\|T_t\| \le \|T_a\|^n \|T_r\| \le b^{n+1} \le be^{Kt}$ with $K = \frac{1}{a}\log b$.

(2) $T_t x - T_s x = T_s[T_{t-s}x - x]$, so

$$\|T_t x - T_s x\| \le \|T_s\|\|T_{t-s}x - x\| \to 0.$$

□

Suppose D is dense in X and $A : D \longrightarrow X$ is closed. $z \in \rho(A)$, the resolvent set, if $z - A$ maps $D = D(A)$ one-to-one onto X. Thus $\rho(A) = \sigma(A)^c$. Write $R(z) = R_z = (zI - A)^{-1}$.

Since A is closed, then R_z is closed. To see this, suppose $x_n \to x$ and $y_n = R_z x_n \to y$. Then

$$Ay_n = zy_n - (z - A)y_n = zy_n - x_n \to zy - x.$$

Since A is closed, $y \in D(A)$ and $Ay = zy - x$, or $(z - A)y = x$. So $y = R_z x$, which proves R_z is closed.

R_z is defined on all of X, so by the Closed graph theorem, R_z is a bounded operator.

Let T be strongly continuous one parameter semigroup. The infinite generator A is defined by

$$Ax = \lim_{h\to 0}\frac{T_h x - x}{h}$$

where we mean that the difference of the two sides goes to 0 in norm. The domain of A consists of those x for which the strong limit exists.

As an example, with T_t defined by 9.0.1, if $f \in C^2$ vanishes at infinity, then using Taylor's theorem,

$$\begin{aligned}
& \frac{T_h f(x) - f(x)}{h} \\
= & \frac{1}{h}\int [f(y)] - f(x)]\frac{1}{\sqrt{2\pi h}} e^{-\frac{(y-x)^2}{2h}}dy + f'(x)\int (y-x)\frac{1}{\sqrt{2\pi h}}e^{-\frac{(y-x)^2}{2h}}dy \\
& +\frac{1}{2}f''(x)\int (y-x)^2\frac{1}{\sqrt{2\pi h}}e^{-\frac{(y-x)^2}{2h}}dy + \int E(h)\frac{1}{\sqrt{2\pi h}}e^{-\frac{(y-x)^2}{2h}}dy \\
= & \frac{1}{2}f''(x) + \frac{E(h)}{h} \to \frac{1}{2}f''(x).
\end{aligned}$$

where $E(h)$ is a remainder term that goes to 0 faster than h; we used standard facts about the Gaussian density. One can improve the above to show that the convergence is uniform, and we can then conclude that $C^2 \subset D(A)$ and $Af = \frac{1}{2}f''$.

Proposition 9.0.4. (1) A commutes with T_t in the sense that if $x \in D(A)$, then $T_t x \in D(A)$ and $AT_t x = T_t Ax$.

(2) $D(A)$ is dense in X.

(3) $D(A^n)$ is dense.

(4) A is closed.

(5) If $\|T_t\| \leq be^{Kt}$ and $Rez > K$, then $z \in \rho(A)$. The resolvent of A is the Laplace transform of T_t.

Proof. (1)

$$\frac{T_{t+h} - T_t}{h}x = T_t\frac{T_h - I}{h}x = \frac{T_h - I}{h}T_t x.$$

If $x \in D(A)$, the middle term converges to $T_t Ax$. So the limit exists in the third term, and therefore $T_t x \in D(A)$. Moreover

$$\frac{d}{dt}T_t x = T_t Ax = AT_t x.$$

(2) We claim,

$$T_t x - x = A\int_0^t T_s x ds.$$

To see this, $T_s x$ is a continuous function of s. Using Riemann sum approximation,

$$\begin{aligned}
\frac{T_h - I}{h}\int_0^t T_s x ds &= \frac{1}{h}\int_0^t [T_{s+h}x - T_s x]ds \\
&= \frac{1}{h}\int_t^{t+h} T_s x ds - \frac{1}{h}\int_0^h T_s x ds \\
&\to T_t x - x.
\end{aligned}$$

So, $\int_0^t T_s x ds \in D(A)$. But $\frac{1}{t}\int_0^t T_s x ds \to x$.

(3) Let ϕ be C^∞ and supported in $(0,1)$. Let

$$x_\phi = \int_0^1 \phi(s)T_s x ds.$$

Then,

$$\begin{aligned}
Ax_\phi &= \int_0^1 \phi(s) AT_s x ds \\
&= \int_0^1 \phi(s) \frac{\partial}{\partial s} T_s x ds \\
&= -\int_0^1 \phi'(s) T_s x ds
\end{aligned}$$

using integration by parts. Repeating, $x_\phi \in D(A^n)$. Now take ϕ_j approximating the identity.

(4) $T_t x - x = \displaystyle\int_0^t T_s Ax ds$: To see this, both are 0 at 0. The derivative on the left is $T_t Ax$, which is the same as the derivative on the right. Let $x_n \in D(A)$, $x_n \to x$, $Ax_n \to y$. Then

$$T_t x_n - x_n = \int_0^t T_s Ax_n ds \to \int_0^t T_s y ds.$$

The left hand term converges to $T_t x - x$. Divide by t and let $t \to 0$. The right hand side converges to y. Therefore, $x \in D(A)$ and $Ax = y$.

(5) Let

$$L(z)x = \int_0^\infty e_{zs} T_s x ds.$$

The Riemann integral converges when $Rez > K$.

$$\begin{aligned}
\|L(z)x\| &\leq \int_0^\infty be^{(K-Rez)s} \|x\| ds \\
&\leq \frac{b}{Rez - K} \|x\|.
\end{aligned}$$

We claim $L(z) = R_z$. Check that $e^{-zt}T_t$ is also a semigroup with infinitesimal generator $A - zI$.

Hence,

$$e^{-zt}T_t - x = (A - zI) \int_0^t e^{-zs} T_s x ds.$$

As $t \to \infty$, the left hand side tends to $-x$ and the right hand side tends to $(A - zI)L(z)x$. Since A is closed, $x = (zI - A)L(z)x$. So $L(z)$ is the right inverse of $(zI - A)$. Similarly, we see that it is also the left inverse. □

9.0.3 Generation of semigroups

Proposition 9.0.5. A strongly continuous semigroup of operators is uniquely defined by its infinitesimal generator.

Proof. If S, T have the same generator, let $x \in D(A)$ and

$$\frac{d}{dt} S_t T_{s-t} x = S(t) AT_{s-t} x - S_t AT_{s-t} x = 0.$$

Therefore,

$$0 = \int_0^s \frac{d}{dr} S_r T_{s-r} x dr = S_s T_0 x - S_0 T_s x,$$

or, $S_s x = T_s x$. Now use the fact that $D(A)$ is dense.

T_t is a contradiction if $\|T_t\| \leq 1$ for all t. □

Proposition 9.0.6. The infinitesimal semigroup of a strongly continuous semigroup of contractions has $(0, \infty) \subset \rho(A)$ and

$$\|R_\lambda\| = \|(\lambda I - A)^{-1}\| \leq \frac{1}{\lambda}.$$

Proof. We already did this: this is the case $b = 1$, $k = 0$. We have

$$\|L(z)x\| \leq \frac{1}{|Rez - K}\|x\|.$$

□

Proposition 9.0.7. Suppose B is an extension of A and there exists $\lambda \in \rho(A) \cap \rho(B)$. Then $A = B$.

Proof. Suppose $x \in D(B) \setminus D(A)$. We know $(\lambda - B)x \in X$, so $(\lambda - A)^{-1}(\lambda - B)x \in D(A) \subset D(B)$. Then,

$$\begin{aligned}(\lambda - B)(\lambda - A)^{-1}(\lambda - B)x &= (\lambda - A)(\lambda - A)^{-1}(\lambda - B)x \\ &= (\lambda - B)x.\end{aligned}$$

Operating both sides with $(\lambda - B)^{-1}$ to obtain $(\lambda - A)^{-1}(\lambda - B)x = x$. So, $x \in D(A)$, a contradiction. □

Unit 10

Course Structure

- Hille-Yosida theorem, Lumer-Phillips lemma, Trotter's theorem, Stone's theorem.

10.1 Hille-Yosida Theorem

Theorem 10.1.1. Let A be a densely defined unbounded operator such that $(0, \infty) \subset \rho(A)$ and $\|R_\lambda\| = \|\lambda I - A)^{-1} \leq \dfrac{1}{\lambda}$.

Then A is the infinitesimal generator of a strongly continuous semigroup of contractions.

Note that saying $(0, \infty) \subset \rho(A)$ implies that $\lambda - A$ is one-to-one and onto from the domain of A to the Banach space, which means the range of $\lambda - A$ is all of the Banach space.

Proof. Note that $nR_n - I = R_nA$ since $R_n(nI - A) = I$. Let $A_n = nAR_n$. Then $A_n = n^2R_n - nI$, so A_n is a bounded operator. Define $T_n(t) = e^{tA_n}$.

Step-1. We show $nR_nx \to x$ for all x.

To prove this,

$$\|nR_nx - x\| = \|R_nA(x)\| \leq \frac{1}{n}\|Ax\|,$$

so the claim is true for $x \in D(A)$. Since $\|nR_n\| \leq 1$ and $D(A)$ is dense in X, this proves the claim.

Step-2. We show that if $x \in D(A)$, then $A_n(x) \to A(x)$

$$A_nx = nAR_nx = nR_nAx \to Ax.$$

Step-3. We show that $T_n(s)x$ converges for all x. We have

$$T_n(t) = e^{tA_n} = e^{-nt}e^{n^2R_nt} = e^{-nt}\sum \frac{(n^2t)^m}{m!}(R_n)^m,$$

so $\|T_n(t)\| \leq e^{nt}e^{nt} = 1$.

A_n and A_m commute with T_n and T_m.

$$\frac{d}{dt}T_n(s-t)T_m(t)x = T_n(s-t)T_m(t)[A_m - A_n]x$$

The norm of the right hand side is bounded by $\|A_n x - A_m x\|$. So

$$\|T_n(s)x - T_m(s)x\| \leq s\|A_n x - A_m x\| \to 0$$

as $n, m \to \infty$. Therefore $T_n(s)x$ converges, say to $T_s x$, uniformly in s. $D(A)$ is dense. So this holds for all x.

$T_n(s)$ is strongly continuous semigroup of contraction, so the same holds for T_s.

Step-4. It remains to show that A is the infinitesimal generator of T. We have

$$T_n(t)x - x = \int_0^t T_n(s)A_n x ds.$$

If $x \in D(A)$, we can let $n \to \infty$ to get

$$T_t x - x = \int_0^t T_s A x ds.$$

If B is the generator of T, dividing by t and letting $t \to 0$, we get $D(A) \subset D(B)$ and $B = A$ on $D(A)$. So B is an extension of A. If $\lambda > 0$, then $\lambda \in \rho(A), \rho(B)$, which implies B cannot be a proper extension by the preceding proposition. □

Alternate Proof

Theorem 10.1.2. A densely defined operator A in a Banach space X is the infinitesimal generator of a semigroup $\{Q(t)\}$ if and only if there are constants c, r so that

$$\|(\lambda I - A)^{-m}\| \leq c(\lambda - r)^{-m} \tag{10.1.1}$$

for all $\lambda > r$ and all positive integers m.

Proof. If A is related to $\{Q(t)\}$, we have

$$(\lambda I - A)^{-1} = R(\lambda),$$

for $\lambda > r$, where

$$R(\lambda)x = \int_0^\infty e^{-\lambda t} Q(t) x dt \tag{10.1.2}$$

is the Laplace transform of $Q(t)x$. Hence $R(\lambda)^2 x$ is the transform of the convolution

$$\int_0^t Q(t-s)Q(s)x ds = tQ(t)x. \tag{10.1.3}$$

Continuing in this way, we find that (10.1.3) implies

$$R(\lambda)^m x = \frac{1}{(m-1)!}\int_0^\infty t^{m-1} e^{-\lambda t} Q(t) x dt$$

for $m = 1, 2, 3, \ldots$. Therefore, with c and r, we get

$$\begin{aligned} \|R(\lambda)^m\| &\leq \frac{c}{(m-1)!}\int_0^\infty t^{m-1} e^{-(\lambda - r)t} dt \\ &= c.(\lambda - r)^{-m} \end{aligned} \tag{10.1.4}$$

This proves the necessity of (10.1.1).

For the converse, set $S(\epsilon) = (I - \epsilon A)^{-1}$, so that (10.1.1) becomes

$$\|S(\epsilon)^m\| \le c(1 - \epsilon r)^{-m} \;\; [0 < \epsilon < \epsilon_0, m = 1, 2, 3 \ldots] \tag{10.1.5}$$

and the relations

$$(I - \epsilon A)S(\epsilon)x = x = S(\epsilon)(I - \epsilon A)x \tag{10.1.6}$$

holds, the first for all $x \in X$, the second for all $x \in \mathscr{D}(A)$.

If $x \in \mathscr{D}(A)$, then $x - S(\epsilon)x = -\epsilon S(\epsilon)x$, so that

$$\lim_{\epsilon \to 0} S(\epsilon)x = x. \tag{10.1.7}$$

But since $\|S(\epsilon)\| \le c(1 - \epsilon_0 r)^{-1}$, $\|S(\epsilon) : 0 < \epsilon < \epsilon_0\}$ is equicontinuous, and hence (10.1.7) holds for all $x \in X$.

Next we set,

$$T(t, \epsilon) = \exp(tAS(\epsilon)) \tag{10.1.8}$$

and claim that

$$\|T(t, \epsilon)\| \le c \exp\left\{\frac{rt}{1 - \epsilon r}\right\} \;\; [0 < \epsilon < \epsilon_0, t > 0]. \tag{10.1.9}$$

Indeed, the relation $\epsilon AS(\epsilon) = S(\epsilon) - 1$ shows that,

$$T(t, \epsilon) = e^{-t/\epsilon} \sum_{m=0}^{\infty} \frac{t^m}{m! \epsilon^m} S(\epsilon)^m. \tag{10.1.10}$$

Now, (10.1.9) follows from (10.1.5) and (10.1.10).

For $x \in \mathscr{D}(A)$, (10.1.6) and (10.1.8) shows that

$$\frac{d}{dt}\{T(t, \epsilon)T(t, \delta)^{-1}x\} = T(t, \epsilon)T(t, \delta)^{-1}(S(\epsilon) - S(\delta))Ax.$$

If we integrate this and apply $T(t, \delta)$ to the result, we obtain

$$T(t, \epsilon)x - T(t, \delta)x = \int_0^t T(u, \epsilon)T(t - u, \delta)(S(t) - S(\delta))Axdu. \tag{10.1.11}$$

If we use (10.1.7) with Ax in place of x, and refer to (10.1.9), we see that the right side of (10.1.11) converge to 0 when $\epsilon \to 0$ and $\delta \to 0$. The limit

$$Q(t)x = \lim_{\epsilon \to 0} T(t, \epsilon)x \tag{10.1.12}$$

exists therefore for every $x \in \mathscr{D}(A)$, uniformly on every bounded subset of $[0, \infty)$. Moreover, (10.1.9) shows that $\|Q(t)\| \le ce^{rt}$.

By equicontinuity, and the assumption that $\mathscr{D}(A)$ is dense, we see now that (10.1.12) holds for all $x \in X$. Since $T(t, \epsilon)$ is defined by (10.1.8), it follows that $\{Q(t)\}$ is a semigroup.

Let $\tilde{A}$ be the infinitesimal generator of $\{Q(t)\}$. Then,

$$(\lambda I - \tilde{A})^{-1}x = \int_0^\infty e^{-\lambda t} Q(t)xdt \;\; (\lambda > r). \tag{10.1.13}$$

On the other hand, $AS(\epsilon)$ is the infinitesimal generator of $\{\exp(tAS(\epsilon))\} = \{T(t,\epsilon)\}$. Thus

$$(\lambda I - AS(\epsilon))^{-1}x = \int_0^\infty e^{-\lambda t} T(t,\epsilon)x dt. \tag{10.1.14}$$

By (10.1.12) this becomes

$$(\lambda I - A)^{-1}x = \int_0^\infty e^{-\lambda t} Q(t)x dt. \tag{10.1.15}$$

Comparison of (10.1.13) and (10.1.15) shows now that $\lambda I - A$ and $\lambda I - \tilde{A}$ have the same inverse for all sufficiently large λ, and this implies that $\tilde{A} = A$. $\square$

10.2 Lumer-Phillips Lemma

Lemma 10.2.1. Let A be densely defined in a Hilbert space B and suppose $(0,\infty) \subset \rho(A)$. Then $\|R_\lambda\| \leq \dfrac{1}{\lambda}$ if and only if $Re\langle x, Ax\rangle \leq 0, \forall x \in D(A)$. If the last property holds, we say A is dissipative. An example is the Laplacian:

$$\langle f, Af\rangle = \int f(x)\Delta f(x)dx = -\int |\bigtriangledown f(x)|^2 dx \leq 0$$

using integration by parts. Another example is if $Af(x) = \displaystyle\sum_{i,j=1}^n \frac{\partial}{\partial x_i}\left(a_{ij}(.)\frac{\partial}{\partial x_j}(.)\right)(x)$.

To verify this, we again use integration by parts.

Proof. Suppose,

$$\|(\lambda I - A)^{-1}u\|^2 \leq \frac{1}{\lambda^2}\|u\|^2.$$

Let $x = (\lambda I - A)^{-1}u$. So,

$$\langle x, x\rangle \leq \frac{1}{\lambda^2}\langle \lambda x - Ax, \lambda x - Ax\rangle.$$

This becomes

$$\begin{aligned} 2Re\langle x, Ax\rangle &= \langle x, Ax\rangle + \langle Ax, x\rangle \\ &\leq \frac{1}{\lambda}\|Ax\|^2. \end{aligned}$$

This is true for all λ. So let $\lambda \to \infty$. For the converse,

$$\begin{aligned} \langle x, Ax\rangle + \langle Ax, x\rangle &= 2Re\langle x, Ax\rangle \\ &\leq 0 \\ &\leq \frac{1}{\lambda}\|Ax\|^2 \end{aligned}$$

for all $\lambda > 0$. $\square$

10.3 Trotter's Theorem

Theorem 10.3.1. Suppose A is the infinitesimal generator of a semigroup of contractions in a Hilbert space. Let B be a densely defined dissipative operator such that $D(A) \subset D(B)$ and there exists $b > 0$ and $a \in (0, 1)$ such that

$$\|Bx\| \leq a\|Ax\| + b\|x\|, x \in D(A).$$

Then $A + B$(defined on $D(A)$) is the generator of a contraction semigroup.

Proof. First, $A + B$ is closed: Let $x_n \to x$ and $y_n = (A + B)x_n \to y$. So,

$$A(x_n - x_m) = y_n - y_m - B(x_n - x_m),$$

and

$$\|A(x_n - x_m)\| \leq \|y_n - y_m\| + a\|A(x_n - x_m)\| + b\|x_n - x_m\|.$$

Since, $a < 1$, then Ax_n converges. Therefore, Bx_n converges. A is closed, so $Ax_n \to Ax$. If $x \in D(A) \subset D(B)$,

$$\|Bx_n - Bx\| \leq a\|A_n x - Ax\| + b\|x_n - x\| \to 0.$$

Then, $(A + B)x_n \to (A + B)x$.

Next, $\lambda \in \rho(A + B)$: By the Lumer-Phillips lemma, A is dissipative. B is also. So, $A + B$ is dissipative. By Lumer-Phillips lemma,

$$\|x\| \leq \frac{1}{\lambda}\|(\lambda I - (A + B)x\|.$$

One immediate consequence of this is that the operator $\lambda - (A + B)$ is one-to-one. Another is that the range of $\lambda - (A + B)$ is closed, because if y_n is in the range and $y_n \to y$, then $y_n = (\lambda - (A + B))x_n$ for some x_n. The inequality shows that $\|x_n - x_m\| \to 0$. If $x_n \to x$, then $y = (\lambda - (A + B))x$, since $A + B$ is a closed operator. Therefore the range of $(A + B) - \lambda I$ is closed.

The range is X: If not, $\exists v \neq 0$ perpendicular to the range. $A - \lambda I$ is invertible, so $\exists x \in D(A)$ such that $(A - \lambda I)x = v$. Then $v + Bx$ is in the range, or $\langle v + Bx, v\rangle = 0$. So, $\|v\|^2 + \langle Bx, v\rangle = 0$, or

$$\|v\|^2 \leq \|Bx\|\|v\|$$

and so $\|v\| \leq \|Bx\|$. Then,

$$\|Ax - \lambda x\| \leq \|Bx\| \leq a\|Ax\| + b\|x\|.$$

Squaring and using the fact that A is dissipative,

$$\|Ax\|^2 + \lambda^2\|x\|^2 \leq a^2\|Ax\|^2 + 2ab\|Ax\|\|x\| + b^2\|x\|.$$

This holds for all $\lambda > 0$, so for λ large enough, $\|x\| = 0$. So, $x = 0$ and the range is the whole space.

Now use the Hille-Yosida theorem. □

10.4 Stone's Theorem

Theorem 10.4.1. (1) Suppose A is self-adjoint and H is a Hilbert space. There exists a strongly continuous semigroup $\mathrm{U}(t)$ of unitary operators with infinitesimal generator A.

(2) Given a strongly continuous group of unitary operators, the generator is of the form iA where A is self-adjoint.

Proof. (1) We know that the spectrum of an unbounded self-adjoint operator is real that $\|(z-A)^{-1}\| \leq \dfrac{1}{|Imz|}$. So if $\lambda > 0$ and $z = -i\lambda$, then

$$\begin{aligned}
\|(\lambda - iA)^{-1}\| &= \|(iz - iA)^{-1}\| \\
&= \|(z-A)^{-1}\| \\
&\leq \frac{1}{|Im(iz)|} \\
&= \frac{1}{\lambda}.
\end{aligned}$$

The resolvent set of iA contains the positive reals. So iA and $-iA$ satisfy the Hille-Yosida theorem. Let $U(t), V(t)$ be the respective semigroups.

V and U are inverses:

$$\begin{aligned}
\frac{d}{dt}U(t)V(t) &= U(t)iAV(t)x - U(t)iAV(t)x \\
&= 0.
\end{aligned}$$

So, $U(t)V(t)x$ is independent of t. When $t = 0$, we get x. So $U(t)V(t)x = x$ if $x \in D(A)$. But $D(A)$ is dense.

Both U and V are contractions. Since $U(t)V(t) = I$, they must be norm preserving. This is because,

$$\|x\| = \|U(t)V(t)x\| \leq \|V(t)x\| \leq \|x\|$$

so, $\|x\| = \|V(t)x\|$ and similarly with U. Since they are invertible. Define $U(t) = V(-t)$ for $t < 0$.

(2) Let $V(t) = U(-t)$. Then $U(t)$ and $V(t)$ are strongly continuous semigroups of contractions, and the infinitesimal generators are additive inverse. So the generators are B, $-B$.

Since both B, $-B$ are infinitesimal generators, all real numbers except 0 are in the resolvent set of B. Take $x \in D(B)$.

$$\|U(t)x\|^2 = (U(t)x, U(t)x) = \|x\|^2.$$

Take the derivative with respect to t:

$$(Bx, x) + (x, Bx) = 0.$$

Letting $A = -iB$ so that $B = iA$, we see that $\langle Ax, x\rangle = \langle x, Ax\rangle$, $\forall x \in D(A)$. Using $\langle Ax, x\rangle = \langle x, Ax\rangle$ with respect to x replaced by $x + y$ and with x replaced by y, we obtain

$$\langle Ax, y\rangle + \langle Ay, x\rangle = \langle x, Ay\rangle + \langle y, Ax\rangle. \tag{10.4.1}$$

Replacing y by iy in above (10.4.1)

$$-i\langle Ax, y\rangle + i\langle Ay, x\rangle = -i\langle x, Ay\rangle + i\langle y, Ax\rangle.$$

Dividing this by i and subtracting from (10.4.1), we have $\langle Ax, y\rangle = \langle x, Ay\rangle$.

Therefore, A is symmetric and A^* is an extension of A. It follows that B^* is an extension of $-B$. The adjoint of $(\lambda - B)^{-1}$ is $(\overline{\lambda} - B^*)^{-1}$, and it follows that $\rho(B^*) = \overline{\rho(B)}$. If $z \neq 0$ and $z \in \mathbb{R}$, then $z \in \rho(B)$. So $z \in \rho(B^*)$. Also, $z \in \rho(-B)$. Again, B^* cannot be a proper extension of $-B$, hence $B^* = -B$, and so $A^* = A$, or A is self-adjoint.

□

Exercise 10.4.2. 1. If $f \in H^2$ and $f(z) = \sum c_n z^n$, we define

$$[Q(t)f](z) = \sum_{n=0}^{\infty} (n+1)^{-t} c_n z^n \quad (0 \leq t < \infty).$$

Show that each $Q(t)$ is self-adjoint and positive. Find the infinitesimal generator A of the semigroup $\{Q(t)\}$. Is A self-adjoint?

2. Define $Q(t) \in B(L^2)$, where $L^2 = L^2(\mathbb{R})$, by $(Q(t)f)(s) = f(s+t)$.

   (a) Prove that each $Q(t)$ is unitary.

   (b) Prove that A is the infinitesimal generator of $\{Q(t)\}$ and $f \in \mathscr{D}(A)$ if and only if $\int |y\hat{f}(y)|^2 dy < \infty$ (where $\hat{f}$ is the Fourier transform of f) and that $Af = f'$ for all $f \in \mathscr{D}(A)$.

Unit 11

Course Structure

- Measures: Class of Sets, Measures, The extension theorems

11.1 Introduction

One of the most fundamental concepts in Euclidean geometry is that of the measure $m(E)$ of a solid body E in one or more dimensions. In one, two, and three dimensions, we refer to this measure as the length, area, or volume of E respectively. In the classical approach to geometry, the measure of a body was often computed by partitioning that body into finitely many components, moving around each component by a rigid motion (e.g. a translation or rotation), and then reassembling those components to form a simpler body which presumably has the same area. One could also obtain lower and upper bounds on the measure of a body by computing the measure of some inscribed or circumscribed body; this ancient idea goes all the way back to the work of Archimedes at least. Such arguments can be justified by an appeal to geometric intuition, or simply by postulating the existence of a measure $m(E)$ that can be assigned to all solid bodies E, and which obeys a collection of geometrically reasonable axioms. One can also justify the concept of measure on "physical" or "reductionistic" grounds, viewing the measure of a macroscopic body as the sum of the measures of its microscopic components. You are already aware of the idea of outer measure of sets in $\mathbb{R}$, which used exactly that. In this unit,

Objectives

After completing this unit, you will be able to:

- define various class of sets and look into some of their examples
- learn about the idea of measure and outer measure functions
- learn different properties of measure
- have some basic idea of extension of measure functions across different class of sets

11.2 Class of sets

Let Ω be a nonempty set and $\mathcal{P}(\Omega) \equiv \{A : A \subset \Omega\}$ be the power set of Ω, i.e., the class of all subsets of Ω.

Definition 11.2.1. A class $\mathcal{C} \subset \mathcal{P}(\Omega)$ is called a semialgebra if

1. $A, B \in \mathcal{C} \Rightarrow A \cap B \in \mathcal{C}$;
2. for any $A \in \mathcal{C}$, there exist sets $B_1, B_2, \ldots, B_k \in \mathcal{C}$, for some $1 \leq k < \infty$, such that $B_i \cap B_j = \emptyset$ for $i \neq j$ and $A^c = \displaystyle\bigcup_{i=1}^{k} B_i$.

Example 11.2.2.
1. $\Omega = \mathbb{R}, \mathcal{C} \equiv \{(a, b], (b, \infty) : \ -\infty \leq a, b < \infty\}$.
2. $\Omega = \mathbb{R}, \mathcal{C} \equiv \{I : \ I \text{ is an interval}\}$.
3. $\Omega = \mathbb{R}^k, \mathcal{C} \equiv \{I_1 \times I_2 \times \ldots I_k : \ I_j \text{ is an interval in } \mathbb{R} \text{ for } 1 \leq j \leq k\}$.

Definition 11.2.3. A collection of sets $\mathcal{F} \subset \mathcal{P}(\Omega)$ is called an algebra if

1. $\Omega \in \mathcal{F}$,
2. $A \in \mathcal{F}$ implies $A^c \in \mathcal{F}$, and
3. $A, B \in \mathcal{F}$ implies $A \cup B \in \mathcal{F}$ (i.e., closure under pairwise unions).

Thus, an algebra is a class of sets containing Ω that is closed under complementation and pairwise (and hence finite) unions. It is easy to see that one can equivalently define an algebra by requiring that properties (a), (b) hold and that the property (c)$'$ $\quad A, B \in \mathcal{F} \Rightarrow A \cap B \in \mathcal{F}$ holds (i.e. closure under finite intersections).

Definition 11.2.4. A class $\mathcal{F} \subset \mathcal{P}(\Omega)$ is called a σ-algebra if it is an algebra and if it satisfies

$$A_n \in \mathcal{F} \quad \text{for} \quad n \geq 1 \Rightarrow \bigcup_{n \geq 1} A_n \in \mathcal{F}.$$

Thus, a σ-algebra is a class of subsets of Ω that contains Ω and is closed under complementation and countable unions. As pointed out in the introductory chapter, a σ-algebra can be alternatively defined as an algebra that is closed under monotone unions as the following shows.

Theorem 11.2.5. Let $\mathcal{F} \subset \mathcal{P}(\Omega)$. Then $\mathcal{F}$ is a σ-algebra if and only if $\mathcal{F}$ is an algebra and satisfies

$$A_n \in \mathcal{F}, A_n \subset A_{n+1} \quad \text{for all } n \Rightarrow \bigcup_{n \geq 1} A_n \in \mathcal{F}.$$

Proof. The 'only if' part is obvious. For the 'if' part, let $\{B_n\}_{n=1}^{\infty} \subset \mathcal{F}$. Then, since $\mathcal{F}$ is an algebra, $A_n \equiv \bigcup_{j=1}^{n} B_j \in \mathcal{F}$ for all n. Further, $A_n \subset A_{n+1}$ for all n and $\bigcup_{n \geq 1} B_n = \bigcup_{n \geq 1} A_n$. Since by hypothesis $\bigcup_n A_n \in \mathcal{F}, \bigcup_n B_n \in \mathcal{F}$. $\square$

Here are some examples of algebras and σ-algebras.

Example 11.2.6. Let $\Omega = \{a, b, c, d\}$. Consider the classes

$$\mathcal{F}_1 = \{\Omega, \emptyset, \{a\}\}$$

and

$$\mathcal{F}_2 = \{\Omega, \emptyset, \{a\}, \{b, c, d\}\}$$

Then, $\mathcal{F}_2$ is an algebra (and also a σ-algebra), but $\mathcal{F}_1$ is not an algebra, since $\{a\}^c \notin \mathcal{F}_1$.

Example 11.2.7. Let Ω be any nonempty set and let

$$\mathcal{F}_3 = \mathcal{P}(\Omega) \equiv \{A : A \subset \Omega\}, \quad \text{the power set of } \Omega$$

and

$$\mathcal{F}_4 = \{\Omega, \emptyset\}$$

Then, it is easy to check that both $\mathcal{F}_3$ and $\mathcal{F}_4$ are σ-algebras. The latter σ-algebra is often called the trivial σ-algebra on Ω.

From the definition it is clear that any σ-algebra is also an algebra and thus $\mathcal{F}_2, \mathcal{F}_3, \mathcal{F}_4$ are examples of algebras, too. The following is an example of an algebra that is not a σ-algebra.

Example 11.2.8. Let Ω be a nonempty set, and let $|A|$ denote the number of elements of a set $A \subset \Omega$. Define. $\mathcal{F}_5 = \{A \subset \Omega : \text{either } |A| \text{ is finite or } |A^c| \text{ is finite }\}$.

Then, note that (i) $\Omega \in \mathcal{F}_5$ (since $|\Omega^c| = |\varphi| = 0$)), (ii) $A \in \mathcal{F}_5$ implies $A^c \in \mathcal{F}_B$ (if $|A| < \infty$, then $|(A^c)^c| = |A| < \infty$ and if $|A^c| < \infty$, then $A^c \in \mathcal{F}_5$ trivially). Next, suppose that $A, B \in \mathcal{F}_B$. If either $|A| < \infty$ or $|B| < \infty$, then

$$|A \cap B| \leq \min\{|A|, |B|\} < \infty$$

so that $A \cap B \in \mathcal{F}_5$. On the other hand, if both $|A^c| < \infty$ and $|B^c| < \infty$, then

$$|(A \cap B)^c| = |A^c \cup B^c| \leq |A^c| + |B^c| < \infty$$

implying that $A \cap B \in \mathcal{F}_5$. Thus, property (c) holds, and $\mathcal{F}_5$ is an algebra. However, if $|\Omega| = \infty$, then $\mathcal{F}_5$ is not a σ-algebra. To see this, suppose that $|\Omega| = \infty$ and $\{\omega_1, \omega_2, \ldots\} \subset \Omega$. Then, by definition, $A_t = \{\omega_1\} \in \mathcal{F}_5$ for all $i \geq 1$, but $A \equiv \bigcup_{i-1}^{\infty} A_{2i-1} = \{\omega_1, \omega_3, \ldots\} \notin \mathcal{F}_5$, since $|A| = |A^c| = \infty$.

Example 11.2.9. Let Ω be a nonempty set and let

$$\mathcal{F}_6 = \{A \subset \Omega : A \text{ is countable or } A^c \text{ is countable }\}.$$

Then, it is easy to show that $\mathcal{F}_6$ is a σ-algebra.

Definition 11.2.10. If $\mathcal{A}$ is a class of subsets of Ω, then the σ-algebra generated by $\mathcal{A}$, denoted by $\sigma(\mathcal{A})$, is defined as

$$\sigma\langle\mathcal{A}\rangle = \bigcap_{\mathcal{F} \in \mathcal{I}(\mathcal{A})} \mathcal{F},$$

where $\mathcal{I}(\mathcal{A}) \equiv \{\mathcal{F} : \mathcal{A} \subset \mathcal{F} \text{ and } \mathcal{F} \text{ is a } \sigma\text{-algebra on } \Omega\}$ is the collection of all σ-algebras consisting the class $\mathcal{A}$.

Note that since the power set $\mathcal{P}(\Omega)$ contains $\mathcal{A}$ and is itself a σ-algebra, the collection $\mathcal{I}(\mathcal{A})$ is not empty and hence, the intersection in the above definition is well defined.

Example 11.2.11. In the setup of Example 11.2.7, $\sigma\langle\mathcal{F}_1\rangle = \mathcal{F}_2$ (why?).

A particularly useful class of σ-algebras are those generated by open sets of a topological space. These are called Borel σ-algebras. A topological space is a pair $(\mathrm{S}, \mathcal{T})$ where S is a nonempty set and $\mathcal{T}$ is a collection of subsets of S such that (i) $\mathrm{S} \in \mathcal{T}$, (ii) $\mathcal{O}_1, \mathcal{O}_2 \in \mathcal{T} \Rightarrow \mathcal{O}_1 \cap \mathcal{O}_2 \in \mathcal{T}$, and (iii) $\{\mathcal{O}_\alpha : \alpha \in I\} \subset \mathcal{T} \Rightarrow \bigcup_{\alpha \in I} \mathcal{O}_\alpha \in \mathcal{T}$. Elements of $\mathcal{T}$ are called open sets.

A metric space is a pair (S, d) where S is a nonempty set and d is a function from $\mathbb{S} \times \mathbb{S}$ to R^+satisfying (i) $d(x, y) = d(y, x)$ for all x, y in $\mathbb{S}$, (ii) $d(x, y) = 0$ iff $x = y$, and (iii) $d(x, z) \leq d(x, y) + d(y, z)$ for all x, y, z in S. Property (iii) is called the triangle inequality. The function d is called a metric on S (ef. see A.4).

Any Euclidean space $\mathbb{R}^n (1 \leq n < \infty)$ is a metric space under any one of the following metrics:

1. For $1 \leq p < \infty, d_p(x,y) = (\sum_{i=1}^n |x_i - y_h|^p)^{1/p}$.

2. $d_\infty(x,y) = \max_{1 \leq 1 \leq n} |x_i - y|$.

3. For $0 < p < 1, d_p(x,y) = (\sum_{1=1}^n |x_1 - y_4|^p)$. A metric space (S,d) is a topological space where s set $\mathcal{O}$ is open if for all $x \in \mathcal{O}$, there is an $\epsilon > 0$ such that $B(x,\epsilon) \equiv \{y : d(y,x) < \epsilon\} \subset \mathcal{O}$.

Definition 11.2.12. The Borel σ-algebra on a topological space S (in particular, on a metric space or an Euclidean space) is defined as the σ-algebra generated by the collection of open sets in S.

Example 11.2.13. Let $\mathcal{B}\left(\mathbb{R}^k\right)$ denote the Borel σ-algebra on $\mathbb{R}^k, 1 \leq k < \infty$. Then,

$$B\left(\mathbb{R}^k\right) \equiv \sigma\left\langle\left\{A : A \text{ is an open subset of } \mathbb{R}^k\right\}\right\rangle$$

is also generated by each of the following classes of sets

$$\begin{aligned}
\mathcal{O}_1 &= \{(a_1,b_1) \times \ldots \times (a_k,b_k) : -\infty \leq a_i < b_i \leq \infty, 1 \leq i \leq k\} \\
\mathcal{O}_2 &= \{(-\infty,x_1) \times \ldots \times (-\infty,x_k) : x_1,\ldots,x_k \in \mathbb{R}\} \\
\mathcal{O}_3 &= \{(a_1,b_1) \times \ldots \times (a_k,b_k) : a_i,b_i \in \mathcal{Q}, a_i < b_i, 1 \leq i \leq k\} \\
\mathcal{O}_4 &= \{(-\infty,x_1) \times \ldots \times (-\infty,x_k) : x_1,\ldots,x_k \in \mathcal{Q}\}
\end{aligned}$$

where $\mathcal{Q}$ denotes the set of all rational numbers. To show this, note that $\sigma\left(\mathcal{O}_i\right) \subset \mathcal{B}\left(\mathbb{R}^k\right)$ for $i = 1,2,3,4$, and hence, it is enough to show that $\sigma\left(\mathcal{O}_i\right\rangle \supset \mathcal{B}\left(\mathbb{R}^k\right)$. Let $\mathcal{G}$ be a σ-algebra containing $\mathcal{O}_3$. Observe that given any open set $A \subset \mathbb{R}^k$, there exist a sequence of sets $\{B_n\}_{n\geq 1}$ in O_3 such that $A = \bigcup_{n>1} B_n$ (Problem 1.9). Since G is a σ algebra snd $B_n \in \mathcal{G}$ for all $n \geq 1, A \in \mathcal{G}$. Thus, G is a σ-algebra consisting all open subsets of $\mathbb{R}^k$, and hence $\mathcal{B}\left(\mathbb{R}^k\right)$. Hence, it follows that

$$\mathcal{B}\left(\mathbb{R}^k\right) \supset \sigma\left\langle\mathcal{O}_1\right\rangle \supset \sigma\left(\mathcal{O}_3\right\rangle = \bigcap_{g:GO_3} \mathcal{G} \supset \mathcal{B}\left(\mathbb{R}^k\right)$$

Next note that any interval $(a,b) \subset \mathbb{R}$ can be expressed in terms of half-spaces of the form $(-\infty,x)$, $x \in \mathbb{R}$ as

$$(a,b) = \bigcup_{n=1}^{\infty}\left[(-\infty,b) \setminus \left(-\infty, a+\frac{1}{n}\right)\right]$$

where for any two sets A and B, $A \setminus B = \{x : \ x \in A, \ x \notin B\}$. It is not difficult to show that this implies $\sigma\langle\mathcal{O}_i\rangle = \mathcal{B}(\mathbb{R}^k)$ for $i = 2,4$.

Let us give the definitions of another two types of classes.

Definition 11.2.14. A class $\mathcal{C}$ of subsets of Ω is a π-system or π-class if $A, B \in \mathcal{C} \Rightarrow A \cap B \in \mathcal{C}$. And $\mathcal{C}$ is called a λ-system or a λ-class if

1. $\Omega \in \mathcal{C}$;

2. $A, B \in \mathcal{C}$ with $A \subset B \Rightarrow B \setminus A \in \mathcal{C}$;

3. $A_n \in \mathcal{C}$ with $A_n \subset A_{n+1}$ for all $n \geq 1 \Rightarrow \bigcup_{n\geq 1} A_n \in \mathcal{C}$.

It is easy to note that every σ-algebra is a λ-system, but an algebra need not be so.

Theorem 11.2.15. If $\mathcal{C}$ is a π-system, then $\lambda\langle\mathcal{C}\rangle = \sigma\langle\mathcal{C}\rangle$.

Exercise 11.2.16. Let Ω be a nonempty set and $\mathcal{P}(\Omega) \equiv \{A : A \subset \Omega\}$ be the power set of Ω. Let $\mathcal{A} \subset \mathcal{P}(\Omega)$. Show that the smallest smallest algebra containing a semialgebra $\mathcal{C}$ is

$$\mathcal{A}(\mathcal{C}) \equiv \{A : \ A = \bigcup_{i=1}^{k} B_i, \ B_i \in \mathcal{C}, \text{ for } i = 1, \ldots, k, \ k < \infty\},$$

i.e., the class of finite unions of sets from $\mathcal{C}$.

11.3 Measures

A set function is an extended real valued function defined on a class of subsets of a set Ω. Measures are nonnegative set functions that, intuitively speaking, measure the content of a subset of Ω. However, a measure has to satisfy certain natural requirements, such as the measure of the union of a countable collection of disjoint sets is the sum of the measures of the individual sets, etc. Formally, one can define measure as the following.

Definition 11.3.1. Let Ω be a nonempty set and $\mathcal{F}$ be an algebra on Ω. Then, a set function μ on $\mathcal{F}$ is called a measure if

1. $\mu(A) \in [0, \infty]$ for all $A \in \mathcal{F}$;
2. $\mu(0) = 0$;
3. for any disjoint collection of sets $A_1, A_2, \ldots, \in \mathcal{F}$ with $\bigcup_{n \geq 1} A_n \in \mathcal{F}$,

$$\mu\left(\bigcup_{n \geq 1} A_n\right) = \sum_{n-1}^{\infty} \mu(A_n).$$

The above conditions on μ are equivalent to finite additivity and monotone continuity from below.

Theorem 11.3.2. Let Ω be a nonempty set and $\mathcal{F}$ be an algebra of subsets of Ω and μ be a set function on $\mathcal{F}$ with values in $[0, \infty]$ and with $\mu(\emptyset) = 0$. Then, μ is a measure iff μ satisfies

(d) (finite additivity) for all $A_1, A_2 \in \mathcal{F}$ with $A_1 \cap A_2 = \$, \mu(A_1 \cup A_2) = \mu(A_1) + \mu(A_2)$, and

(e) (monotone continuity from below or, m.c.f.b., in short) for any collection $\{A_n\}_{n \geq 1}$ of sets in $\mathcal{F}$ such that $A_n \subset A_{n+1}$ for all $n \geq 1$ and $\bigcup_{n \geq 1} A_n \in \mathcal{F}$,

$$\mu\left(\bigcup_{n \geq 1} A_n\right) = \lim_{n \to \infty} \mu(A_n).$$

Proof. Let μ be a measure on $\mathcal{F}$. Since μ satisfies (c), taking $A_3, A_4, \ldots$ to be $\emptyset$ yields (d). This implies that for A and B in $\mathcal{F}, A \subset B \Rightarrow \mu(B) = \mu(A) + \mu(B \backslash A) \geq \mu(A)$, i.e., μ is monotone. To establish (e), note that if $\mu(A_n) = \infty$ for some $n = n_0$, then $\mu(A_n) = \infty$ for all $n \geq n_0$ and $\mu\left(\bigcup_{n \geq 1} A_n\right) = \infty$ and (e) holds in this case. Hence, suppose that $\mu(A_n) < \infty$ for all $n \geq 1$. Setting $B_n = A_n \backslash A_{n-1}$ for $n \geq 1$ (with

$A_0 = \emptyset$), by (d), $\mu(B_n) = \mu(A_n) - \mu(A_{n-1})$. Since $\{B_n\}_{n\geq 1}$ is a disjoint collection of sets in $\mathcal{F}$ with $\bigcup\limits_{n\geq 1} B_n = \bigcup\limits_{n\geq 1} A_n$, by (c),

$$\begin{aligned}
\mu\left(\bigcup_{n\geq 1} A_n\right) &= \mu\left(\bigcup_{n\geq 1} B_n\right) \\
&= \sum_{n=1}^{\infty} \mu(B_n) \\
&= \lim_{N\to\infty} \sum_{n=1}^{N} [\mu(A_n) - \mu(A_{n-1})] \\
&= \lim_{N\to\infty} \mu(A_N),
\end{aligned}$$

and so (e) holds in this case too.

Conversely, let μ satisfy $\mu(\emptyset) = 0$ and (d) and (e). Let $\{A_n\}_{n\geq 1}$ be a disjoint collection of sets in $\mathcal{F}$ with $\bigcup\limits_{i\geq 1} A_i \in \mathcal{F}$. Let $C_n = \bigcup\limits_{j=1}^{n} A_j$ for $n \geq 1$. Since $\mathcal{F}$ is an algebra, $C_n \in \mathcal{F}$ for all $n \geq 1$. Also, $C_n \subset C_{n+1}$ for all $n \geq 1$. Hence, $\bigcup_{n\geq 1} C_n = \bigcup_{j\geq 1} A_j$. By (e),

$$\begin{aligned}
\mu\left(\bigcup_{j\geq 1} A_j\right) &= \mu\left(\bigcup_{n\geq 1} C_n\right) \\
&= \lim_{n\to\infty} \mu(C_n) \\
&= \lim_{n\to\infty} \sum_{j=1}^{n} \mu(A_j) \quad (\text{ by (d)}) \qquad (11.3.1) \\
&= \sum_{j=1}^{\infty} \mu(A_j).
\end{aligned}$$

Thus, (c) holds. □

The definition of measure given above is valid when $\mathcal{F}$ is a σ-algebra. However, very often one may start with a measure on an algebra A and then extend it to a measure on the σ-algebra $\sigma\langle A\rangle$. This is why the definition of a measure on an algebra is given here. Similarly, one may begin with defining a measure on a class of subsets of Ω that form only a semialgebra. Let us now discuss a few relevant definitions and examples.

Definition 11.3.3. A measure μ is called finite or infinite according as $\mu(\Omega) < \infty$ or $\mu(\Omega) = \infty$. A finite measure with $\mu(\Omega) = 1$ is called a probability measure. A measure μ on a σ-algebra $\mathcal{F}$ is called σ-finite if there exist a countable collection of sets $A_1, A_2, \ldots, \in \mathcal{F}$, not necessarily disjoint, such that

$$\text{(a)} \bigcup_{n\geq 1} A_n = \Omega \;\text{ and }\; \text{(b)}\; \mu(A_n) < \infty \;\text{ for all }\; n \geq 1.$$

Example 11.3.4. (The counting measure). Let Ω be a nonempty set and $\mathcal{F}_3 = \mathcal{P}(\Omega)$ be the set of all subsets of Ω (see example 11.2.7). Define

$$\mu(A) = |A|, \quad A \in \mathcal{F}_3,$$

where $|A|$ denotes the number of elements in A. It is easy to check that μ satisfies the requirements (a)-(c) of a measure. This measure μ is called the counting measure on Ω. Note that μ is finite iff Ω is finite and it is σ-finite if Ω is countably infinite.

Let us now discuss a few properties of the measure.

Theorem 11.3.5. Let μ be a measure on an algebra $\mathcal{F}$, and let $A, B, A_1, \ldots, A_k \in \mathcal{F}, 1 \leq k < \infty$. Then,

(a)$'$ (Monotonicity) $\mu(A) \leq \mu(B)$ if $A \subset B$;

(b)$'$ (Finite subadditivity) $\mu\left(A_1 \cup \ldots \cup A_k\right) \leq \mu\left(A_1\right) + \ldots + \mu\left(A_k\right)$;

(c)$'$ (Inclusion-exclusion formula) If $\mu\left(A_i\right) < \infty$ for all $i = 1, \ldots, k$, then

$$\begin{aligned}\mu\left(A_1 \cup \ldots \cup A_k\right) &= \sum_{i=1}^{k} \mu\left(A_i\right) - \sum_{1 \leq i < j < k} \mu\left(A_i \cap A_j\right) \\ &\quad + \ldots + (-1)^{k-1} \mu\left(A_1 \cap \ldots \cap A_k\right).\end{aligned}$$

Proof. $\mu(B) = \mu\left(A \cup \left(A^c \cap B\right)\right) = \mu(A) + \mu(B \backslash A) \geq \mu(A)$, by (a) and (c) of the definition of measure. This proves (a)$'$.

To prove (b)$'$, note that if either $\mu(A)$ or $\mu(B)$ is finite, then $\mu(A \cap B) < \infty$, by (a)$'$. Hence, using the countable additivity property (c), we have

$$\begin{aligned}\mu(A \cup B) &= \mu(A) + \mu(B \backslash A) \\ &= \mu(A) + [\mu(B \backslash A) + \mu(A \cap B)] - \mu(A \cap B) \\ &= \mu(A) + \mu(B) - \mu(A \cap B).\end{aligned} \tag{11.3.2}$$

Hence, (b)$'$ follows from (11.3.2) and by induction.

To prove (c)$'$, we note that the case $k = 2$ follows from (11.3.2). Next, suppose that (c)$'$ holds for all sets $A_1, \ldots, A_k \in \mathcal{F}$ with $\mu\left(A_i\right) < \infty$ for all $i = 1, \ldots, k$ for some $k = n, n \in \mathbb{N}$. To show that it holds for $k = n + 1$, we have by (11.3.2),

$$\begin{aligned}&\mu\left(\bigcup_{i=1}^{n+1} A_i\right) \\ = \;& \mu\left(\bigcup_{i=1}^{n} A_i\right) + \mu\left(A_{n+1}\right) - \mu\left(\bigcup_{i=1}^{n}\left(A_i \cap A_{n+1}\right)\right) \\ = \;& \left\{\sum_{i=1}^{n} \mu(A_i) - \sum_{1<i<j\leq n} \mu(A_i \cap A_j) + \ldots + (-1)^{n-1} \mu(A_1 \cap \ldots A_n)\right\} + \mu(A_{n+1}) \\ & - \left[\sum_{i=1}^{n} \mu(A_i \cap A_{n+1}) - \sum_{1\leq i<j\leq n} \mu(A_i \cap A_j \cap A_{n+1}) + \ldots + (-1)^{n-1} \mu(A_1 \cap \ldots A_{n+1})\right] \\ = \;& \sum_{i=1}^{n+1} \mu(A_i) - \sum_{1\leq i<j\leq n+1} \mu(A_i \cap A_j) + \ldots + (-1)^{n} \mu\left(\bigcap_{j=1}^{n+1} A_j\right).\end{aligned}$$

Hence, by induction, the proof is done. □

Theorem 11.3.6. Let μ be a measure on an algebra $\mathcal{F}$.

1. (Monotone continuity from above) Let $\{A_n\}_{n\geq 1}$ be a sequence of sets in $\mathcal{F}$ such that $A_{n+1} \subset A_n$ for all $n \geq 1$ and $A \equiv \bigcap_{n\geq 1} A_n \in \mathcal{F}$. Also, let $\mu(A_{n_0}) < \infty$ for some $n_0 \in \mathbb{N}$. Then,

$$\lim_{n\to\infty} \mu(A_n) = \mu(A)$$

2. (Countable subadditivity) If $\{A_n\}_{n\geq 1}$ is a sequence of sets in $\mathcal{F}$ such that $\bigcup_{n\geq 1} A_n \in \mathcal{F}$, then

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mu(A_n).$$

Proof. To prove part 1, without loss of generality (w.l.o.g.), assume that $n_0 = 1$, i.e., $\mu(A_1) < \infty$. Let $C_n = A_1 \backslash A_n$ for $n \geq 1$, and $C_\infty = A_1 \backslash A$. Then C_n and C_∞ belong to $\mathcal{F}$ and $C_n \uparrow C_\infty$. By theorem 11.3.2 (e), (i.e., by the m.c.f.b. property), $\mu(C_n) \uparrow \mu(C_\infty)$ and by (d), (i.e., finite additivity), $\mu(C_n) = \mu(A_1) - \mu(A_n)$ for all $1 \leq n < \infty$, due to the fact $\mu(A_1) < \infty$. This proves 1.

To prove 2, let $D_n = \bigcup_{i=1}^{n} A_i,\ n \geq 1$. Then $D_n \uparrow D \equiv \bigcup_{i\geq 1}^{n} A_i$. Hence, by m.c.f.b. and finite subadditivity,

$$\mu(D) = \lim_{n\to\infty} \mu(D_n) \leq \lim_{n\to\infty} \sum_{i=1}^{n} \mu(A_i) = \sum_{n=1}^{\infty} \mu(A_n).$$

□

11.4 Extension theorems

Definition 11.4.1. A set function μ on a semialgebra $\mathcal{C}$, taking values in $[0, \infty]$ is called a measure if

1. $\mu(\emptyset) = 0$;

2. for any sequence of sets $\{A_n\}_{n\geq 1} \subset \mathcal{C}$, with $\bigcup_{n\geq 1} A_n \in \mathcal{C}$, and $A_i \cap A_j = \emptyset$ for $i \neq j$, $\mu\left(\bigcup_{n\geq 1} A_n\right)$ $= \sum_{n=1}^{\infty} \mu(A_n)$.

Theorem 11.4.2. Let μ be a measure on a semialgebra $\mathcal{C}$. Let $\mathcal{A} \equiv \mathcal{A}(\mathcal{C})$ be the smallest algebra generated by $\mathcal{C}$. For each $A \in \mathcal{A}$, set

$$\bar{\mu}(A) = \sum_{i=1}^{k} \mu(B_i),$$

if the set A has the representation $A = \bigcup_{i=1}^{k} B_i$ for some $B_1, \ldots, B_k \in \mathcal{C},\ k < \infty$ with $B_i \cap B_j = \emptyset$ for $i \neq j$. Then,

(i) $\bar{\mu}$ is independent of the representation of A as $A = \bigcup_{i=1}^{k} B_i$;

(ii) $\bar{\mu}$ is finitely additive on $\mathcal{A}$, i.e., $A, B \in \mathcal{A}, A \cap B = \emptyset \Rightarrow \bar{\mu}(A \cup B) = \bar{\mu}(A) + \bar{\mu}(B)$; and

(iii) $\bar{\mu}$ is countably additive on $\mathcal{A}$, i.e., if $A_n \in \mathcal{A}$ for all $n \geq 1, A_i \cap A_j = \emptyset$ for all $i \neq j$, and $\bigcup_{n\geq 1} A_n \in \mathcal{A}$, then

$$\bar{\mu}\left(\bigcup_{n\geq 1} A_n\right) = \sum_{n=1}^{\infty} \bar{\mu}\left(A_n\right).$$

Proof. Parts (i) and (ii) are easy to verify. Turning to part (iii), let each $n \geq 1$, $A_n = \bigcup_{j=1}^{k_n} B_{nj}$, $B_{nj} \in \mathcal{C}$, $\{B_{nj}\}_{j=1}^{k_n}$ disjoint. Since $\bigcup_{n\geq 1} A_n \in \mathcal{A}$ then there exist disjoint sets $\{B_i\}_{i=1}^{k} \subset \mathcal{C}$ such that $\bigcup_{n\geq 1} A_n = \bigcup_{i=1}^{k} B_i$. Now

$$\begin{aligned} B_i = B_i \cap \left(\bigcup_{n\geq 1} A_n\right) &= \bigcup_{n\geq 1} \left(B_i \cap A_n\right) \\ &= \bigcup_{n\geq 1} \bigcup_{j=1}^{k_n} \left(B_i \cap B_{nj}\right). \end{aligned}$$

Since for all $i, B_i \in \mathcal{C}, B_i \cap B_{nj} \in \mathcal{C}$ for all j, n and μ is a measure on $\mathcal{C}$

$$\mu\left(B_i\right) = \sum_{n\geq 1} \sum_{j=1}^{k_n} \mu\left(B_i \cap B_{nj}\right).$$

Thus,

$$\begin{aligned} \bar{\mu}\left(\bigcup_{n\geq 1} A_n\right) &= \sum_{i=1}^{k} \mu\left(B_i\right) \\ &= \sum_{i=1}^{k} \sum_{n\geq 1} \left(\sum_{j=1}^{k_n} \mu\left(B_i \cap B_{nj}\right)\right) \\ &= \sum_{n\geq 1} \left(\sum_{i=1}^{k} \sum_{j=1}^{k_n} \mu(B_i \cap B_{nj}\right) \\ &= \sum_{n\geq 1} \overline{\mu}(A_n), \end{aligned}$$

since

$$A_n = A_n \cap \bigcup_{i=1}^{k} B_i = \bigcup_{i=1}^{k} \bigcup_{j=1}^{k_n} (B_i \cap B_{nj}).$$

□

Definition 11.4.3. Given a measure μ on a semialgebra $\mathcal{C}$, the outer measure induced by μ is the set function μ^*, defined on $\mathcal{P}(\Omega)$, as

$$\mu^*(A) = \inf\left\{\sum_{n=1}^{\infty} \mu(A_n) : \{A_n\}_{n\geq 1} \subset \mathcal{C},\ A \subset \bigcup_{n\geq 1} A_n\right\}.$$

Thus, a given set A is covered by countable unions of sets from $\mathcal{C}$ and the sums of the measures on such covers are computed and $\mu^*(A)$ is the greatest lower bound one can get in this way. It is not difficult to show that on $\mathcal{C}$ and $\mathcal{A}$, this is not an overestimate. That is, $\mu^* = \mu$ on $\mathcal{C}$ and on $\mathcal{A}$, $\mu^* = \overline{\mu}$. Now, let us define measurable sets.

Definition 11.4.4. A set A is said to be μ^*-measurable if

$$\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \cap A^c)$$

for all $E \subset \Omega$.

Exercise 11.4.5. Show that μ^* satisfies the following properties.

1. $\mu^*(\emptyset) = 0$;
2. $A \subset B \Rightarrow \mu^*(A) \leq \mu^*(B)$;
3. For any $\{A_n\}_{n\geq 1} \subset \mathcal{P}(\Omega)$,
$$\mu^*\left(\bigcup_{n\geq 1} A_n\right) \leq \sum_{n=1}^{\infty} \mu^*(A_n).$$

Any set function $\mu^* : \mathcal{P}(\Omega) \to [0, \infty]$ that satisfies the above three properties is called an outer measure on Ω.

Theorem 11.4.6. Let μ^* be an outer measure on Ω, i.e., it satisfies

1. $\mu^*(\emptyset) = 0$;
2. $A \subset B \Rightarrow \mu^*(A) \leq \mu^*(B)$;
3. For any $\{A_n\}_{n\geq 1} \subset \mathcal{P}(\Omega)$,
$$\mu^*\left(\bigcup_{n\geq 1} A_n\right) \leq \sum_{n=1}^{\infty} \mu^*(A_n).$$

Let $\mathcal{M} \equiv \mathcal{M}_{\mu^*} \equiv \{A : A \text{ is } \mu^*\text{-measurable}\}$. Then

(i) $\mathcal{M}$ is a σ-algebra,

(ii) μ^* restricted to $\mathcal{M}$ is a measure, and

(iii) $\mu^*(A) = 0 \Rightarrow \mathcal{P}(A) \subset \mathcal{M}$.

Proof. From μ^*-measurability, it follows that $\emptyset \in \mathcal{M}$ and that $A \in \mathcal{M} \Rightarrow A^c \in \mathcal{M}$. Next, it will be shown that $\mathcal{M}$ is closed under finite unions. Let $A_1, A_2 \in \mathcal{M}$. Then, for any $E \subset \Omega$,

$$\begin{aligned}
\mu^*(E) =& \mu^*\left(E \cap A_1\right) + \mu^*\left(E \cap A_1^c\right) \quad (\text{ since } A_1 \in \mathcal{M}) \\
=& \mu^*\left(E \cap A_1 \cap A_2\right) + \mu^*\left(E \cap A_1 \cap A_2^c\right) \\
&+ \mu^*\left(E \cap A_1^c \cap A_2\right) + \mu^*\left(E \cap A_1^c \cap A_2^c\right) \quad (\text{ since } A_2 \in \mathcal{M}).
\end{aligned}$$

But $(A_1 \cap A_2) \cup (A_1 \cap A_2^c) \cup (A_1^c \cap A_2) = A_1 \cup A_2$. Since μ^* is subadditive, it follows that

$$\mu^*\left(E \cap (A_1 \cup A_2)\right) \leq \mu^*\left(E \cap A_1 \cap A_2\right) + \mu^*\left(E \cap A_1 \cap A_2^c\right) + \mu^*\left(E \cap A_1^c \cap A_2\right).$$

Thus

$$\mu^*(E) \geq \mu^*\left(E \cap (A_1 \cup A_2)\right) + \mu^*\left(E \cap (A_1 \cup A_2)^c\right).$$

The subadditivity of μ^* yields the opposite inequality and so, $A_1 \cup A_2 \in \mathcal{M}$ and hence, $\mathcal{M}$ is an algebra. To show that $\mathcal{M}$ is a σ-algebra, it suffices to show that $\mathcal{M}$ is closed under monotone unions, i.e., $A_n \in \mathcal{M}, A_n \subset A_{n+1}$ for all $n \geq 1 \Rightarrow A \equiv \bigcup_{n\geq 1} A_n \in \mathcal{M}$. Let $B_1 = A_1$ and $B_n = A_n \cap A_{n-1}^c$ for all $n \geq 2$. Then for all $n \geq 1$, $B_n \in \mathcal{M}$ (since $\mathcal{M}$ is an algebra), $\bigcup_{j=1}^{n} B_j = A_n$ and $\bigcup_{j=1}^{\infty} B_j = A$. Hence for any $E \subset \Omega$,

$$\begin{aligned}
\mu^*(E) &= \mu^*\left(E \cap A_n\right) + \mu^*\left(E \cap A_n^c\right) \\
&= \mu^*\left(E \cap A_n \cap B_n\right) + \mu^*\left(E \cap A_n \cap B_n^c\right) + \mu^*\left(E \cap A_n^c\right) && \text{(since } B_n \in \mathcal{M}) \\
&= \mu^*\left(E \cap B_n\right) + \mu^*\left(E \cap A_{n-1}\right) + \mu^*\left(E \cap A_n^c\right) \\
&= \sum_{j=1}^{n} \mu^*\left(E \cap B_j\right) + \mu^*\left(E \cap A_n^c\right) && \text{(by iteration)} \\
&\geq \sum_{j=1}^{n} \mu^*\left(E \cap B_j\right) + \mu^*\left(E \cap A^c\right) && \text{(by monotonicity).}
\end{aligned}$$

Now letting $n \to \infty$, and using the subadditivity of μ^* and the fact that $\bigcup_{j=1}^{\infty} B_j = A$, one gets

$$\begin{aligned}
\mu^*(E) &\geq \sum_{j=1}^{\infty} \mu^*\left(E \cap B_j\right) + \mu^*\left(E \cap A^c\right) \\
&\geq \mu^*(E \cap A) + \mu^*\left(E \cap A^c\right).
\end{aligned}$$

This completes the proof of part (i).

To prove part (ii), let $\{B_n\}_{n\geq 1} \subset \mathcal{M}$ and $B_i \cap B_j = \emptyset$ for $i \neq j$. Let $A_j = \bigcup_{i=j}^{\infty} B_i, j \in \mathbb{N}$. Then, by (i), $A_j \in \mathcal{M}$ for all $j \in \mathbb{N}$ and so

$$\begin{aligned}
\mu^*\left(A_1\right) &= \mu^*\left(A_1 \cap B_1\right) + \mu^*\left(A_1 \cap B_1^c\right) \quad \text{(since } B_1 \in \mathcal{M}) \\
&= \mu^*\left(B_1\right) + \mu^*\left(A_2\right) \\
&= \mu^*\left(B_1\right) + \mu^*\left(B_2\right) + \mu^*\left(A_3\right) \quad \text{(by iteration)} \\
&= \sum_{i=1}^{n} \mu^*\left(B_i\right) + \mu^*\left(A_{n+1}\right) \quad \text{(by iteration)} \\
&\geq \sum_{i=1}^{n} \mu^*\left(B_i\right) \quad \text{for all } n \geq 1.
\end{aligned}$$

Now letting $n \to \infty$, one has $\mu^*\left(A_1\right) \geq \sum_{i=1}^{\infty} \mu^*\left(B_i\right)$. By subadditivity of μ^*, the opposite inequality holds and so

$$\mu^*\left(A_1\right) = \mu^*\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} \mu^*\left(B_i\right),$$

proving (ii).

As for (iii), note that by monotonicity, $\mu^*(A) = 0,\ B \subset A \Rightarrow \mu^*(B) = 0$ and hence, for any E, $\mu^*(E \cap B) = 0$. Since $\mu^*(E) \geq \mu^*(E \cap B^c)$, this implies

$$\mu^*(E) \geq \mu^*(E \cap B^c) + \mu^*(E \cap B).$$

The opposite inequality holds by the subadditivity of μ^*. So $B \in \mathcal{M}$ and hence (iii) is proved. $\square$

Few Probable Questions

1. Show that the measure function defined on an algebra $\mathcal{F}$ over non-empty set Ω satisfies monotone continuity from below.
2. Show that the measure μ satisfies finite subadditivity.
3. Define an extension of measure μ on a semialgebra $\mathcal{C}$ to the smallest algebra generated by $\mathcal{C}$. Is the newly defined function countably additive? Justify your answer.
4. Let M_{μ^*} be the set of all μ^*-measurable sets in Ω. Show that it forms a σ-algebra.

Unit 12

Course Structure

- Caratheodory extension of measure, Completeness of measure, Lebesgue-Stieltjes measures

---

12.1 Introduction

In Mathematics, a null set is a set that is negligible in some sense. In measure theory, we again have a null set which more or less can be similar in terms of measure. By definition, a set $E \subset \Omega$ is a null set for a measure μ on Ω if $E \in \mathcal{M}$ and $\mu(E) = 0$. In general, an arbitrary subset A of E need not be measurable, but if A happens to be measurable, then monotonicity implies that $\mu(A) = 0$. A complete measure is one such that every subset A of every null set E is measurable. Complete measures are often more convenient to work with than incomplete measures. Fortunately, if we have a incomplete measure in hand, there is a way to extend μ to a larger σ-algebra in such a way that the extended measure is complete. This is what this unit is all about. We shall start with the definition of complete measure and gradually move on developing the theory.

Objectives

After reading this unit, you will be able to

- define complete measure and see few of its examples
- define complete extension of measure on an incomplete measure space

12.2 Caratheodory extension of measures

Definition 12.2.1. A measure space $(\Omega, \mathcal{F}, \nu)$, where Ω is a non-empty set, $\mathcal{F}$ is a σ-algebra over Ω and ν is a measure function on $\mathcal{F}$, is called complete if for any $A \in \mathcal{F}$ with $\nu(A) = 0 \Rightarrow \mathcal{P}(A) \subset \mathcal{F}$.

By part (iii) of theorem 11.3.2, $(\Omega, \mathcal{M}, \mu^*)$ is a complete measure space.

Theorem 12.2.2. (Caratheodory's extension theorem). Let μ be a measure on a semialgebra $\mathcal{C}$ and let μ^* be the set function induced by μ as defined in the previous unit. Then,

(i) μ^* is an outer measure,

(ii) $\mathcal{C} \subset \mathcal{M}_{\mu^*}$, and

(iii) $\mu^* = \mu$ on $\mathcal{C}$, where $\mathcal{M}_\mu$ is as in Theorem 1.3.2.

Proof. The proof of (i) involves verifying the conditions satisfied by μ^*, which is left as an exercise in the preceding unit. To prove (ii), let $A \in \mathcal{C}$. Let $E \subset \Omega$ and $\{A_n\}_{n\geq 1} \subset \mathcal{C}$ be such that $E \subset \bigcup_{n\geq 1} A_n$. Then, for all $i \in \mathbb{N}$, $A_i = (A_i \cap A) \cup (A_i \cap B_1) \cup \ldots \cup (A_i \cap B_k)$ where $B_1, \ldots, B_k$ are disjoint sets in $\mathcal{C}$ such that $\bigcup_{j=1}^{k} B_j = A^c$. Since μ is finitely additive on $\mathcal{C}$,

$$\begin{aligned} \mu(A_i) &= \mu(A_i \cap A) + \sum_{j=1}^{k} \mu(A_i \cap B_j) \\ \Rightarrow \quad \sum_{n=1}^{\infty} \mu(A_n) &= \sum_{n=1}^{\infty} \mu(A_n \cap A) + \sum_{n=1}^{\infty}\sum_{j=1}^{k} \mu(A_n \cap B_j) \\ &\geq \mu^*(E \cap A) + \mu^*(E \cap A^c), \end{aligned}$$

since $\{A_n \cap A\}_{n\geq 1}$ and $\{A_n \cap B_j : 1 \leq j \leq k, n \geq 1\}$ are both countable subcollections of $\mathcal{C}$ whose unions cover $E \cap A$ and $E \cap A^c$, respectively. From the definition of $\mu^*(E)$, it now follows that

$$\mu^*(E) \geq \mu^*(E \cap A) + \mu^*(E \cap A^c).$$

Now the subadditivity of μ^* completes the proof of part (ii).

To prove (iii), let $A \in \mathcal{C}$. Then, by definition, $\mu^*(A) \leq \mu(A)$. If $\mu^*(A) = \infty$, then $\mu(A) = \infty = \mu^*(A)$. If $\mu^*(A) < \infty$, then by the definition of 'infimum,' for any $\epsilon > 0$, there exists $\{A_n\}_{n\geq 1} \subset \mathcal{C}$ such that $A \subset \bigcup_{n\geq 1} A_n$ and

$$\mu^*(A) \leq \sum_{n=1}^{\infty} \mu(A_n) \leq \mu^*(A) + \epsilon.$$

But $A = A \cap \left(\bigcup_{n\geq 1} A_n\right) = \bigcup_{n\geq 1} (A \cap A_n)$. We note that the set function $\bar{\mu}$ defined on $\mathcal{A}(\mathcal{C})$ defined in the previous unit is a measure and it coincides with μ on $\mathcal{C}$. Since $A,\ A \cap A_n \in \mathcal{C}$ for all $n \geq 1$, thus, by countable subadditivity applied to $\bar{\mu}$, we get

$$\mu(A) = \bar{\mu}(A) \leq \sum_{n=1}^{\infty} \bar{\mu}(A \cap A_n) \leq \sum_{n=1}^{\infty} \bar{\mu}(A_n) = \sum_{n=1}^{\infty} \mu(A_n) \leq \mu^*(A) + \epsilon.$$

Since $\epsilon > 0$ is arbitrary, this yields, $\mu(A) \leq \mu^*(A)$. □

Thus, given a measure μ on a semialgebra $\mathcal{C} \subset \mathcal{P}(\Omega)$, there is a complete measure space $(\Omega, \mathcal{M}_{\mu^*}, \mu^*)$ such that $\mathcal{M}_{\mu^*} \supset \mathcal{C}$ and μ^* restricted to $\mathcal{C}$ equals μ. For this reason, μ^* is called an extension of μ. The measure space $(\Omega, \mathcal{M}_{\mu^*}, \mu^*)$ is called the Caratheodory extension of μ. Since $\mathcal{M}_{\mu^*}$ is a $\sigma-$ algebra and contains $\mathcal{C}, \mathcal{M}_{\mu^*}$ must contain $\sigma\langle\mathcal{C}\rangle$, the σ-algebra generated by $\mathcal{C}$, and thus, $(\Omega, \sigma\langle\mathcal{C}\rangle, \mu^*)$ is also a measure space. However, the latter may not be complete (see Section 1.4).

12.3 Lebesgue-Stieltjes measures

Let us apply the above method to construct Lebesgue-Stieltjes measures on $\mathbb{R}$ and $\mathbb{R}^n$.

12.3.1 Lebesgue-Stieltjes measures on $\mathbb{R}$

Let $F : \mathbb{R} \to \mathbb{R}$ be nondecreasing. For $x \in \mathbb{R}$, let $F(x+) \equiv \lim_{y \downarrow x} F(y)$, and $F(x-) \equiv \lim_{y \uparrow x} F(y)$. Set $F(\infty) = \lim_{x \uparrow \infty} F(x)$ and $F(-\infty) = \lim_{x \downarrow -\infty} F(y)$. Let

$$\mathcal{C} \equiv \{(a, b] : -\infty \leq a \leq b < \infty\} \cup \{(a, \infty) : -\infty \leq a < \infty\}$$

Define

$$\mu_F((a, b]) = F(b+) - F(a+),$$
$$\mu_F((a, \infty)) = F(\infty) - F(a+).$$

Then, it may be verified that (i) $\mathcal{C}$ is a semialgebra; (ii) μ_F is a measure on $\mathcal{C}$. (For (ii), one needs to use the Heine-Borel theorem. See Problems 1.22 and 1.23.)

Let $\left(\mathbb{R}, \mathcal{M}_{\mu_F^*}, \mu_F^*\right)$ be the Caratheodory extension of μ_F, i.e., the measure space constructed as in the above two theorems.

12.4 Completeness of Measures

Theorem 12.4.1. Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. Let

$$\tilde{\mathcal{F}} = \{A : \ B_1 \subset A \subset B_2 \ \text{ for some } B_1, B_2 \in \mathcal{F} \ \text{ satisfying } \mu(B_2 \setminus B_1) = 0\}.$$

For every $A \in \tilde{\mathcal{F}}$, set $\tilde{\mu}(A) = \mu(B_1) = \mu(B_2)$ for any pair of sets $B_1, B_2 \in \mathcal{F}$ with $B_1 \subset A \subset B_2$ and $\mu(B_2 \setminus B_1) = 0$. Then

1. $\tilde{\mathcal{F}}$ is a σ-algebra and $\mathcal{F} \subset \tilde{\mathcal{F}}$,
2. $\tilde{\mu}$ is well defined,
3. $(\Omega, \tilde{\mathcal{F}}, \tilde{\mu})$ is a complete measure space and $\tilde{\mu} = \mu$ on $\mathcal{F}$.

Proof.

1. Since $A \in \tilde{\mathcal{F}}$, there exist $B_1, B_2 \in \mathcal{F}$ with $B_1 \subset A \subset B_2$ and $\mu(B_2 \setminus B_1) = 0$. Clearly $B_2^c \subset A^c \subset B_1^c$ and $B_1^c, \ B_2^c \in \mathcal{F}$ and $\mu(B_1^c \setminus B_2^c) = \mu(B_2 \setminus B_1) = 0$ and so $A^c \in \tilde{\mathcal{F}}$. Next, let $\{A_n\}_{n=1}^{\infty} \subset \tilde{\mathcal{F}}$ and $A = \bigcup_{n \geq 1} A_n$. Then, for each n there exist $B_{1n}, B_{2n} \in \mathcal{F}$ with $B_{1n} \subset A_n \subset B_{2n}$ and $\mu(B_{2n} \setminus B_{1n}) = 0$. Let $B_1 = \bigcup_{n \geq 1} B_{1n}$ and $B_2 = \bigcup_{n \geq 1} B_{2n}$. Then $B_1 \subset A \subset B_2, \ B_1, B_2 \in \mathcal{F}$ and $B_2 \setminus B_1 \subset \bigcup_{n \geq 1} (B_{2n} \setminus B_{1n})$ and hence $\mu(B_2 \setminus B_1) \leq \sum_{n=1}^{\infty} \mu(B_{2n} \setminus B_{1n}) = 0$. Thus, $A \in \tilde{\mathcal{F}}$ and hence $\tilde{\mathcal{F}}$ is a σ-algebra. Clearly, $\mathcal{F} \subset \tilde{\mathcal{F}}$ since for every $A \in \mathcal{F}$, one may take $B_1 = B_2 = A$.

2. Let $B_1 \subset A \subset B_2, B_1' \subset A \subset B_2', B_1, B_1', B_2, B_2' \in \mathcal{F}$ and $\mu\left(B_2 \setminus B_1\right) = 0 = \mu\left(B_2' \setminus B_1'\right)$. Then $B_1 \cup B_1' \subset A \subset B_2 \cap B_2'$ and $(B_2 \cap B_2') \setminus (B_1 \cup B_1') = (B_2 \cap B_2') \cap (B_1^c \cap B_1'^c) \subset B_2 \cap B_1^c$. Thus

   $$\mu\left(\left[B_2 \cap B_2'\right] \setminus \left[B_1 \cup B_1'\right]\right) = 0.$$

   Hence, $\mu\left(B_2\right) = \mu\left(B_1\right) + \mu\left(B_2 \setminus B_1\right) = \mu\left(B_1\right) \leq \mu\left(B_1 \cup B_1'\right) = \mu\left(B_2 \cap B_2'\right) \leq \mu\left(B_2'\right)$. By symmetry $\mu\left(B_2'\right) \leq \mu\left(B_2\right)$ and so $\mu\left(B_2\right) = \mu\left(B_2'\right)$. But $\mu\left(B_2\right) = \mu\left(B_1\right)$ and $\mu\left(B_2'\right) = \mu\left(B_1'\right)$ and also all four quantities agree.

3. It remains to show that $\tilde{\mu}$ is countably additive and complete on $\tilde{\mathcal{F}}$. Let $\{A_n\}_{n=1}^{\infty}$ be a disjoint sequence of sets from $\tilde{\mathcal{F}}$ and let $A = \bigcup_{n\geq 1} A_n$. Let $\{B_{1n}\}_{n\geq 1}, \{B_{2n}\}_{n\geq 1}, B_1, B_2$ be as in the proof of (a). Then, the fact that $\{A_n\}_{n=1}^{\infty}$ are disjoint implies $\{B_{1n}\}_{n=1}^{\infty}$ are also disjoint. And since $B_1 = \bigcup_{n\geq 1} B_{1n}$ and μ is a measure on $(\Omega, \mathcal{F})$,

$$\mu(B_1) \equiv \sum_{n=1}^{\infty} (B_{1n})$$

Also, by definition of B_{1n} 's, $\mu(B_{1n}) = \tilde{\mu}(A_n)$ for all $n \geq 1$, and by (i), $\tilde{\mu}(A) = \mu(B_1)$. Thus,

$$\tilde{\mu}(A) = \mu(B_1) = \sum_{n=1}^{\infty} (B_{1n}) = \sum_{n=1}^{\infty} \tilde{\mu}(A_n),$$

establishing the countable additivity of $\tilde{\mu}$. Next, suppose that $A \in \tilde{\mathcal{F}}$ and $\tilde{\mu}(A) = 0$. Then there exist $B_1, B_2 \in \mathcal{F}$ such that $B_1 \subset A \subset B_2$ and $\mu(B_2 \backslash B_1) = 0$. Further, by definition of $\tilde{\mu}, \mu(B_2) = \tilde{\mu}(A) = 0$. If $D \subset A$, then $\emptyset \subset D \subset B_2$ and $\mu(B_2 \backslash \emptyset) = 0$. Therefore, $D \in \tilde{\mathcal{F}}$ and hence $(\Omega, \tilde{\mathcal{F}}, \tilde{\mu})$ is complete. Finally, if $A \in \mathcal{F}$, then take $B_1 = B_2 = A$ and so, $\tilde{\mu}(A) = \mu(B_1) = \mu(A)$, and thus, $\tilde{\mu} = \mu$ on $\mathcal{F}$. Hence, the proof of the theorem is complete.

$\square$

Few Probable Questions

1. State and prove the Caratheodory extension theorem.

2. For any measure space $(\Omega, \mathcal{F}, \mu)$, is it possible to define a complete extension to the measure μ? If yes, justify your answer.

3. Let $A \in \mathcal{M}_{\mu^*}$ and $\mu^*(A) < \infty$. Show that for each $\epsilon > 0$, there exist $B_1, B_2, \ldots, B_k \in \mathcal{C}$, $k < \infty$ with $B_i \cap B_j = \emptyset$ for $1 \leq i \neq j \leq k$, such that

$$\mu^* \left(A \Delta \bigcup_{j=1}^{k} B_j \right) < \epsilon.$$

Unit 13

Course Structure

- Integrations: Measurable transformations, Induced measures, distribution functions, Integration.

13.1 Introduction

You are already aware of the notion of measurable functions in space of all Lebesgue measurable sets over $\mathbb{R}$. This can be generalised for any arbitrary measure spaces. In fact, a measurable function is a function between the underlying sets of two measurable spaces that preserves the structure of the spaces: the preimage of any measurable set is measurable. This is in direct analogy to the definition that a continuous function between topological spaces preserves the topological structure: the preimage of any open set is open.

Objectives

After reading this unit, you will be able to

- define measurable transformations and get to know several examples of them
- define induced measure
- define integration of measurable function with respect to a measure μ
- come across various convergence theorems in measure

13.2 Measurable transformations

Definition 13.2.1. Let Ω be a nonempty set and let $\mathcal{F}$ be a σ-algebra on Ω. Then the pair $(\Omega, \mathcal{F})$ is called a measurable space. If μ is a measure on $(\Omega, \mathcal{F})$, then the triple $(\Omega, \mathcal{F}, \mu)$ is called a measure space.

If in addition, μ is a probability measure, then $(\Omega, \mathcal{F}, \mu)$ is called a probability space.

Definition 13.2.2. Let $(\Omega, \mathcal{F})$ be a measurable space. Then a function $f : \Omega \to \mathbb{R}$ is called $\mathcal{F}$-measurable if for each $a \in \mathbb{R}$,

$$f^{-1}((-\infty, a]) = \{x : f(x) \leq a\} \in \mathcal{F}.$$

However, if $(\Omega, \mathcal{F})$ is a probability space. Then a function $X : \Omega \to \mathbb{R}$ is called a random variable, if the event

$$X^{-1}((-\infty, a]) = \{x : \ X(x) \leq a\} \in \mathcal{F},$$

for each $a \in \mathbb{R}$.

The following definition generalizes the above between two measurable spaces.

Definition 13.2.3. Let $(\Omega_i, \mathcal{F}_i)$, $i = 1, 2$ be measurable spaces. Then, a mapping $T : \Omega_1 \to \Omega_2$ is called measurable with respect to the σ-algebras $\langle \mathcal{F}_1, \mathcal{F}_2 \rangle$ (or $\langle \mathcal{F}_1, \mathcal{F}_2 \rangle$-measurable) if

$$T^{-1}(A) \in \mathcal{F}_1, \quad \text{for all} \quad A \in \mathcal{F}_2.$$

Example 13.2.4. Let $\Omega = \{a, b, c, d\}$, $\mathcal{F}_2 = \{\Omega, \emptyset, \{a\}, \{b, c, d\}\}$ and let $\mathcal{F}_2$ =the set of all subsets of Ω. Define the mappings $T_i : \Omega \to \Omega$, $i = 1, 2$, by

$$T_1(x) = a, \ \ x \in \Omega$$

and

$$\begin{aligned} T_2(x) &= a, \ \text{if} \ x = a, b \\ &= c, \ \text{if} \ x = c, d. \end{aligned}$$

Then T_1 is $\langle \mathcal{F}_2, \mathcal{F}_3 \rangle$-measurable since for any $A \in \mathcal{F}_3$, $T_1^{-1}(A) = \Omega$ or $\emptyset$ according as $a \in A$ or $a \notin A$. By similar arguments, it follows that T_2 is $\langle \mathcal{F}_3, \mathcal{F}_2 \rangle$-measurable. However, T_2 is not $\langle \mathcal{F}_2, \mathcal{F}_3 \rangle$-measurable since $T_2^{-1}(\{a\}) = \{a, b\} \notin \mathcal{F}_2$.

Theorem 13.2.5. Let $(\Omega_i, \mathcal{F}_i), i = 1, 2, 3$ be measurable spaces.

(i) Suppose that $\mathcal{F}_2 = \sigma\langle \mathcal{A} \rangle$ for some class of subsets $\mathcal{A}$ of Ω_2. If $T : \Omega_1 \to \Omega_2$ is such that $T^{-1}(A) \in \mathcal{F}_1$ for all $A \in \mathcal{A}$, then T is $\langle \mathcal{F}_1, \mathcal{F}_2 \rangle$ measurable.

(ii) Suppose that $T_1 : \Omega_1 \to \Omega_2$ is $\langle \mathcal{F}_1, \mathcal{F}_2 \rangle$-measurable and $T_2 : \Omega_2 \to \Omega_3$ is $\langle \mathcal{F}_2, \mathcal{F}_3 \rangle$-measunable. Let $T = T_2 \circ T_1 : \Omega_1 \to \Omega_3$ denote the composition of T_1 and T_2, defined by $T(\omega_1) = T_2(T_1(\omega_1)), \omega_1 \in \Omega_1$. Then, T is $\langle \mathcal{F}_1, \mathcal{F}_3 \rangle$-measurable.

Proof. (i) Define the collection of sets

$$\mathcal{F} = \left\{A \in \mathcal{F}_2 : T^{-1}(A) \in \mathcal{F}_1\right\}$$

Then,

(a) $T^{-1}(\Omega_2) = \Omega_1 \in \mathcal{F}_1 \Rightarrow \Omega_2 \in \mathcal{F}$.

(b) If $A \in \mathcal{F}$, then $T^{-1}(A) \in \mathcal{F}_1 \Rightarrow \left(T^{-1}(A)\right)^c \in \mathcal{F}_1 \Rightarrow T^{-1}(A^c) = \left(T^{-1}(A)\right)^c \in \mathcal{F}_1$, implying $A^c \in \mathcal{F}$.

(c) If $A_1, A_2, \ldots, \in \mathcal{F}$, then, $T^{-1}(A_i) \in \mathcal{F}_1$ for all $i \geq 1$. Since $\mathcal{F}_1$ is a σ-algebra $T^{-1}\left(\bigcup_{n\geq 1} A_n\right) = \bigcup_{n\geq 1} T^{-1}(A_n) \in \mathcal{F}_1$. Thus, $\bigcup_{n\geq 1} A_n \in \mathcal{F}$. (See also Problem 2.1 on de Morgan's laws.)

Hence, by (a), (b), (c), $\mathcal{F}$ is a σ-algebra and by hypothesis $\mathcal{A} \subset \mathcal{F}$. Hence, $\mathcal{F}_2 = \sigma\langle \mathcal{A} \rangle \subset \mathcal{F} \subset \mathcal{F}_2$. Thus, $\mathcal{F} = \mathcal{F}_2$ and T is $\langle \mathcal{F}_1, \mathcal{F}_2 \rangle -$ measurable.

(ii) Let $A \in \mathcal{F}_3$. Then, $T_2^{-1}(A) \in \mathcal{F}_2$, since T_2 is $\langle \mathcal{F}_2, \mathcal{F}_3 \rangle$-measurable. Also, by the $\langle \mathcal{F}_1, \mathcal{F}_2 \rangle$-measurability of $T_1, T^{-1}(A) = T_1^{-1}\left(T_2^{-1}(A)\right) \in \mathcal{F}_1$, showing that T is $\langle \mathcal{F}_1, \mathcal{F}_3 \rangle$-measurable.

□

13.3 Induced measures, distribution functions

Definition 13.3.1. Let $(\Omega_i, \mathcal{F}_i)$, $i = 1, 2$ be measurable spaces and let $T : \Omega_1 \to \Omega_2$ be a $\langle \mathcal{F}_1, \mathcal{F}_2 \rangle$-measurable mapping from Ω_1 to Ω_2. Then, for any measure μ on $(\Omega_1, \mathcal{F}_1)$, the set function μT^{-1}, defined by

$$\mu T^{-1}(A) = \mu\left(T^{-1}(A)\right), \quad A \in \mathcal{F}_2$$

is a measure on $\mathcal{F}_2$.

Exercise 13.3.2. Check whether μT^{-1} satisfies all the conditions for being a measure.

Definition 13.3.3. The measure μT^{-1} is called the measure induced by T (or the induced measure of T) on $\mathcal{F}_2$.

13.4 Integration

Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $f : \Omega \to \mathbb{R}$ be a measurable function. We will define the integral of f with respect to measure μ.

Definition 13.4.1. A function $f : \Omega \to \mathbb{R} \equiv [-\infty, \infty]$ is called simple if there exist a finite set (of distinct elements) $\{c_1, \ldots, c_k\} \in \mathbb{R}$ and sets $A_1, \ldots, A_k \in \mathcal{F}$, $k \in \mathbb{N}$ such that f can be written as

$$f = \sum_{i=1}^{k} c_i I_{A_i},$$

where, I_A denotes the characteristic function of the set A.

Definition 13.4.2. (The integral of a simple nonnegative function). Let $f : \Omega \to [0, \infty]$ be a simple non-negative function on $(\Omega, \mathcal{F}, \mu)$ with the representation (3.1). The integral of f w.r.t. μ, denoted by $\int f d\mu$, is defined as

$$\int f d\mu \equiv \sum_{i=1}^{k} c_i \mu\left(A_i\right).$$

It may be verified that the value of the integral above does not depend on the representation of f. That is, if f can be expressed as $f = \sum_{j=1}^{l} d_j I_{B_j}$ for some $d_1, \ldots, d_l \in \overline{\mathbb{R}}_+$(not necessarily distinct) and for some sets $B_1, \ldots, B_l \in \mathcal{F}$, then $\sum_{i=1}^{k} c_i \mu\left(A_i\right) = \sum_{j=1}^{l} d_j \mu\left(B_j\right)$, so that the value of the integral remains unchanged (verify). Also note that for the f above,

$$0 \leq \int f d\mu \leq \infty$$

The following result is an easy consequence of the definition and the above discussion.

Theorem 13.4.3. Let f and g be two simple non-negative functions on $(\Omega, \mathcal{F}, \mu)$. Then

(i) (Linearity) For $\alpha \geq 0, \beta \geq 0, \int (\alpha f + \beta g) d\mu = \alpha \int f d\mu + \beta \int g d\mu$.

(ii) (Monotonicity) If $f \geq g$ a.e. (μ), i.e., $\mu(\{x : \ x \in \Omega, f(x) < g(x)\}) = 0$, then $\int f d\mu \geq \int g d\mu$.

(iii) If $f = g$ a.e. (μ), that is, $\mu(\{x : \ x \in \Omega, \ f(x) \neq g(x)\}) = 0$, then $\int f d\mu = \int g d\mu$.

Proof. Left as exercise. □

Definition 13.4.4. (The integral of a non-negative measurable function). Let $f : \Omega \to [0, \infty]$ be a non-negative measurable function on $(\Omega, \mathcal{F}, \mu)$. The integral of f with respect to μ, also denoted by $\int f d\mu$, is defined as

$$\int f d\mu = \lim_{n\to\infty} f_n d\mu \tag{13.4.1}$$

where $\{f_n\}$ is any sequence of non-negative simple functions such that $f_n(x) \uparrow f(x)$ for all x.

The sequence $\{f_n\}$ is is non-decreasing, and hence the right side of the above equation is well defined, that is, it is the same for all such approximating sequences of functions as established in the theorem below.

Theorem 13.4.5. Let $\{f_n\}_{n\geq 1}$ and $\{g_n\}_{n\geq 1}$ be two sequences of simple non-negative measurable functions on $(\Omega, \mathcal{F}, \mu)$ to $[0, \infty]$such that as $n \to \infty$, $f_n(x) \uparrow f(x)$ and $g_n(x) \uparrow f(x)$ for all $x \in \Omega$. Then

$$\lim_{n\to\infty} \int f_n d\mu = \lim_{n\to\infty} \int g_n d\mu$$

Proof. Fix $N \in \mathbb{N}$ and $0 < \rho < 1$. It will now be shown that

$$\lim_{n\to\infty} \int f_n d\mu \geq \rho \int g_N d\mu. \tag{13.4.2}$$

Suppose that g_N has the representation $g_N \equiv \sum_{i=1}^{k} d_i I_{B_1}$. Let $D_n = \{x \in \Omega : f_n(x) \geq \rho g_N(x)\}, n \geq 1$. Since $f_n(x) \uparrow f(x)$ for all $x, D_n \uparrow D \equiv \{x : f(x) \geq \rho g_N(x)\}$. Also since $g_N(x) \leq f(x)$ and $0 < \rho < 1, D = \Omega$. Now writing $f_n = f_n I_{D_n} + f_n I_{D_n^c}$, it follows from the previous theorem that

$$\begin{aligned}
\int f_n d\mu &\geq \int f_n I_{D_n} d\mu \geq \rho \int g_N I_{D_n} d\mu \\
&= \rho \sum_{i=1}^{k} d_i \mu\left(B_i \cap D_n\right).
\end{aligned} \tag{13.4.3}$$

By the m.c.f.b. property, for each $i \in \mathbb{N}, \mu\left(B_i \cap D_n\right) \uparrow \mu\left(B_i \cap \Omega\right) = \mu\left(B_i\right)$ as $n \to \infty$. Since the sequence $\left\{\int f_n d\mu\right\}_{n\geq 1}$ is non-decreasing, taking limits in (13.4.3), we get (13.4.2). Next, letting $\rho \uparrow 1$ yields

$$\lim_{n\to\infty} \int f_n d\mu \geq \int g_N d\mu$$

for each $N \in \mathbb{N}$ and hence,

$$\lim_{n\to\infty} \int f_n d\mu \geq \int g_n d\mu.$$

By symmetry, we get the desired result. □

We should make a remark of an alternative definition of the integral of a non-negative measurable function which is given as follows.

$$\int f d\mu = \sup \left\{ \int g d\mu : \ g \ \text{ non-negative and simple, } g \leq f \right\}.$$

This definition is equivalent to (13.4.1) (verify). Also it needs to be mentioned that the properties of linearity, monotonicity and non-negativity are valid for the integrals of non-negative measurable functions as well.

Theorem 13.4.6. (The monotone convergence theorem or MCT). Let $\{f_n\}_{n\geq 1}$ and f be non-negative measurable functions on $(\Omega, \mathcal{F}, \mu)$ such that $f_n \uparrow f$ a.e. (μ). Then

$$\int f d\mu = \lim_{n\to\infty} \int f_n d\mu$$

Proof. Let $\{g_n\}_{n\geq 1}$ be a sequence of non-negative simple functions on $(\Omega, \mathcal{F}, \mu)$ such that $g_n(x) \uparrow f(x)$ for all x. By hypothesis, there exists a set $A \in \mathcal{F}$ such that $\mu(A^c) = 0$ and for x in $A, f_n(x) \uparrow f(x)$. Fix $k \in \mathbb{N}$ and $0 < \rho < 1$. Let $D_n = \{x : \ x \in A, f_n(x) \geq \rho g_k(x)\}, n \geq 1$. Then, $D_n \uparrow D \equiv \{x : x \in A, f(x) \geq \rho g_k(x)\}$. Since $g_k(x) \leq f(x)$ for all x, it follows that $D = A$. Now, by the non-negativity of the integral of non-negative measurable functions, we get

$$\int f_n d\mu \geq \int f_n I_{D_\mathrm{n}} d\mu \geq \rho \int g_k I_{D_\mathrm{n}} d\mu \text{ for all } n \geq 1.$$

By m.c.f.b., $\int g_k I_{D_n} d\mu \uparrow \int g_k I_A d\mu = \int g_k d\mu$ as $n \to \infty$, yielding

$$\liminf_{n\to\infty} \int f_n d\mu \geq \rho \int g_k d\mu,$$

for all $0 < \rho < 1$ and all $k \in \mathbb{N}$. Letting $\rho \uparrow 1$ first and then $k \uparrow \infty$, from the definition of integral of non-negative measurable function, one gets

$$\liminf_{n\to\infty} \int f_n d\mu \geq \int f d\mu.$$

By monotonicity,

$$\int f_n d\mu \leq \int f d\mu \ \text{ for all } n \geq 1$$

and the proof is done. $\square$

Corollary 13.4.7. Let $\{h_n\}_{n\geq 1}$ be a sequence of non-negative measurable functions on a measure space $(\Omega, \mathcal{F}, \mu)$. Then

$$\int \left(\sum_{n=1}^{\infty} h_n\right) d\mu = \sum_{n=1}^{\infty} \int h_n d\mu.$$

Proof. Let $f_n = \displaystyle\sum_{i=1}^{n} h_i, n \geq 1$, and let $f = \sum_{i=1}^{\infty} h_i$. Then, $0 \leq f_n \uparrow f$. By the MCT,

$$\int f_n d\mu \uparrow \int f d\mu$$

But by linearity of integrals,

$$\int f_n d\mu = \sum_{i=1}^{n} \int h_i d\mu.$$

Hence, the result follows. $\square$

Corollary 13.4.8. Let f be a non-negative measurable function on a measurable space $(\Omega, \mathcal{F}, \mu)$. For $A \in \mathcal{F}$, let

$$\nu(A) \equiv \int f I_A d\mu.$$

Then, ν is a measure on $(\Omega, \mathcal{F})$.

Proof. Let $\{A_n\}_{n\geq 1}$ be a sequence of disjoint sets in $\mathcal{F}$. Let $h_n = fI_{A_n}$ for $n \geq 1$. Then by the preceding result,

$$\begin{aligned}
\nu\left(\bigcup_{n\geq 1} A_n\right) &= \int fI_{\left[\bigcup_{n\geq 1} A_n\right]} d\mu = \int f \cdot \left[\sum_{n=1}^{\infty} I_{A_n}\right] d\mu \\
&= \int \left[\sum_{n=1}^{\infty} h_n\right] d\mu = \sum_{n=1}^{\infty} \int h_n d\mu = \sum_{n=1}^{\infty} \nu\left(A_n\right).
\end{aligned}$$

□

Theorem 13.4.9. (Fatou's lemma). Let $\{f_n\}_{n\geq 1}$ be a sequence of non-negative measurable functions on $(\Omega, \mathcal{F}, \mu)$. Then

$$\liminf_{n\to\infty} \int f_n d\mu \geq \int \liminf_{n\to\infty} f_n d\mu.$$

Proof. Let $g_n(x) = \inf\{f_j(x) : \ j \geq n\}$. Then $\{g_n\}_{n\geq 1}$ is a sequence of non-negative, non-decreasing measurable functions on $(\Omega, \mathcal{F}, \mu)$ such that $g_n(x) \uparrow g(x) \equiv \liminf\limits_{n\to\infty} f_n(x)$. By MCT,

$$\int g_n d\mu \uparrow \int g d\mu.$$

But by monotonicity,

$$\int f_n d\mu \geq \int g_n d\mu$$

for each $n \geq 1$ and hence the result follows. □

Definition 13.4.10. (The integral of a measurable function). Let f be a real valued measurable function on a measure space $(\Omega, \mathcal{F}, \mu)$. Let $f^+ = fI_{f\geq 0}$ and $f^- = -fI_{f<0}$. The integral of f with respect to μ, denoted by $\int f d\mu$, is defined as

$$\int f d\mu = \int f^+ d\mu - \int f^- d\mu,$$

provided that at least one of the integrals on the right side is finite.

It is to be noted that both f^+ and f^- are non-negative measurable functions and $f = f^+ - f^-$ and $|f| = f^+ + f^-$. Finally we are in a position to define integrable functions.

Definition 13.4.11. A measurable function f on a measure space $(\Omega, \mathcal{F}, \mu)$ is said to be integrable with respect to μ if $\displaystyle\int |f| d\mu < \infty$.

Since $|f| = f^+ + f^-$, it follows that f is integrable if and only if both f^+ and f^- are integrable, that is, $\displaystyle\int f^+ d\mu < \infty$ and $\displaystyle\int f^- d\mu < \infty$. Finally, if $A \subset \mathcal{F}$, then the integral of f over A with respect to μ is denoted by $\displaystyle\int_A f d\mu$ and is defined by

$$\int_A f d\mu = \int f I_A d\mu$$

provided the right side is well defined. We shall conclude this unit with this definition. In the next unit, we shall explore further properties related to integrals.

Few Probable Questions

1. Show that the integral of non-negative measurable functions satisfies linearity.
2. State and prove Monotone Convergence theorem.
3. State and prove Fatou's lemma.
4. Define integrable function. Let f and g be two integrable functions such that $f = g$ a.e. (μ). Can we say whether $\int f d\mu = \int g d\mu$.

Unit 14

Course Structure

- More on Convergence

14.1 Introduction

In the previous unit, we defined the integration of a measurable function starting with the simple functions and developing using non-negative measurable functions. It is imperative that the definition of integral that we defined in the previous unit satisfies the property of linearity, monotonicity, etc. and is a routine exercise. In this unit, we shall explore further properties of integral such as the convergence theorems.

Objectives

After reading this unit, you will be able to

- discuss additional properties of integrals
- learn various definitions of convergence in measure spaces and explore their properties

14.2 Further properties related to Integration

Theorem 14.2.1. Let f be a measurable function on $(\Omega, \mathcal{F}, \mu)$ and let f be non-negative a.e. (μ). Then

$$\int f d\mu = 0 \quad \text{iff} \quad f = 0 \quad \text{a.e. } (\mu).$$

Proof. If $f = 0$ a.e. (μ), then the result is trivially true (verify!). For the converse part, let $D = \{\omega : f(\omega) > 0\}$ and $D_n = \left\{\omega : f(\omega) > \frac{1}{n}\right\}, n \geq 1$. Then $D = \bigcup_{n\geq 1} D_n$. Since $f \geq f I_{D_n}$ a.e. (μ),

$$0 = \int f d\mu \geq \int f I_{D_n} d\mu \geq \frac{1}{n}\mu\left(D_n\right) \Rightarrow \mu\left(D_n\right) = 0 \quad \text{for each} \quad n \geq 1.$$

Also $D_n \uparrow D$ and so by m.c.f.b.,

$$\mu(D) = \lim_{n\to\infty} \mu\left(D_n\right) = 0.$$

Hence, the result follows. □

Theorem 14.2.2. If f is integrable over a measure space $(\Omega, \mathcal{F}, \mu)$, then $|f| < \infty$ a.e. (μ).

Proof. Let $C_n = \{x : |f(x)| > n\}$, $n \geq 1$ and let $C = \{x : |f(x)| = \infty\}$. Then $C_n \downarrow C$ and

$$\int |f| d\mu \geq \int |f| I_{C_n} d\mu \geq n\mu(C_n)$$

which implies that $\mu(C_n) \leq \dfrac{\int |f| d\mu}{n}$. Since $\displaystyle\int |f| d\mu < \infty$, $\lim_{n\to\infty} \mu(C_n) = 0$. Hence, by m.c.f.a, $\mu(C) = \lim_{n\to\infty} \mu(C_n) = 0$. $\square$

We are now in a position to prove the extended dominated convergence theorem.

Theorem 14.2.3. (The extended dominated convergence theorem or EDCT). Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and let $f_n, g_n : \Omega \to \mathbb{R}$ be $\langle \mathcal{F}, \mathbb{R} \rangle$-measurable functions such that $|f_n| \leq g_n$ a.e. (μ) for all $n \geq 1$. Suppose that

(i) $g_n \to g$ a.e. (μ) and $f_n \to f$ a.e. (μ);

(ii) g_n, g are integrable and $\int |g_n| \, d\mu \to \int |g| d\mu$ as $n \to \infty$.

Then, f is integrable and

$$\lim_{n\to\infty} \int f_n d\mu = \int f d\mu \quad \text{and} \quad \lim_{n\to\infty} \int |f_n - f| \, d\mu = 0.$$

Proof. By Fatou's lemma,

$$\int |f| d\mu \leq \liminf_{n\to\infty} \int |f_n| \, d\mu \leq \liminf_{n\to\infty} \int |g_n| \, d\mu = \int |g| d\mu < \infty.$$

Hence, f is integrable. For proving the second part, let $h_n = f_n + g_n$ and $\gamma_n = g_n - f_n$, $n \geq 1$. Then $\{h_n\}$ and $\{\gamma_n\}$ are sequences of non-negative integrable functions. By Fatou's lemma and (ii),

$$\begin{aligned}
\int (f+g) d\mu &= \int \liminf_{n\to\infty} h_n d\mu \\
&\leq \liminf_{n\to\infty} \int h_n d\mu \\
&= \liminf_{n\to\infty} \left[\int g_n d\mu + \int f_n d\mu \right] \\
&= \int g d\mu + \liminf_{n\to\infty} \int f_n d\mu.
\end{aligned}$$

Similarly,

$$\int (g - f) d\mu \leq \int g d\mu - \limsup_{n\to\infty} \int f_n d\mu.$$

By the linearity of integrals we have from the above two equations,

$$\int f d\mu \leq \liminf_{n\to\infty} \int f_n d\mu$$
$$\limsup_{n\to\infty} \int f_n d\mu \leq \int f d\mu,$$

which yields

$$\lim_{n\to\infty} \int f_n d\mu = \int f d\mu.$$

For the last part, one should apply the above argument to f_n and g_n replaced by $|f - f_n|$ and $g_n + |f|$ respectively. $\square$

Corollary 14.2.4. (Lebesgue's dominated convergence theorem, or DCT). If $|f_n| \leq g$ a.e. (μ) for all $n \geq 1$, $\int g d\mu < \infty$ and $f_n \to f$ a.e. (μ), then f is integrable and

$$\lim_{n\to\infty} \int f_n d\mu = \int f d\mu \text{ and } \lim_{n\to\infty} \int |f_n - f| \, d\mu = 0.$$

Corollary 14.2.5. (The bounded convergence theorem, or BCT). Let $\mu(\Omega) < \infty$. If there exist a $0 < k < \infty$ such that $|f_n| \leq k$ a.e. (μ) and $f_n \to f$ a.e. (μ), then

$$\lim_{n\to\infty} \int f_n d\mu = \int f d\mu \text{ and } \lim_{n\to\infty} \int |f_n - f| \, d\mu = 0.$$

Proof. Take $g(\omega) \equiv k$ for all $\omega \in \Omega$ in the previous corollary. □

Exercise 14.2.6. If f is integrable, check whether $|f| < \infty$ a.e. (μ).

14.3 More properties related to Convergence

Let $\{f_n\}_{n\geq 1}$ and f be measurable functions from a measure space $(\Omega, \mathcal{F}, \mu)$ to $[0, \infty]$. Let us define the notion of convergence on the measure space.

Definition 14.3.1. $\{f_n\}_{n\geq 1}$ is said to converge pointwise to f if

$$\lim_{n\to\infty} f_n(x) = f(x), \;\; \forall x \in \Omega.$$

Definition 14.3.2. $\{f_n\}_{n\geq 1}$ is said to converge to f almost everywhere (μ), denoted by $f_n \to f$ a.e. (μ), if there exists a set $B \in \mathcal{F}$ such that $\mu(B) = 0$ and

$$\lim_{n\to\infty} f_n(x) = f(x), \;\; \forall x \in B^c.$$

Let us consider some more notions of convergence.

Definition 14.3.3. $\{f_n\}_{n\geq 1}$ is said to converge to f in measure (w.r.t. μ) denoted by $f_n \xrightarrow{m} f$ if for each $\epsilon > 0$,

$$\lim_{n\to\infty} \mu\left(\{|f_n - f| > \epsilon\}\right) = 0.$$

Definition 14.3.4. $\{f_n\}_{n\geq 1}$ is said to converge to f uniformly over Ω if

$$\lim_{n\to\infty} \sup\{|f_n(x) - f(x)| : \; x \in \Omega\} = 0.$$

Definition 14.3.5. $\{f_n\}_{n\geq 1}$ is said to converge to f nearly uniformly (μ) if for every $\epsilon > 0$, there exists a set $A \in \mathcal{F}$ such that $\mu(A) < \epsilon$ and $f_n \to f$ uniformly on A^c, $f_n \to f$ uniformly , that is,

$$\sup\{|f_n(x) - f(x)| : \; x \in A^c\} \to 0$$

as $n \to \infty$.

Theorem 14.3.6. Suppose that $\mu(\Omega) < \infty$. Then $f_n \to f$ a.e. (μ) implies $f_n \xrightarrow{m} f$.

Proof. Left as an exercise. □

Theorem 14.3.7. Let $f_n \xrightarrow{m} f$. Then there exists a subsequence $\{n_k\}$ such that $f_{n_k} \xrightarrow{m} f$ a.e. (μ).

Proof. Since $f_n \xrightarrow{m} f$, for each integer $k \geq 1$, there exists an n_k such that for all $n \geq n_k$,

$$\mu\left(\{|f_n - f| > 2^{-k}\}\right) < 2^{-k}. \tag{14.3.1}$$

Without any loss of generality assume that $n_{k+1} > n_k$ for all $k \geq 1$. Let $A_k = \{|f_n - f| > 2^{-k}\}$. Then by corollary 13.4.7,

$$\int \left(\sum_{k=1}^{\infty} I_{A_k}\right) d\mu = \sum_{k=1}^{\infty} \int I_{A_k} d\mu = \sum_{k=1}^{\infty} \mu(A_k),$$

which is finite by (14.3.1). Hence, by theorem 14.2.2, $\sum_{k=1}^{\infty} I_{A_k} < \infty$ a.e. (μ). We observe that

$$\sum_{k=1}^{\infty} I_{A_k}(x) < \infty \Rightarrow |f_{n_k}(x) - f(x)| \leq 2^{-k}$$

for all large k which implies that

$$\lim_{k\to\infty} f_{n_k}(x) = f(x).$$

Hence the result. □

Theorem 14.3.8. (Scheffe's theorem). Let $\{f_n\},\ f$ be a collection of nonnegative measurable functions on a measure space $(\Omega, \mathcal{F}, \mu)$. Let $f_n \to f$ a.e. (μ), $\int f_n d\mu \to \int f d\mu$ and $\int f d\mu < \infty$. Then

$$\lim_{n\to\infty} \int |f_n - f| d\mu = 0.$$

Proof. Let $g_n = f - f_n$, $n \geq 1$. Since $f_n \to f$ a.e. (μ), both g_n^+ and g_n^- converge to zero a.e. (μ). Further, $0 \leq g_n^+ \leq f$ and by hypothesis $\int f d\mu < \infty$. Thus, by the dominated convergence theorem, it follows that

$$\int g_n^+ d\mu \to 0.$$

Next, by hypothesis we note that $\int g_n d\mu \to 0$. Thus,

$$\int g_n^- d\mu = \int g_n^+ d\mu - \int g_n d\mu \to 0$$

and hence

$$\int |g_n| d\mu = \int g_n^+ d\mu + \int g_n^- d\mu.$$

□

Theorem 14.3.9. Let f be integrable over the measure space $(\Omega, \mathcal{F}, \mu)$. Then for every $\epsilon > 0$, there exists a $\delta > 0$ such that $\mu(A) < \delta \Rightarrow \int_A |f| d\mu < \epsilon$.

Proof. Fix a number $\epsilon > 0$. By dominated convergence theorem, there exists a $t > 0$ such that $\int_{\{|f|>t\}} |f| d\mu < \epsilon/2$. Hence, for any $A \in \mathcal{F}$ with $\mu(A) \leq \delta = \frac{\epsilon}{2t}$,

$$\begin{aligned} \int_A |f| d\mu &\leq \int_{A\cap\{|f|>t\}} |f| d\mu + \int_{\{|f|>t\}} |f| d\mu \\ &\leq t\mu(A) + \int_{\{|f|>t\}} |f| d\mu \\ &\leq \epsilon. \end{aligned}$$

Hence the theorem. □

Few Probable Questions

1. State and prove the extended dominated convergence theorem.
2. Define convergence in measure. If $\mu(\Omega) < \infty$ that show that $f_n \to f$ a.e. (μ) implies $f_n \xrightarrow{m} f$.
3. State and prove Scheffe's theorem.

Unit 15

Course Structure

- L^p-Spaces, Dual spaces, Banach and Hilbert spaces

15.1 Introduction

L^p spaces are the special spaces of measurable functions over a given measure space $(\Omega, \mathcal{F}, \mu)$. Let us start with the definition of L^p space. Let $0 < p < \infty$. Then

$$\begin{aligned} L^p(\Omega, \mathcal{F}, \mu) &= \{f : |f|^p \text{ is integrable with respect to } \mu\} \\ &= \left\{f : \int |f|^p d\mu < \infty\right\}. \end{aligned}$$

Also, we can define

$$L^\infty(\Omega, \mathcal{F}, \mu) = \{f : \mu(\{|f| > K\}) = 0 \text{ for some } K \in (0, \infty)\}.$$

For the sake of simplicity, if the underlying measure space is specified, we will only use the notation L^p. We will check upon the properties of L^p spaces in details.

Objectives

After reading this unit, you will be able to

- define L^p spaces and develop it into a complete metric space
- define dual space of L^p

15.2 L^p Spaces

Unless otherwise specified, we will take the measure space to be $(\Omega, \mathcal{F}, \mu)$. We start with the following theorem.

Theorem 15.2.1. Let $f, g \in L^1$. Then

1. $\displaystyle\int (\alpha f + \beta g) d\mu = \alpha \int f d\mu + \beta \int g d\mu$ for any $\alpha, \beta \in \mathbb{R}$;

2. $f \geq g$ a.e. (μ) implies $\displaystyle\int f d\mu \geq \int g d\mu$;

3. $f = g$ a.e. (μ) implies $\displaystyle\int f d\mu = \int g d\mu$.

The following result will show that L^p is a vector space over $\mathbb{R}$ for general $0 < p \leq \infty$.

Theorem 15.2.2. For every $0 < p \leq \infty$,

$$f, g \in L^p \Rightarrow af + bg \in L^p, \quad \text{for } a, b \in \mathbb{R}.$$

Let us define the following for convergence in L^p space.

Definition 15.2.3. Let $1 \leq p < \infty$. Then $\{f_n\}$ converges to f in L^p, denoted by $f_n \xrightarrow{L^p} f$, if $\displaystyle\int |f_n|^p d\mu < \infty$ for all $n \geq 1$, $\displaystyle\int |f|^p d\mu < \infty$ and

$$\lim_{n \to \infty} \int |f_n - f|^p d\mu = 0.$$

Clearly, the above is equivalent to $\|f_n - f\|_p \to 0$ as $n \to \infty$, where, for any $\mathcal{F}$ measurable function g and any $1 \leq p < \infty$,

$$\|g\|_p = \left(\int |g|^p d\mu \right)^{\min\{\frac{1}{p}, 1\}}.$$

For $p = 1$, this is also called convergence in absolute deviation and for $p = 2$, convergence in mean square. Further, $\{f_n\}$ converges to f in L^∞ if

$$\lim_{n \to \infty} \|f_n - f\|_\infty = 0,$$

where for any $\mathcal{F}$ measurable function g on $(\Omega, \mathcal{F}, \mu)$,

$$\|g\|_\infty = \inf\{K : \ K \in (0, \infty), \ \mu(\{|g| > K\}) = 0\}.$$

Theorem 15.2.4. Let $\{f_n\}, f$ be measurable functions on a measure space $(\Omega, \mathcal{F}, \mu)$. Let $f_n \xrightarrow{L^p} f$ for some $1 \leq p < \infty$. Then $f_n \xrightarrow{m} f$.

Proof. For each $\epsilon > 0$, let $A_n = \{|f_n - f| \geq \epsilon\}, \ n \geq 1$. Then

$$\int |f_n - f|^p d\mu \geq \int_{A_n} |f_n - f|^p d\mu \geq \epsilon^p \mu(A_n).$$

Since $f_n \to f$ in L^p, $\displaystyle\int |f_n - f|^p d\mu \to 0$ and hence $\mu(A_n) \to 0$. □

We are familiar with the idea of metric spaces and the distance function in a metric space. Let us first define the norm function in L^p.

Definition 15.2.5. Let $f \in L^p$, where $0 < p < \infty$. Then

$$\|f\|_p = \left(\int |f|^p d\mu \right)^{\min\{\frac{1}{p}, 1\}}$$

and for $p = \infty$,

$$\|f\|_\infty = \sup\{k : \ \mu(\{|f| > k\}) > 0\}$$

which is called the essential supremum of f.

Theorem 15.2.6. For $f, g \in L^p$, $0 < p \le \infty$, let

$$d_p(f,g) = \|f-g\|_p.$$

Then for any $f, g, h \in L^p$, $1 \le p \le \infty$,

1. $d_p(f,g) = d_p(g,f) \ge 0$;
2. $d_p(f,h) \le d_p(f,g) + d_p(g,h)$.

However, $d_p(f,g) = 0$ implies only that $f = g$ a.e. (μ). Thus, the above theorem says that the function d_p satisfies two conditions for being a metric. Also, it will satisfy the last condition of being a metric if we additionally define the following.

Definition 15.2.7. For $f, g \in L^p$, f is called equivalent to g and is written as $f \sim g$ if $f = g$ a.e. (μ).

It is easy to verify that the relation '$\sim$' s an equivalence relation. Thus, it partitions L^p into disjoint equivalence classes such that in each class all elements are equivalent. The notion of distance between these classes may be defined as follows:

$$d_p([f],[g]) = d_p(f,g),$$

where $[f]$ and $[g]$ denote, respectively, the equivalence classes of functions containing f and g. It can be verified that this is a metric on the set of equivalence classes. In what follows, the equivalence class $[f]$ is identified with the element f. With this identification, (L^p, d_p) becomes a metric space for $1 \le p \le \infty$.

However, for $0 < p < 1$, if we define

$$d_p(f,g) = \int |f-g|^p d\mu,$$

then (L^p, d_p) becomes a metric space.

15.2.1 Some Useful Results

Now let us state certain useful results that will be useful in the sequel.

Theorem 15.2.8. (Markov's inequality). Let f be a nonnegative measurable function on a measure space $(\Omega, \mathcal{F}, \mu)$. Then for any $0 < t < \infty$,

$$\mu(\{f \ge t\}) \le \frac{\int f d\mu}{t}.$$

Definition 15.2.9. A function $\phi : (a,b) \to \mathbb{R}$ is called convex if for all $0 \le \lambda \le 1$, $a < x \le y < b$,

$$\phi(\lambda x + (1-\lambda)y) \le \phi(x) + (1-\lambda)\phi(y).$$

Geometrically, this means that for the graph of $y = \phi(x)$ on (a,b), for each fixed $t \in (0,\infty)$, the chord over the interval $(x, x+t)$ turns in the counterclockwise direction as x increases.

More precisely, the following result holds.

Theorem 15.2.10. Let $\phi : (a,b) \to \mathbb{R}$. Then ϕ is convex if and only if $a < x_1 < x_2 < x_3 < b$,

$$\frac{\phi(x_2)-\phi(x_1)}{x_2-x_1} \le \frac{\phi(x_3)-\phi(x_2)}{x_3-x_2},$$

which is equivalent to

$$\frac{\phi(x_2)-\phi(x_1)}{x_2-x_1} \le \frac{\phi(x_3)-\phi(x_1)}{x_3-x_1} \le \frac{\phi(x_3)-\phi(x_2}{x_3-x_2}.$$

Let us state the following properties of a convex function which can be deduced from the above theorem.

Theorem 15.2.11. Let $\phi : (a, b) \to \mathbb{R}$ be convex. Then

1. For each $x \in (a, b)$,

$$\phi'_{+}(x) = \lim_{y \downarrow x} \frac{\phi(y) - \phi(x)}{y - x}, \quad \phi'_{-}(x) = \lim_{y \uparrow x} \frac{\phi(y) - \phi(x)}{y - x}$$

   exist and are finite.

2. Further, $\phi'_{-}(\cdot) \leq \phi'_{+}(\cdot)$ and both are nondecreasing on (a, b).

3. $\phi'(\cdot)$ exists except on the countable set of discontinuity points of ϕ'_{+} and ϕ'_{-} .

4. For any $a < c < d < b$, ϕ is Lipschitz on $[c, d]$, that is, here exists a constant $K < \infty$ such that

$$|\phi(x) - \phi(y)| \leq K|x - y|$$

   for all $c \leq x,\ y \leq d$.

5. For any $a < c,\ x < b$,

$$\phi(x) - \phi(c) \geq \phi'_{+}(c)(x - c) \text{ and } \phi(x) - \phi(c) \geq \phi'_{-}(c)(x - c).$$

Theorem 15.2.12. (Holder's inequality). Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. Let $1 < p < \infty$, $f \in L^p$ and $g \in L^q$, where $q = \dfrac{p}{p-1}$. Then

$$\int |fg| d\mu \leq \left(\int |f|^p d\mu \right)^{\frac{1}{p}} \left(\int |g|^q d\mu \right)^{\frac{1}{q}},$$

that is,

$$\|fg\|_1 \leq \|f\|_p \|g\|_q.$$

If $\|fg\|_1 \neq 0$, then the inequality in the first equation above holds if and only if $|f|^p = c|g|^q$ a.e. (μ) for some constant $c \in (0, \infty)$.

For $p = 1$ or $q = \infty$, we have the following.

$$\|fg\|_1 = \int |fg| d\mu \leq \|f\|_1 \|g\|_\infty.$$

If equality holds, then $|f|(\|g\|_\infty - |g|) = 0$ a.e. (μ) and hence $|g| = \|g\|_\infty$ on the set $\{|f| > 0\}$ a.e. (μ).

Corollary 15.2.13. (Cauchy-Schwarz inequality). Let $f, g \in L^2$. Then

$$\int |fg| d\mu \leq \left(\int |f|^2 d\mu \right)^{\frac{1}{2}} \left(\int |g|^2 d\mu \right)^{\frac{1}{2}},$$

that is,

$$\|fg\|_1 \leq \|f\|_2 \|g\|_2.$$

As an application of Holder's inequality, one can get the following.

Theorem 15.2.14. (Minkowski's inequality). Let $1 < p < \infty$ and $f, g \in L^p$. Then

$$\left(\int |f+g|^p d\mu\right)^{\frac{1}{p}} \leq \left(\int |f|^p d\mu\right)^{\frac{1}{p}} + \left(\int |g|^p d\mu\right)^{\frac{1}{p}}$$

that is,

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

Now let us discuss the properties of L^p spaces in details. Recall that a metric space (X, d) is complete when every Cauchy sequence in (X, d) converges to an element in X. In what follows, we will show that L^p forms a complete metric space with respect to the metric defined.

Theorem 15.2.15. For $0 < p < \infty$, (L^p, d_p) is complete.

Proof. **Step I:**Let $\{f_n\}$ be a Cauchy sequence in L^p for $0 < p < \infty$. Let $\{\epsilon_k\}$ and $\{\delta_k\}$ be sequences of positive numbers decreasing to zero. Since $\{f_n\}$ is Cauchy, for each $k \geq 1$, there exists an integer n_k such that

$$\int |f_n - f_m|^p d\mu \leq \epsilon_k, \;\; \forall n, m \geq n_k. \tag{15.2.1}$$

Without any loss of generality, let us assume that $n_{k+1} > n_k$ for each $k \geq 1$. Then by Markov's inequality

$$\mu(\{|f_{n_{k+1}} - f_{n_k}| \geq \delta_k\}) \leq \delta_k^{-p} \int |f_{n_{k+1}} - f_{n_k}|^p d\mu \leq \delta_k^{-p} \epsilon_k. \tag{15.2.2}$$

Let $A_k = \{|f_{n_{k+1}} - f_{n_k}| \geq \delta_k\},\; k = 1, 2, \ldots$ and $A = \limsup_{k\to\infty} A_k = \bigcap_{j=1}^{\infty} \bigcup_{k \geq j} A_k$. If $\{\epsilon_k\}$ and $\{\delta_k\}$ satisfy

$$\sum_{k=1}^{\infty} \delta_k^{-p} \epsilon_k < \infty, \tag{15.2.3}$$

then by equation (15.2.2), $\sum_{k=1}^{\infty} \mu(A_k) < \infty$ and hence $\mu(A) = 0$.

Note that for $x \in A^c$, $|f_{n_{k+1}}(x) - f_{n_k}(x)| < \delta_k$ for all large k. Thus, if $\sum_{k=1}^{\infty} \delta_k < \infty$, then for $x \in A^c$, $\{f_{n_k}(x)\}$ is a Cauchy sequence in $\mathbb{R}$ and hence, it converges to some $f(x)$ in $\mathbb{R}$. Setting $f(x) = 0$ for $x \in A$ one gets

$$\lim_{k\to\infty} f_{n_k} = f \;\text{ a.e. } (\mu).$$

A choice of $\{\epsilon_k\}$ and $\{\delta_k\}$ such that $\sum_{k=1}^{\infty} \delta_k < \infty$ and (15.2.3) holds is given by $\epsilon = 2^{-(p+1)k}$ and $\delta_k = 2^{-k}$.

Step II: By Fatou's lemma, Step I, and equation (15.2.1), for any fixed $k \geq 1$,

$$\epsilon_k \geq \liminf_{j\to\infty} |f_{n_k} - f_{n_{k+j}}|^p d\mu \geq \int |f_{n_k} - f|^p d\mu.$$

Since $f_{n_k} \in L^p$, this shows that $f \in L^p$. Now letting $k \to \infty$, $\lim_{k\to\infty} d_p(f_{n_k}, f) = 0$.

Step III: By triangle inequality, for any fixed $k \geq 1$,

$$d_p(f_n, f) \leq d_p(f_n, f_{n_k}) + d_p(f_{n_k}, f).$$

By equation (15.2.1) and Step II, for $n \geq n_k$, the right hand side of the above equation is $\leq 2\epsilon$, where

$$\epsilon = \begin{cases} \epsilon_k, & 0 < p < 1 \\ \epsilon_k^{1/p}, & 1 \leq p < \infty. \end{cases}$$

Now, letting $k \to \infty$, we get, $\lim_{n\to\infty} d_p(f_n, f) = 0$. □

Exercise 15.2.16. Prove that L^∞ is a complete space.

15.3 Dual Spaces

Let $1 \leq p < \infty$. Let $g \in L^q(\mu)$, where $q = \frac{p}{(p-1)}$ if $1 < p < \infty$ and $q = \infty$ if $p = 1$. Let

$$T_g(f) = \int fg d\mu, \quad f \in L^p(\mu). \tag{15.3.1}$$

By Holder's inequality, $\int |fg| d\mu < \infty$ and so $T_g(\cdot)$ is well defined. Clearly T_g is linear, i.e.,

$$T_g(\alpha_1 f_1 + \alpha_2 f_2) = \alpha_1 T_g(f_1) + \alpha_2 T_g(f_2) \tag{15.3.2}$$

for all $\alpha_1, \alpha_2 \in \mathbb{R}$ and $f_1, f_2 \in L^p(\mu)$.

Definition 15.3.1.

1. A function $T : L^p(\mu) \to \mathbb{R}$ that satisfies (15.3.2) is called a linear functional.
2. A linear functional T on $L^p(\mu)$ is called bounded if there is a constant $c \in (0, \infty)$ such that
$$|T(f)| \leq c\|f\|_p \quad \text{for all} \quad f \in L^p(\mu).$$
3. The norm of a bounded linear functional T on $L^p(\mu)$ is defined as
$$\|T\| = \sup \{|Tf| : f \in L^p(\mu), \|f\|_p = 1\}.$$
By Holder's inequality,
$$|T_g(f)| \leq \|g\|_q \|f\|_p \quad \text{for all} \quad f \in L^p(\mu),$$
and hence, T_g is a bounded linear functional on $L^p(\mu)$. This implies that if $d_p(f_n, f) \to 0$, then
$$|T_g(f_n) - T_g(f)| \leq \|g\|_q d_p(f_n, f) \to 0,$$
i.e., T_g is continuous on the metric space $(L^p(\mu), d_p)$.

Definition 15.3.2. The set of all continuous linear functionals on L^p is called the dual space of L^p and is denoted by $(L^p)^*$.

Theorem 15.3.3. (Riesz representation theorem). Let $1 \leq p < \infty$. Let $T : L^p \to \mathbb{R}$ be linear and continuous. Then, there exists a g in L^q such that $T = T_g$, that is,

$$T(f) = T_g(f) = \int fg d\mu \ \text{ for all } f \in L^p \tag{15.3.3}$$

where $q = \dfrac{p}{p-1}$ for $1 < p < \infty$ and $q = \infty$ if $p = 1$.

Such a representation is not valid for $p = \infty$, that is, there exist continuous linear functionals T on L^∞ for which there is no $g \in L^1$ such that $T(f) = \displaystyle\int fg d\mu$ for all $f \in L^\infty$.

Few Probable Questions

1. Establish a metric function on L^p with proper justifications.
2. Show that L^p is a complete space.

Unit 16

Course Structure

- Product of two measure spaces. Fubini's theorem.

16.1 Introduction

In this unit, we will probe the question as to whether there is a natural way to define measure on the product of two sets Ω_1 and Ω_2 which reflects the structure of the original measure space.

Objectives

After reading this unit, you will be able to

- define product of measure spaces and learn related results
- state and prove the Fubini's theorem

16.2 Product of two measure spaces

Let us start with the following definition.

Definition 16.2.1. Let $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ be two measurable spaces. The set $A \times B$ with $A \in \mathcal{F}_1$ and $B \in \mathcal{F}_2$ is called a measurable rectangle. The collection of measurable rectangles will be denoted by $\mathcal{C}$. The product σ-algebra of $\mathcal{F}_1$ and $\mathcal{F}_2$ on $\Omega_1 \times \Omega_2$, denoted by $\mathcal{F}_1 \times \mathcal{F}_2$ is the smallest σ-algebra generated by $\mathcal{C}$, i.e.,

$$\mathcal{F}_1 \times \mathcal{F}_2 = \sigma\langle\{A \times B : \ A \in \mathcal{F}_1, \ B \in \mathcal{F}_2\}\rangle.$$

$(\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2)$ is called the product measurable space.

Now, if μ_1 and μ_2 are measures defined on the measurable spaces $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ respectively, then we can define the a new function μ on $\mathcal{C}$ as follows.

$$\mu(A \times B) = \mu_1(A) \cdot \mu_2(B)$$

for all $A \in \mathcal{F}_1$ and $B \in \mathcal{F}_2$. One can extend it to a measure on the algebra $\mathcal{A}$ of all finite unions of disjoint measurable rectangles simply by assigning the μ-measure of a finite union of disjoint measurable rectangles as the sum of the μ-measures of the corresponding individual measurable rectangles. Then, by the extension theorem, it can be further extended to a complete measure on a σ-algebra containing $\mathcal{F}_1 \times \mathcal{F}_2$ as defined in the above definition. However, if $A \in \mathcal{F}_1 \times \mathcal{F}_2$ is not a measurable rectangle, then we need some further approach to evaluate $\mu(A)$.

Exercise 16.2.2. Let $(\Omega_1, \mathcal{F}_1, \mu)$ and $(\Omega_1, \mathcal{F}_1, \nu)$ are two measure spaces. Define outer measure τ^* on $\Omega = \Omega_1 \times \Omega_2$ as follows

$$\tau^*(E) = \inf \left\{ \sum_{j=1}^{\infty} \mu(A_j)\nu(B_j) : \ A_j \in \mathcal{F}_1, \ B_j \in \mathcal{F}_2 \ j \in \mathbb{N} \text{ and } E \subset \bigcup_{j=1}^{\infty} A_j \times B_j \right\},$$

for every $E \subseteq \Omega$. Show that τ^* as defined above is an outer measure on Ω.

Definition 16.2.3. Let $A \in \mathcal{F}_1 \times \mathcal{F}_2$. Then, for any $x_1 \in \Omega_1$, the set

$$A_{1x_1} = \{x_2 \in \Omega_2 : \ (x_1, x_2) \in A\} \tag{16.2.1}$$

is called the x_1-section of A and for any $x_2 \in \Omega_2$, the set $A_{2x_2} = \{x_1 \in \Omega_1 : \ (x_1, x_2) \in A\}$ is called the x_2-section of A.

If $f : \Omega_1 \times \Omega_2 \to \Omega_3$ is a $\langle \mathcal{F}_1 \times \mathcal{F}_2, \mathcal{F}_3 \rangle$ measurable mapping from $\Omega_1 \times \Omega_2$ into some measurable space $(\Omega_3, \mathcal{F}_3)$, then the x_1-section of f is the function $f_{1x_1} : \Omega_2 \to \Omega_3$, given by

$$f_{1x_1}(x_2) = f(x_1, x_2), \ \ x_2 \in \Omega_2. \tag{16.2.2}$$

The x_2-sections of f can be similarly defined.

Theorem 16.2.4. Let $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2)$ be a product space, $A \in \mathcal{F}_1 \times \mathcal{F}_2$ and let $f : \Omega_1 \times \Omega_2 \to \Omega_3$ be a $\langle \mathcal{F}_1 \times \mathcal{F}_2, \mathcal{F}_3 \rangle$-measurable function.

1. For every $x_1 \in \Omega_1$, $A_{1x_1} \in \mathcal{F}_2$ and for every $x_2 \in \Omega_2$, $A_{2x_2} \in \mathcal{F}_1$.

2. For every $x_1 \in \Omega_1$, f_{1x_1} is $\langle \mathcal{F}_2, \mathcal{F}_3 \rangle$-measurable and for every $x_2 \in \Omega_2$, f_{2x_2} is $\langle \mathcal{F}_1, \mathcal{F}_3 \rangle$-measurable.

Proof. Let $x_1 \in \Omega_1$ be fixed. We define a function $g : \Omega_2 \to \Omega_1$ as follows.

$$g(x_2) = (x_1, x_2), \ x_2 \in \Omega_2.$$

Note that for any measurable rectangle $A = A_1 \times A_2 \in \mathcal{F}_1 \times \mathcal{F}_2$,

$$A_{1x_1} = \begin{cases} A_2, & \text{if } x_1 \in A_1 \\ \emptyset, & \text{if } x_1 \notin A_1 \end{cases}$$

and hence $g^{-1}(A_1 \times A_2) \in \mathcal{F}_2$. Since the class of all measurable rectangles generates $\mathcal{F}_1 \times \mathcal{F}_2$, for fixed $x_1 \in \Omega_1$, g is $\langle \mathcal{F}_1, \mathcal{F}_1 \times \mathcal{F}_2 \rangle$-measurable. Therefore, for $A \in \mathcal{F}_1 \times \mathcal{F}_2$, $A_{1x_1} = g^{-1}(A) \in \mathcal{F}_2$ and for f as given, $f_{1x_1} = f \circ g$ is $\langle \mathcal{F}_2, \mathcal{F}_3 \rangle$-measurable. This proves 1 and 2 for x_1-sections. Similar proof follows for x_2-sections. $\square$

Next we suppose that μ_1 and μ_2 are measures on $(\Omega_1, \mathcal{F}_1)$ and $(\Omega_2, \mathcal{F}_2)$ respectively. Then for any set $A \in \mathcal{F}_1 \times \mathcal{F}_2$, for all $x_1 \in \Omega_1$, $A_{1x_1} \in \mathcal{F}_2$ and hence $\mu_2(A_{1x_1})$ is well defined. If it were an $\mathcal{F}_1$-measurable function, then one might have defined a set function on $\mathcal{F}_1 \times \mathcal{F}_2$ by

$$\mu_{12}(A) = \int_{\Omega_1} \mu_2(A_{1x_1}\mu_1 dx_1. \tag{16.2.3}$$

Similarly, reversing the order of μ_1 and μ_2, we may define the set function

$$\mu_{21}(A) = \int_{\Omega_2} \mu_1(A_{2x_2}\mu_2 dx_2. \tag{16.2.4}$$

provided that $\mu_1(A_{2x_2})$ is $\mathcal{F}_2$-measurable. Also, it should be noted that for the measurable rectangles $A = A_1 \times A_2$, $\mu_{12}(A) = \mu_1(A_1)\mu_2(A_2) = \mu_{21}(A)$ and thus both μ_{12} and μ_{21} coincide with the product measure μ on the class $\mathcal{C}$ of all measurable rectangles. This implies that if the product measure μ is unique on $\mathcal{F}_1 \times \mathcal{F}_2$, and μ_{12} and μ_{21} are measures on $\mathcal{F}_1 \times \mathcal{F}_2$, then they must be equal to μ on $\mathcal{F}_1 \times \mathcal{F}_2$. Then one can evaluate $\mu(A)$ for any set $A \in \mathcal{F}_1 \times \mathcal{F}_2$ using the equations (16.2.3) and (16.2.4).

Theorem 16.2.5. Let $(\Omega_1, \mathcal{F}_1, \mu_1)$ and $(\Omega_2, \mathcal{F}_2, \mu_2)$ be σ-finite measure spaces. Then

1. for all $A \in \mathcal{F}_1 \times \mathcal{F}_2$, the functions $\mu_2(A_{1x_1})$ and $\mu_1(A_{2x_2})$ are respectively $\mathcal{F}_1$ and $\mathcal{F}_2$-measurable.

2. The functions μ_{12} and μ_{21} as given in (16.2.3) and (16.2.4), are measures on $\mathcal{F}_1 \times \mathcal{F}_2$ satisfying $\mu_{12}(A) = \mu_{21}(A)$ for all $A \in \mathcal{F}_1 \times \mathcal{F}_2$.

3. Further, $\mu_{12} = \mu_{21} = \mu$ is σ-finite and it is the only measure satisfying

$$\mu(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2) \text{ for all } A_1 \times A_2 \in \mathcal{C}.$$

Proof. First let us assume that μ_1 and μ_2 are finite measures. Also, let

$$\mathcal{S} = \{A \in \mathcal{F}_1 \times \mathcal{F}_2 : \ \mu_2(A_{1x_1}) \text{ is a } \langle \mathcal{F}_1, \mathcal{B}(\mathcal{R})\rangle - \text{measurable function}\}.$$

For $A = \Omega_1 \times \Omega_2$, $\mu_2(A_{1x_1}) = \mu_2(\Omega_2)$ for all $x_1 \in \Omega_1$ and hence $\Omega_1 \times \Omega_2 \in \mathcal{S}$. Next, let $A, B \in \mathcal{S}$ with $A \subset B$. Then we can check that

$$(A \setminus B)_{1x_1} = A_{1x_1} \setminus B_{1x_1}.$$

Since μ_2 is finite and $A, B \in \mathcal{S}$ so

$$\mu_2((A \setminus B)_{1x_1}) = \mu_2(A_{1x_1} \setminus B_{1x_1}) = \mu_2(A_{1x_1}) - \mu_2(B_{1x_1})$$

is $\langle \mathcal{F}_1, \mathcal{B}(\mathbb{R})\rangle$-measurable. Thus, $A \setminus B \in \mathcal{S}$. Finally, let $\{B_n\}$ be a monotonically increasing sequence of sets in $\mathcal{S}$. Then, for any $x_1 \in \Omega_1$, $(B_n)_{1x_1} \subset (B_{n+1})_{1x_1}$ for all $n \geq 1$. Thus, by MCT,

$$\infty > \mu_2\left(\left(\bigcup_{n\geq 1} B_n\right)\right) = \mu_2\left(\bigcup_{n\geq 1} (B_n)_{1x_1}\right) = \lim_{n\to\infty} \mu_2((B_n)_{1x_1})$$

for all $x_1 \in \Omega_1$. This implies that $\mu_2\left(\left(\bigcup_{n\geq 1} B_n\right)\right)$ is $\langle \mathcal{F}_1, \mathcal{B}(\mathbb{R})\rangle$-measurable and hence, $bigcup_{n\geq 1} B_n \in \mathcal{S}$. Thus, $\mathcal{S}$ is a λ-system. Now, for $A = A_1 \times A_2 \in \mathcal{C}$, $\mu_2(A_{1x_1}) = \mu_2(A_2) I_{A_1}(x_1)$ and hence $\mathcal{C} \subset \mathcal{S}$. Since $\mathcal{C}$

is a π-system, it follows that $\mathcal{S} = \mathcal{F}_1 \times \mathcal{F}_2$. Thus, $\mu_2(A_{1x_1})$, considered as a function of x_1, is $\langle \mathcal{F}_1, \mathcal{B}(\mathcal{R})\rangle$-measurable for all $A \in \mathcal{F}_1 \times \mathcal{F}_2$. This proves 1.

Next, we prove part 2. By 1, μ_{12} is a well-defined set function on $\mathcal{F}_1 \times \mathcal{F}_2$. It is easy to check that μ_{12} is a well-defined measure on $\mathcal{F}_1 \times \mathcal{F}_2$. Similarly, μ_{21} is a well-defined measure on $\mathcal{F}_1 \times \mathcal{F}_2$. Since $\mu_{12}(A) = \mu_{21}(A) = \mu_1(A_1)\mu_2(A_2)$ for all $A = A_1 \times A_2 \in \mathcal{C}$ and $\mathcal{C}$ is a π-system generating $\mathcal{F}_1 \times \mathcal{F}_2$, it follows that $\mu_{12}(A) = \mu_{21}(A)$ for all $A \in \mathcal{F}_1 \times \mathcal{F}_2$. Thus, 2 is proved for finite $\mu_1(\Omega_1)$ and $\mu_2(\Omega_2)$.

Next let us assume that μ_i's are σ-finite. Then there exist disjoint sets $\{B_{in}\}_{n\geq 1}$ in $\mathcal{F}_i$ such that $\bigcup_{n\geq 1} B_{in} = \Omega_i$ and $\mu_i(B_{in}) < \infty$ for all $n \geq 1$, $i = 1, 2$. We define finite measures

$$\mu_{in}(D) = \mu_i(D \cap B_{in}), \;\; D \in \mathcal{F}_i,$$

for $n \geq 1, \; i = 1, 2$. The above arguments replacing μ_i by μ_{in} implies the $\langle \mathcal{F}_1, \mathcal{B}(\mathbb{R})\rangle$-measurability of $\mu_{2n}(A_{1x_1})$ for any $A \in \mathcal{F}_1 \times \mathcal{F}_2$, $n \geq 1$. Since μ_2 is a measure on $\mathcal{F}_2$,

$$\mu_2(A_{1x_1}) = \sum_{n=1}^{\infty} \mu_{2n}(A_{x_1})$$

and hence, considered as a function of x_1, it is $\langle \mathcal{F}_1, \mathcal{B}(\mathbb{R})\rangle$-measurable for all $A \in \mathcal{F}_1 \times \mathcal{F}_2$. Thus, the set function μ_{12} of (16.2.3) is well-defined and σ-finite as well. We can say the same for μ_{21} of (16.2.4). Now, let $\mu_{12}^{(m,n)}$ and $\mu_{21}^{(m,n)}$ denote the set functions defined by (16.2.3) and (16.2.4) respectively with μ_1 and μ_2 being replaced by μ_{1m} and μ_{2m}, for $m \geq 1, \; n \geq 1$. By repeated use of the MCT,

$$\begin{aligned}
\mu_{12}(A) &= \int_{\Omega_1} \mu_2(A_{1x_1})\mu_1 dx_1 \\
&= \sum_{m=1}^{\infty} \left(\int_{B_{1_m}} \sum_{n=1}^{\infty} \mu_2(A_{1x_1} \cap B_{2n}) \right) \mu_1 dx_1 \\
&= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \int_{B_{1m}} \mu_2(A_{1x_1} \cap B_{2n}) \mu_1 dx_1 \\
&= \sum_{m=1}^{\infty} \sum_{n=1}^{\infty} \mu_{12}^{(m,n)}(A), \quad A \in \mathcal{F}_1 \times \mathcal{F}_2 \qquad (16.2.5)
\end{aligned}$$

and similarly,

$$\mu_{21}(A) = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \mu_{21}^{(m,n)}(A), \quad A \in \mathcal{F}_1 \times \mathcal{F}_2. \qquad (16.2.6)$$

Since $\mu_{12}^{(m,n)}$ and $\mu_{21}^{(m,n)}$ are finite measures, it is easy to check that μ_{12} and μ_{21} are measures on $\mathcal{F}_1 \times \mathcal{F}_2$. Also, by the finite case,

$$\mu_{12}^{(m,n)}(A_1 \times A_2) = \mu_{21}^{(m,n)}(A_1 \times A_2) \quad \text{for all } n \geq 1, \; m \geq 1$$

and hence

$$\mu_{12}(A_1 \times A_2) = \mu_{21}(A_1 \times A_2) \quad \text{for all } A_1 \times A_2 \mathcal{C}.$$

Next $\{B_{1m} \times B_{2n} : \; m \geq 1, \; n \geq 1\}$ is a partition of $\Omega_1 \times \Omega_2$ by $\mathcal{F}_1 \times \mathcal{F}_2$ sets and by (16.2.5) and (16.2.6), for all $m \geq 1, n \geq 1$,

$$\mu_{12}(B_{1m} \times B_{2n}) = \mu_1(B_{1m})\mu_2(B_{2n}) = \mu_{21}(B_{1m} \times B_{2n}) < \infty.$$

Hence, μ_{12} and μ_{21} are σ-finite on $\mathcal{F}_1 \times \mathcal{F}_2$. Since μ_{12} and μ_{21} are equal on $\mathcal{C}$ and $\mathcal{C}$ is a π-system generating the product σ-algebra, it follows that $\mu_{12} = \mu_{21}$ on $\mathcal{F}_1 \times \mathcal{F}_2$ and it is the unique measure satisfying $\mu(A_1 \times A_2) = \mu_1(A_1)\mu_2(A_2)$ for all $A_1 \times A_2 \in \mathcal{C}$. This completes the proof. $\square$

Definition 16.2.6. The unique measure μ on $\mathcal{F}_1 \times \mathcal{F}_2$ in the above theorem is called the product measure and is denoted by $\mu_1 \times \mu_2$. The measure space $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2, \mu_1 \times \mu_2)$ is called the product measure space.

16.3 Fubini-Tonelli theorems

Let $f : \Omega_1 \times \Omega_2 \to \mathbb{R}$ be a $\langle \mathcal{F}_1 \times \mathcal{F}_2, \mathcal{B}(\mathbb{R})\rangle$-measurable function. Equations (16.2.3) and (16.2.4) suggest that the integral of f w.r.t. $\mu_1 \times \mu_2$ may be evaluated as iterated integrals, using the formulas

$$\int_{\Omega_1\times\Omega_2} f(x_1,x_2)\,\mu_1 \times \mu_2\,(d(x_1,x_2)) = \int_{\Omega_2}\left[\int_{\Omega_1} f(x_1,x_2)\,\mu_1(dx_1)\right]\mu_2(dx_2) \tag{16.3.1}$$

and

$$\int_{\Omega_1\times\Omega_2} f(x_1,x_2)\,\mu_1 \times \mu_2\,(d(x_1,x_2)) = \int_{\Omega_1}\left[\int_{\Omega_2} f(x_1,x_2)\,\mu_2(dx_2)\right]\mu_1(dx_1). \tag{16.3.2}$$

Here, the left sides of both (16.3.1) and (16.3.2) are simply the integral of f on the space $\Omega = \Omega_1 \times \Omega_2$ w.r.t. the measure $\mu = \mu_1 \times \mu_2$. The expressions on the right sides of (16.3.1) and (16.3.2) are, however, iterated integrals, where integrals of sections of f are evaluated first and then the resulting sectional integrals are integrated again to get the final expression. Conditions for the validity of (16.3.1) and (16.3.2) are provided by the Fubini-Tonelli theorems stated below.

Theorem 16.3.1. (Tonelli's theorem). Let $(\Omega_i, \mathcal{F}_i, \mu_i), i = 1, 2$ be σ-finite measure spaces and let $f : \Omega_1 \times \Omega_2 \to \mathbb{R}_+$be a nonnegative $\langle \mathcal{F}_1 \times \mathcal{F}_2\rangle$-measurable function. If $\overline{\mathbb{R}} = [-\infty, \infty]$, then

$$g_1(x_1) \equiv \int_{\Omega_2} f(x_1,x_2)\,\mu_2(dx_2) : \Omega_1 \to \overline{\mathbb{R}} \quad \text{is} \quad \left\langle \mathcal{F}_1, \mathcal{B}(\overline{\mathbb{R}})\right\rangle\text{-measurable} \tag{16.3.3}$$

and

$$g_2(x_2) \equiv \int_{\Omega_1} f(x_1,x_2)\,\mu_1(dx_1) : \Omega_2 \to \overline{\mathbb{R}} \quad \text{is } \left\langle \mathcal{F}_2, \mathcal{B}(\overline{\mathbb{R}})\right\rangle\text{-measurable.} \tag{16.3.4}$$

Further

$$\int_{\Omega_1\times\Omega_2} f d\mu = \int_{\Omega_1} g_1 d\mu_1 = \int_{\Omega_2} g_2 d\mu_2 \tag{16.3.5}$$

where $\mu = \mu_1 \times \mu_2$.

Proof. If $f = I_A$ for some A in $\mathcal{F}_1 \times \mathcal{F}_2$, the result follows from the previous theorem. By the linearity of integrals, the result now holds for all simple nonnegative functions f. For a general nonnegative function f, there exist a sequence $\{f_n\}_{n\geq 1}$ of nonnegative simple functions such that $f_n(x_1,x_2) \uparrow f(x_1,x_2)$ for all $(x_1,x_2) \in \Omega_1\times\Omega_2$. Write $g_{1n}(x_1) = \displaystyle\int_{\Omega_1} f_n(x_1,x_2)\,\mu_2(dx_2)$. Then, g_{1n} is $\mathcal{F}_1$-measurable for all $n \geq 1$, g_{1n} 's are nondecreasing, and by the MCT,

$$\begin{aligned} g_1(x_1) &\equiv \int_{\Omega_1} f(x_1,x_2)\,\mu_2(dx_2) \\ &= \lim_{n\to\infty} \int f_n(x_1,x_2)\,\mu_2(dx_2) \\ &= \lim_{n\to\infty} g_{1n}(x_1) \end{aligned} \tag{16.3.6}$$

for all $x_1 \in \Omega_1$. Thus, g_1 is $\langle \mathcal{F}_1, \mathcal{B}(\overline{\mathbb{R}}) \rangle$-measurable. Since (16.3.5) holds for simple functions, $\displaystyle\int f_n d\mu = \int g_{1n} d\mu_1$ for all $n \geq 1$. Hence, by repeated applications of the MCT, it follows that

$$\begin{aligned}
\int f d\mu &= \lim_{n\to\infty} \int f_n d\mu \\
&= \lim_{n\to\infty} \int g_{1n} d\mu_1 \\
&= \int \left(\lim_{n\to\infty} g_{1n} \right) d\mu_1 \\
&= \int g_1 d\mu_1.
\end{aligned}$$

The proofs of (16.3.4) and the second equality in (16.3.5) are similar. □

Theorem 16.3.2. (Fubini's theorem). Let $(\Omega_i, \mathcal{F}_i, \mu_i), i = 1, 2$ be σ-finite measure spaces and let $f \in L^1(\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2, \mu_1 \times \mu_2)$. Then there exist sets $B_i \in \mathcal{F}_i, i = 1, 2$ such that

1. $\mu_i(\Omega_i \backslash B_i) = 0$ for $i = 1, 2$,
2. for $x_1 \in B_1, f(x_1, \cdot) \in L^1(\Omega_2, \mathcal{F}_2, \mu_2)$, the function

$$g_1(x_1) = \begin{cases} \int_{\Omega_2} f(x_1, x_2) \mu_2 dx_2 & \text{for } x_1 \in B_1 \\ 0, \ \text{for } x_1 \in B_1^c \end{cases}$$

   is $\mathcal{F}_1$-measurable and

$$\int_{\Omega_1} g_1 d\mu_1 = \int_{\Omega_1 \times \Omega_2} f d(\mu_1 \times \mu_2), \tag{16.3.7}$$

3. for $x_2 \in B_2,\ f(\cdot, x_2) \in L^1(\Omega_1, \mathcal{F}_1, \mu_1)$, the function

$$g_2(x_2) = \begin{cases} \int_{\Omega_1} f(x_1, x_2) \mu_1 dx_1 & \text{for } x_2 \in B_2 \\ 0, \ \text{for } x_2 \in B_2^c \end{cases}$$

   is $\mathcal{F}_2$-measurable and

$$\int_{\Omega_2} g_2 d\mu_2 = \int_{\Omega_1 \times \Omega_2} f d(\mu_1 \times \mu_2). \tag{16.3.8}$$

Proof. By Tonelli's theorem

$$\int_{\Omega_1 \times \Omega_2} |f| d(\mu_1 \times \mu_2) = \int_{\Omega_1} \left(\int_{\Omega_2} |f(x_1, x_2)| \mu_2(dx_2) \right) \mu_1(dx_1).$$

So $\displaystyle\int_{\Omega_1 \times \Omega_2} |f| d(\mu_1 \times \mu_2) < \infty$ implies that $\mu_1(B_1^c) = 0$ where $B_1 = \{x_1 : \displaystyle\int |f(x_1, \cdot)| d\mu_2 < \infty\Big\}$. Also, by Tonelli's theorem

$$g_{11}(x_1) \equiv \int_{\Omega_2} f^+(x_1, \cdot) d\mu_2 \quad \text{and} \quad g_{12}(x_1) \equiv \int_{\Omega_2} f^-(x_1, \cdot) d\mu_2$$

are both $\mathcal{F}_1$-measurable and

$$\int_{\Omega_1} g_{11} d\mu_1 = \int_{\Omega_1 \times \Omega_2} f^+ d\left(\mu_1 \times \mu_2\right), \quad \int_{\Omega_1} g_{12} d\mu_1 = \int_{\Omega_1 \times \Omega_2} f^- d\left(\mu_1 \times \mu_2\right). \tag{16.3.9}$$

Since g_1 defined in 2 can be written as $g_1 = \left(g_{11} - g_{12}\right) I_{B_1}, g_1$ is $\mathcal{F}_1$ measurable. Also,

$$\begin{aligned}
\int_{\Omega_1} \left|g_1\right| d\mu_1 &\leq \int_{\Omega_1} g_{11} d\mu_1 + \int_{\Omega_1} g_{12} d\mu_1 \\
&= \int_{\Omega_1 \times \Omega_2} f^+ d\left(\mu_1 \times \mu_2\right) + \int_{\Omega_1 \times \Omega_2} f^- d\left(\mu_1 \times \mu_2\right) \\
&< \infty.
\end{aligned}$$

Further, as $\int_{\Omega_1 \times \Omega_2} |f| d\mu_1 \times \mu_2 < \infty$, by (16.3.9), g_{11} and $g_{12} \in L^1\left(\Omega_1, \mathcal{F}_1, \mu_1\right)$. Noting that $\mu_1\left(B_1^c\right) = 0$, one gets

$$\begin{aligned}
\int_{\Omega_1} g_1 d\mu_1 &= \int_{\Omega_1} \left(g_{11} - g_{12}\right) I_{B_1} d\mu_1 \\
&= \int_{\Omega_1} g_{11} I_{B_1} d\mu_1 - \int_{\Omega_1} g_{12} I_{B_1} d\mu_1 \\
&= \int_{\Omega_1} g_{11} d\mu_1 - \int_{\Omega_1} g_{12} d\mu_1
\end{aligned}$$

which, by (16.3.9), equals $\displaystyle\int_{\Omega_1 \times \Omega_2} f^+ d(\mu_1 \times \mu_2) - \int_{\Omega_1 \times \Omega_2} f^- d(\mu_1 \times \mu_2) = \int_{\Omega_1 \times \Omega_2} f d(\mu_1 \times \mu_2)$. Thus, 2 is established as well as 1 for $i = 1$. Similarly, we can prove 1 for $i = 2$ as well as 3. $\square$

Few Probable Questions

1. Define product measure. Show that it is unique.
2. State and prove Tonelli's theorem.
3. State and prove Fubini's theorem.

Unit 17

Course Structure

- Decomposition and Differentiations: Radon-Nikodym theorem

17.1 Introduction

This unit deals mainly with the Radon-Nikodym theorem and its implications. The Radon–Nikodym theorem essentially states that, under certain conditions, any measure ν can be expressed in this way with respect to another measure μ on the same space. The theorem is named after Johann Radon, who proved the theorem for the special case where the underlying space is $\mathbb{R}^n$ in 1913; and Otto Nikodym who proved the general case in 1930.

Objectives

After reading this unit, you will be able to

- define absolute continuity of measure
- define singular measure
- state Radon-Nikodym theorem and define Radon-Nikodym derivative along with its various properties

17.2 Differentiation

Let us start with the definition below.

Definition 17.2.1. Let $(\Omega, \mathcal{F})$ be a measurable space and let μ and ν be two measures on $(\Omega, \mathcal{F})$. The measure μ is said to be dominated by ν or absolutely continuous with respect to ν, written as $\mu \ll \nu$ if

$$\nu(A) = 0 \Rightarrow \mu(A) = 0 \text{ for all } A \in \mathcal{F}.$$

Example 17.2.2. Let f be a non-negative measurable function on a measure space $(\Omega, \mathcal{F}, \nu)$. Let

$$\mu(A) = \int_A f d\nu \text{ for all } A \in \mathcal{F}.$$

Then, μ is a measure on $(\Omega, \mathcal{F})$ and $\nu(A) = 0 \Rightarrow \mu(A) = 0$ for all $A \in \mathcal{F}$ and hence $\mu \ll \nu$.

The Radon-Nikodym theorem is a sort of converse to the above example. It says that if μ and ν are σ-finite measures on a measurable space $(\Omega, \mathcal{F})$ and if $\mu \ll \nu$, then there is a non-negative measurable function f on $(\Omega, \mathcal{F})$ such that

$$\mu(A) = \int_A f d\nu \text{ for all } A \in \mathcal{F}.$$

Definition 17.2.3. Let $(\Omega, \mathcal{F})$ be a measurable space and let μ and ν be two measures on $(\Omega, \mathcal{F})$. Then μ is called singular w.r.t. ν, written as $\mu \perp \nu$ if there exists a set $B \in \mathcal{F}$ such that

$$\mu(B) = 0 \text{ and } \nu(B^c) = 0.$$

It should be noted that μ singular with respect to ν implies hat ν is singular with respect to μ. Thus, the property of being singular is symmetric. However, the property of being absolutely continuous is not. It should also be noted that if $\mu \perp \nu$ and B is a set satisfying the singularity condition as given in the above definition, then for all $A \in \mathcal{F}$,

$$\mu(A) = \mu(A \cap B^c) \text{ and } \nu(A) = \nu(A \cap B). \tag{17.2.1}$$

Example 17.2.4. Let μ be the Lebesgue measure restricted to $(-\infty, 0]$, that is,

$$\mu(A) = \text{the Lebesgue measure of } A \cap (-\infty, 0];$$

and another measure ν is defined as follows

$$\nu(A) = \int_{A \cap (0,\infty)} \mathrm{e}^{-x}\, dx.$$

Then $\mu((0, \infty)) = 0$ and $\nu((-\infty, 0]) = 0$ and the singularity condition holds with $B = (-\infty, 0]$.

Suppose that μ and ν are are two finite measures on a measurable space $(\Omega, \mathcal{F})$. Then H. Lebesgue showed that μ can be decomposed as a sum of two measures, i.e.,

$$\mu = \mu_a + \mu_s$$

where $\mu_a \ll \nu$ and $\mu_s \perp \nu$.

Theorem 17.2.5. Let $(\Omega, \mathcal{F})$ be a measurable space and let μ_1 and μ_2 be two σ-finite measures on $(\Omega, \mathcal{F})$.

1. (The Lebesgue decomposition theorem). The measure μ_1 can be uniquely decomposed as

$$\mu_1 = \mu_{1a} + \mu_{1s} \tag{17.2.2}$$

where μ_{1a} and μ_{1s} are σ-finite measures on $(\Omega, \mathcal{F})$ such that $\mu_{1a} \ll \mu_{1s}$ and $\mu_{1s} \perp \mu_2$.

2. (The Radon-Nikodym theorem). There exists a non-negative measurable function h on $(\Omega, \mathcal{F})$ such that

$$\mu_{1a}(A) = \int_A h d\mu_2 \text{ for all } A \in \mathcal{F}. \tag{17.2.3}$$

Proof. Case 1: Suppose that μ_1 and μ_2 are finite measures. Let μ be the measure $\mu = \mu_1 + \mu_2$ and let $H = L^2$. Define a linear function T on H by

$$T(f) = \int f d\mu_1. \tag{17.2.4}$$

Then, by the Cauchy-Schwarz inequality applied to the functions f and $g \equiv 1$,

$$\begin{aligned} |T(f)| &\leq \left(\int f^2 d\mu_1\right)^{\frac{1}{2}} (\mu_1(\Omega))^{\frac{1}{2}} \\ &\leq \left(\int f^2 d\mu\right)^{\frac{1}{2}} (\mu_1(\Omega))^{\frac{1}{2}}. \end{aligned}$$

This shows that T is a bounded linear functional on H with $\|T\| \leq M \equiv (\mu_1(\Omega))^{\frac{1}{2}}$. By the Riesz representation theorem, there exists a $g \in L^2$ such that

$$T(f) = \int fg d\mu \tag{17.2.5}$$

for all $f \in L^2$. Let $f = I_A$ for $A \in \mathcal{F}$. Then equations (17.2.4) and (17.2.5) give

$$\mu_1(A) = T(I_A) = \int_A g d\mu.$$

But, $0 \leq \mu_1(A) \leq \mu(A)$ for all $A \in \mathcal{F}$. Hence the function g in L^2 satisfies

$$0 \leq \int_A g d\mu \leq \mu(A) \text{ for all } A \in \mathcal{F}. \tag{17.2.6}$$

Let $A_1 = \{0 \leq g < 1\}$, $A_2 = \{g = 1\}$, $A_3 = \{g \notin [0,1]\}$. Then equation (17.2.6) implies that $\mu(A_3) = 0$. Now we define measures μ_{1a} and μ_{1s} as follows.

$$\mu_{1a}(A) = \mu_1(A \cap A_1), \;\; \mu_{1s}(A) = \mu_1(A \cap A_2), \;\; A \in \mathcal{F}. \tag{17.2.7}$$

Next we show that $\mu_{1a} \ll \mu_2$ and $\mu_{1s} \perp \mu_2$. By equations (17.2.4) and (17.2.5), for all $f \in H$,

$$\begin{aligned} \int f d\mu_1 = \int g d\mu &= \int fg d\mu_1 + \int fg d\mu_2 \\ \Rightarrow \int f(1-g) d\mu_1 &= \int fg d\mu_2. \end{aligned} \tag{17.2.8}$$

Setting $f = I_{A_2}$ yields

$$0 = \mu_2(A_2).$$

From equation (17.2.7), since $\mu_{1s}(A_2^c) = 0$, it follows that $\mu_{1s} \perp \mu_2$. Now, fix $n \geq 1$ and $A \in \mathcal{F}$. Let $f = I_{A \cap A_1}(1 + g + \ldots + g^{n-1})$. Then (17.2.8) implies that

$$\int_{A \cap A_1} (1 - g^n) d\mu_1 = \int_{A \cap A_1} g(1 + g + \ldots + g^{n-1}) d\mu_2.$$

Now, letting $n \to \infty$ and using the MCT on both sides yield

$$\mu_{1a}(A) = \int_A I_{A_1} \frac{g}{1-g} d\mu_2. \tag{17.2.9}$$

Setting $h \equiv \dfrac{g}{1-g} I_{A_1}$ completes the proof of (17.2.2) and (17.2.3).

Case 2: Suppose that μ_1 and μ_2 are σ-finite measures. Then there exists a countable partition $\{D_n\} \subset \mathcal{F}$ of Ω such that $\mu_1(D_n)$ and $\mu_2(D_n)$ are both finite for all $n \geq 1$. Let

$$\mu_1^{(n)}(\cdot) = \mu_1(\cdot \cap D_n) \text{ and } \mu_2^{(n)}(\cdot) = \mu_2(\cdot \cap D_n).$$

Then applying Case 1 to $\mu_1^{(n)}$ and $\mu_2^{(n)}$ for each $n \geq 1$, one gets measures $\mu_{1a}^{(n)}$, $\mu_{1s}^{(n)}$ and a function h_n such that

$$\mu_1^{(n)}(\cdot) = \mu_{1a}^{(n)}(\cdot) + \mu_{1s}^{(n)}(\cdot), \tag{17.2.10}$$

where, for $A \in \mathcal{F}$,

$$\mu_{1a}^{(n)}(A) = \int_A h_n d\mu_2^{(n)} = \int_A h_n I_{D_n} d\mu_2$$

and $\mu_{1s}^{(n)} \perp \mu_2^{(n)}$. Since $\mu_1(\cdot) = \sum_{n=1}^{\infty} \mu_1^{(n)}(\cdot)$, it follows from (17.2.10) that

$$\mu_1(\cdot) = \mu_{1a}(\cdot) + \mu_{1s}(\cdot) \tag{17.2.11}$$

where $\mu_{1a}(A) = \sum_{n=1}^{\infty} \mu_{1a}^{(n)}(A)$ and $\mu_{1s}(\cdot) = \sum_{n=1}^{\infty} \mu_{1s}^{(n)}(\cdot)$. By MCT,

$$\mu_{1a}(A) = \int_A h d\mu_2, \quad A \in \mathcal{F},$$

where $h = \sum_{n=1}^{\infty} h_n I_{D_n}$.

Clearly, $\mu_{1a} \ll \mu_2$. The verification of the singularity of μ_{1s} and μ_2 is left as an exercise.

It remains to prove the uniqueness of the decomposition. Let

$$\mu_1 = \mu_a + \mu_s = \mu_a' + \mu_s'$$

be two decompositions of μ_1 where μ_a and μ_a' are absolutely continuous with respect to μ_2 and μ_s and μ_s' are singular with respect to μ_2. By definition, there exist sets B and B' in $\mathcal{F}$ such that

$$\mu_2(B) = 0, \ \mu_2(B') = 0, \text{ and } \mu_s(B^c) = 0, \ \mu_s'(B'^c) = 0.$$

Let $D = B \cup B'$. Then $\mu_2(D) = 0$ and $\mu_s(D^c) \leq \mu_s(B^c) = 0$. Similarly, $\mu_s'(D^c) \leq \mu_s'(B'^c) = 0$. Also, $\mu_2(D) = 0$ implies $\mu_a(D) = 0 = \mu_a'(D)$. Thus, for any $A \in \mathcal{F}$,

$$\mu_a(A) = \mu_a(A \cap D^c) \text{ and } \mu_a'(A) = \mu_a'(A \cap D^c).$$

Also

$$\begin{aligned} \mu_s(A \cap D^c) &\leq \mu_s(A \cap B^c) = 0. \\ \mu_s'(A \cap D^c) &\leq \mu_s'(A \cap B'^c) = 0. \end{aligned}$$

Thus, $\mu(A \cap D^c) = \mu_a(A \cap D^c) + \mu_s(A \cap D^c) = \mu_a(A)$ and $\mu(A \cap D^c) = \mu_a'(A \cap D^c) + \mu_s'(A \cap D^c) = \mu_a'(A \cap D^c) = \mu_a'(A)$. Hence, $\mu_a(A) = \mu(A \cap D^c) = \mu_a'(A)$ for every $A \in \mathcal{F}$. That is, $\mu_a = \mu_a'$ and hence $\mu_s = \mu_s'$.

□

The σ-finiteness in the above theorem can't be dropped. For example, let μ be the Lebesgue measure and ν be the counting measure on $[0,1]$. Then $\mu \ll \nu$ but there does not exist a non-negative $\mathcal{F}$-measurable function h such that $\mu(A) = \int_A h d\nu$ (Check!).

Definition 17.2.6. Let μ and ν be measures on a measurable space $(\Omega, \mathcal{F})$ and let h be a non-negative measurable function such that

$$\mu(A) = \int_A h d\nu \text{ for all } A \in \mathcal{F}.$$

Then h is called the Radon-Nikodym derivative of μ w.r.t. ν and is written as

$$\frac{d\mu}{d\nu} = h.$$

If $\mu(\Omega) < \infty$, and there exist two non-negative $\mathcal{F}$-measurable functions h_1 and h_2 such that

$$\mu(A) = \int_A h_1 d\nu = \int_A h_2 d\nu$$

for all $A \in \mathcal{F}$, then $h_1 = h_2$ a.e. (ν) and thus the Radon-Nikodym derivative is unique upto equivalence a.e. (ν). This also extends to the case when μ is σ-finite.

Theorem 17.2.7. Let $\nu, \mu, \mu_1, \mu_2, \ldots$ be σ-finite measures on a measurable space $(\Omega, \mathcal{F})$.

1. If $\mu_1 \ll \mu_2$ and $\mu_2 \ll \mu_3$, then $\mu_1 \ll \mu_3$ and
$$\frac{d\mu_1}{d\mu_3} = \frac{d\mu_1}{d\mu_2}\frac{d\mu_2}{d\mu_3} \;\; a.e.\ (\mu_3).$$

2. Suppose that μ_1 and μ_2 are dominated by μ_3. Then for any $\alpha, \beta \geq 0$, $\alpha\mu_1 + \beta\mu_2$ is dominated by μ_3 and
$$\frac{d(\alpha\mu_1 + \beta\mu_2)}{d\mu_3} = \alpha\frac{d\mu_1}{d\mu_3} + \beta\frac{d\mu_2}{d\mu_3} \;\; a.e.\ (\mu_3).$$

3. If $\mu \ll \nu$ and $\frac{d\mu}{d\nu} > 0$ a.e. (ν), then $\nu \ll \mu$ and
$$\frac{d\nu}{d\mu} = \left(\frac{d\mu}{d\nu}\right)^{-1} \;\; a.e.\ (\mu).$$

4. Let $\{\mu_n\}$ be a sequence of measures and $\{\alpha_n\}$ be a sequence of positive real numbers, that is, $\alpha_n > 0$ for all $n \geq 1$. Define $\mu = \sum_{n=1}^{\infty} \alpha_n \mu_n$.

   (a) Then, $\mu \ll \nu$ iff $\mu_n \ll \nu$ for each $n \geq 1$ and in this case,
   $$\frac{d\mu}{d\nu} = \sum_{n=1}^{\infty} \alpha_n \frac{d\mu_n}{d\nu} \;\; a.e.\ (\nu).$$

   (b) $\mu \perp \nu$ iff $\mu_n \perp \nu$ for all $n \geq 1$.

Few Probable Questions

1. State and prove he Lebesgue decomposition theorem.

2. State and prove the Radon-Nikodym theorem.

Unit 18

Course Structure

- Signed and Complex measures

18.1 Introduction

A signed measure on a measurable space is a set function which has all the properties of a measure, except that of non-negativity. There are two slightly different concepts of a signed measure, depending on whether or not one allows it to take infinite values. Signed measures are usually only allowed to take finite real values, while some textbooks allow them to take infinite values. For example, if μ and ν are measures on $(\Omega, \mathcal{F})$, then $\lambda = \alpha\mu + \beta\nu$ where $\alpha, \beta \geq 0$ is a measure on $(\Omega, \mathcal{F})$. However, if $\alpha = 1,\ \beta = -1$, then λ may tale both positive as well as negative values. However, the situation can be serious when $\mu(E) = +\infty = \nu(E)$. This situation allows us to introduce the concept of signed measures.

Objectives

After reading this unit, you will be able to

- grasp the idea of signed measures and show that every finite signed measure can be expressed as the difference of two finite measures;
- define positive and negative sets with respect to a finite signed measure;
- state and prove Hahn decomposition theorem;
- define complex measure.

18.2 Signed Measures

Let μ_1 and μ_2 be two finite measures on a measurable space $(\Omega, \mathcal{F})$. Let

$$\nu(A) = \mu_1(A) - \mu_2(A), \quad \text{for all } A \in \mathcal{F}. \tag{18.2.1}$$

Then $\nu : \mathcal{F} \to \mathbb{R}^* = [-\infty, +\infty]$ satisfies the following:

1. $\nu(\emptyset) = 0$;

2. For any $\{A_n\} \subset \mathcal{F}$ with $A_i \cap A_j = \emptyset$ for $i \neq j$, and $\sum_{i=1}^{\infty} |\nu(A_i)| < \infty$

$$\nu(A) = \sum_{i=1}^{\infty} \nu(A_i). \tag{18.2.2}$$

3. Let

$$\|\nu\| = \sup \left\{ \sum_{i=1}^{\infty} |\nu(A_i)| : \{A_n\} \subset \mathcal{F}, \ A_i \cap A_j = \emptyset \text{ for } i \neq j, \ \bigcup_{n \geq 1} A_n = \Omega \right\}. \tag{18.2.3}$$

   Then $\|\nu\|$ is finite.

Note that 3 holds because $\|\nu\| \leq \mu_1(\Omega) + \mu_2(\Omega) < \infty$.

Definition 18.2.1. A set function $\nu : \mathcal{F} \to \mathbb{R}^*$ satisfying 1, 2 and 3 above is called a finite signed measure.

It will be shown below that every finite signed measure can be expressed as the difference of two finite measures.

Theorem 18.2.2. Let ν be a finite signed measure on $(\Omega, \mathcal{F})$. Let

$$|\nu|(A) = \sup \left\{ \sum_{i=1}^{\infty} |\nu(A_i)| : \{A_n\} \subset \mathcal{F}, \ A_i \cap A_j = \emptyset \text{ for } i \neq j, \ \bigcup_{n \geq 1} A_n = A \right\}. \tag{18.2.4}$$

Then $|\nu|$ is a finite measure on $(\Omega, \mathcal{F})$.

Proof. From 3 of the definition, it follows that $|\nu(\Omega)| < \infty$. Thus it is enough to verify that $|\nu|$ is countably additive. Let $\{A_n\}$ be a countable family of disjoint sets in $\mathcal{F}$. Let $A = \bigcup_{n \geq 1} A_n$. By the definition of $|\nu|$, for all $\epsilon > 0$ and $n \in \mathbb{N}$, there exists a countable family $\{A_{n_j}\}$ of disjoint sets in $\mathcal{F}$ with $A_n = \bigcup_{j \geq 1} A_{n_j}$ such that $\sum_{j=1}^{\infty} |\nu(A_{n_j})| > |\nu|(A_n) - \frac{\epsilon}{2^n}$. Hence,

$$\sum_{n=1}^{\infty} \sum_{j=1}^{\infty} |\nu(A_{n_j})| > \sum_{n=1}^{\infty} |\nu|(A_n) - \epsilon.$$

Note that $\{A_{n_j}\}$ is a countable family of disjoint sets in $\mathcal{F}$ such that $A = \bigcup_{n \geq 1} A_n = \bigcup_{n \geq} \bigcup_{j \geq 1} A_{n_j}$. It follows from the definition of $|\nu|$ that

$$|\nu|(A) \geq \sum_{n=1}^{\infty} \sum_{j=1}^{\infty} |\nu(A_{n_j})| > \sum_{n=1}^{\infty} |\nu|(A_n) - \epsilon.$$

Since this is true for for all $\epsilon > 0$, it follows that

$$|\nu|(A) \geq \sum_{n=1}^{\infty} |\nu|(A_n). \tag{18.2.5}$$

To get the opposite inequality, let $\{B_j\}$ be a countable family of disjoint sets in $\mathcal{F}$ such that $\bigcup_{j\geq 1} B_j = A = \bigcup_{n\geq 1} A_n$. Since $B_j = B_j \cap A = \bigcup_{j\geq 1}(B_j \cap A_n)$ and ν satisfies (18.2.2)

$$\nu(B_j) = \sum_{n=1}^{\infty} \nu(B_j \cap A_n) \text{ for all } j \geq 1.$$

Thus,

$$\begin{aligned}\sum_{j=1}^{\infty} |\nu(B_j)| &\leq \sum_{j=1}^{\infty}\sum_{n=1}^{\infty} |\nu(B_j \cap A_n)| \\ &= \sum_{n=1}^{\infty}\sum_{j=1}^{\infty} |\nu(B_j \cap A_n)|. \end{aligned} \tag{18.2.6}$$

Note that for each A_n, $\{B_j \cap A_n\}_{j\geq 1}$ is a countable family of disjoint sets in $\mathcal{F}$ such that $A_n = \bigcup_{j\geq 1}(B_j \cap A_n)$. Hence, from equation (18.2.4), it follows that $|\nu|(A_n) \geq \sum_{j=1}^{\infty} |\nu(B_j \cap A_n)|$ and hence, $\sum_{n=1}^{\infty} |\nu|(A_n) \geq \sum_{n=1}^{\infty}\sum_{j=1}^{\infty}$ $|\nu(B_j \cap A_n)|$. From (18.2.6), it follows that $\sum_{n=1}^{\infty} |\nu|(A_n) \geq \sum_{j=1}^{\infty} |\nu(B_j)|$. This being true for every such family $\{B_j\}$, it follows that from (18.2.4) that

$$|\nu|(A_n) \leq \sum_{i=1}^{\infty} |\nu|(A_i) \tag{18.2.7}$$

and with (18.2.5), this completes the proof. □

Definition 18.2.3. The measure $|\nu|$ defined by (18.2.4) is called the total variation measure or absolute measure of the signed measure ν.

Next, define the set functions

$$\nu^+ = \frac{|\nu| + \nu}{2}, \quad \nu^- = \frac{|\nu| - \nu}{2}. \tag{18.2.8}$$

It can also be verified that both ν^+ and ν^- are finite measures on $(\Omega, \mathcal{F})$.

Definition 18.2.4. The measures ν^+ and ν^- are called the positive and negative variation measures of the signed measure ν, respectively.

It follows from (18.2.8) that

$$\nu = \nu^+ - \nu^-. \tag{18.2.9}$$

Thus every finite signed measure ν on $(\Omega, \mathcal{F})$ s the difference of two finite measures, as claimed earlier.

Note that both ν^+ and ν^- are dominated by $|\nu|$ and all three measures are finite. By the Radon-Nikodym theorem, there exist functions h_1 and h_2 in $L^1(\Omega, \mathcal{F}, |\nu|)$ such that

$$\frac{d\nu^+}{d|\nu|} = h_1, \text{ and } \frac{d\nu^-}{d|\nu|} = h_2. \tag{18.2.10}$$

This and (18.2.9) imply that for any $A \in \mathcal{F}$,

$$\nu(A) = \int_A h_1 d|\nu| - \int_A h_2 d|\nu| = \int_A h d|\nu|, \tag{18.2.11}$$

where $h = h_1 - h_2$. Thus every finite signed measure ν on $(\Omega, \mathcal{F})$ can be expressed as

$$\nu(A) = \int_A f d\mu, \quad A \in \mathcal{F} \tag{18.2.12}$$

for some finite measure μ on $(\Omega, \mathcal{F})$ and some $f \in L^1(\Omega, \mathcal{F}, \mu)$.

Conversely, it is easy to verify that a set function ν defined in (18.2.12) for some finite measure μ on $(\Omega, \mathcal{F})$ and some $f \in L^1(\Omega, \mathcal{F}, \mu)$ is a finite signed measure. This leads to the following result.

Theorem 18.2.5.

1. A set function ν on a measurable space $(\Omega, \mathcal{F})$ is a finite signed measure iff there exist two finite measures μ_1 and μ_2 on $(\Omega, \mathcal{F})$ such that $\nu = \mu_1 - \mu_2$.

2. A set function ν on a measurable space $(\Omega, \mathcal{F})$ is a finite signed measure if there exists a finite measure μ on $(\Omega, \mathcal{F})$ and some $f \in L^1(\Omega, \mathcal{F}, \mu)$ such that for all $A \in \mathcal{F}$,

$$\nu(A) = \int_A f d\mu.$$

Definition 18.2.6. Let ν be a finite signed measure on a measurable space on $(\Omega, \mathcal{F})$. A set $A \in \mathcal{F}$ is called a positive set for ν if for any $B \subset A$, $B \in \mathcal{F}$, $\nu(B) \geq 0$. A set $A \in \mathcal{F}$ is called a negative set for ν if for any $B \subset A$, $B \in \mathcal{F}$, $\nu(B) \leq 0$. A is a null set if it is both positive as well as negative with respect to ν.

Remark 18.2.7. If ν is a finite signed measure on $(\Omega, \mathcal{F})$, then $-\nu$ is also so. We have the following.

1. If a set is positive with respect to ν, then it is negative with respect to $-\nu$;

2. If a set is negative with respect to ν, then it is positive with respect to $-\nu$;

3. If a set is null with respect to ν, then it is so with respect to $-\nu$;

4. for $\alpha \in \mathbb{R}$, $\alpha\nu$ is a finite signed measure.

Exercise 18.2.8.

1. Show that the countable union of sets, positive with respect to ν, is also a positive set.

2. Show that the countable union of sets, negative with respect to ν, is also a negative set.

3. Show that the countable union of null sets, is also a null set.

Theorem 18.2.9. Let $(\Omega, \mathcal{F})$ be a measurable space and $E \in \mathcal{F}$ with $0 < \nu(E) < +\infty$. Then there exists a positive set $A \subset E$ such that $\nu(A) > 0$.

Proof. If E is a positive set with respect to ν, then E does not contain any negative set and in that case, we have $A = E$, which is our desired set.

Otherwise, let E contains a set negative with respect to ν-measure. Let n_1 be the least positive integer such that there is a set $E_1 \subset E$ with $\nu(E_1) < -\dfrac{1}{n_1}$. If $E \setminus E_1$ is not a positive set, then let n_2 be the least positive integer so that there is a set $E_2 \subset E \setminus E_1$ with $\nu(E_2) < -\dfrac{1}{n_2}$. Continuing this process inductively, we can find a least positive integer n_k so that there is a set $E_k \subset E \setminus \bigcup_{i=1}^{k-1} E_i$ such that $\nu(E_k) < -\dfrac{1}{n_k}$. If this process does not stop, we take $A = E \setminus \bigcup_{i=1}^{\infty} E_i$. Then $E = A \cup \left(\bigcup_{i=1}^{\infty} E_i \right)$. The sets A and $\{E_i\}$ are disjoint. Hence we have

$$\nu(E) = \nu(A) + \sum_{i=1}^{\infty} \nu(E_i).$$

Since $0 < \nu(E) < +\infty$, the series $\sum_{i=1}^{\infty} \nu(E_i)$ converges absolutely. Hence, $\sum_{i=1}^{\infty} \nu(E_i) > -\infty$. Hence,

$$-\infty < \sum_{i=1}^{\infty} \nu(E_i) < -\sum_{i=1}^{\infty} \frac{1}{n_i}$$

$$\Rightarrow \quad \sum_{i=1}^{\infty} \frac{1}{n_i} < +\infty.$$

In particular, $\lim_{i \to \infty} n_i = +\infty$ and $n_i > 1$ for $i \geq i_0$. We show that A is a positive set with respect to ν. Given $\epsilon > 0$, we have $(n_i - 1)^{-1} < \epsilon$ for large value of i. As $A = E \setminus \bigcup_{i=1}^{\infty} E_i$, then A has no subset with ν-measure less that $-(n_i - 1)^{-1}$ which is greater than $-\epsilon$. Since ϵ is arbitrary, then A has no subset with negative ν-measure. Hence, A is a positive set with respect to ν. $\square$

Theorem 18.2.10. (Hahn Decomposition theorem). Let ν be a signed measure on a measurable space $(\Omega, \mathcal{F})$. Then there is a positive set A and a negative set B such that $\Omega = A \cup B$ and $A \cap B = \emptyset$.

Proof. We note that ν can not take both values $+\infty$ and $-\infty$. Without any loss of generality, we may assume that $+\infty$ is the infinite value omitted by ν, that is, $\nu(E) < +\infty$ for all $E \in \mathcal{F}$ (otherwise we take $-\nu$, the result for $-\nu$ implying the result for ν). Let $\lambda = \sup\{\nu(A) : A \text{ is a positive set with respect to } \nu\}$. Since empty set is positive, hence $\lambda \geq 0$. Then there is a sequence $\{A_n\}$ of positive sets such that $\lambda = \lim_{n \to \infty} \nu(A_n)$. Write $A = \bigcup_{n=1}^{\infty} A_n$. Since each A_n is a positive set and a countable union of positive sets is positive, hence A is a positive set. Then

$$\nu(A) \leq \lambda. \tag{18.2.13}$$

Since $A \setminus A_n \subset A$ for each n, hence $\nu(A \setminus A_n) \geq 0$ for each n. Now as $A = (A \setminus A_n) \cup A_n$, hence

$$\nu(A) = \nu(A \setminus A_n) + \nu(A_n) \geq \nu(A_n) \;\; \forall n.$$

Hence

$$\nu(A) \geq \lambda. \tag{18.2.14}$$

Combining (18.2.13) and (18.2.14), $\nu(A) = \lambda$, $0 \leq \lambda < \infty$. Thus, we see that the value of λ is attained by positive set, namely A.

Now let $B = A^c$. Suppose E is a positive subset of B. Since E and A are both positive sets, $E \cup A$ is also positive, then

$$\lambda \geq \nu(E \cup A) = \nu(E) + \nu(A) = \nu(E) + \lambda.$$

This implies that $\nu(E) = 0$ since $0 \leq \lambda < \infty$. Hence, B contains no positive subsets of positive ν-measure. Hence B is a negative set with the desired property. $\square$

Definition 18.2.11. Let ν be a signed measure on the measurable space $(\Omega, \mathcal{F})$. A decomposition of Ω into two disjoint sets A and B such that A is positive and B is negative with respect to ν is called the Hahn decomposition of Ω with respect to ν. We use $\{A, B\}$ as a notation for Hahn decomposition of Ω. We sometimes use the notation for A^+ and A^- respectively in the Hahn decomposition with respect to ν.

Note 18.2.12. Hahn Decomposition need not be unique. Let $\Omega = \{a, b, c\}$ and $\mathcal{F} = \mathcal{P}(\Omega)$, the power set of Ω and $\nu = \delta_a - \delta_b$ where δ_a and δ_b are defined as follows.

$$\delta_a(E) = \begin{cases} 1, & a \in E \\ 0, & a \notin E \end{cases} \quad , \quad \delta_b(E) = \begin{cases} 1, & b \in E \\ 0, & b \notin E \end{cases}$$

Consider $A = \{a\}$ and $B = \{b, c\}$. Then $A \cap B = \emptyset$ and $A \cup B = \Omega$. Also,

$$\nu(A) = \delta_a(A) - \delta_b(A) = 1 - 0 = 1.$$

Then $\nu(A) \geq 0$. Also

$$\nu(B) = \delta_a(B) - \delta_b(B) = -1.$$

$\nu(\{b\}) = -1$ and $\nu(\{c\}) = 0$. Thus, $\nu(B) \leq 0$.

Again, if we take $A_1 = \{a, c\}$ and $B_1 = \{b\}$ then $\Omega = A_1 \cup B_1$ and $A_1 \cap B_1 = \emptyset$. Then A_1 is a positive set and B_1 is negative. Hence both $\{A, B\}$ and $\{A_1, B_1\}$ are Hahn decompositions of Ω with respect to ν.

18.3 Complex Measures

Complex measures are defined analogously to signed measures, except that they are only permitted to take finite complex values.

Definition 18.3.1. Let $(\Omega, \mathcal{F})$ be a measurable space. A complex measure ν on Ω is a function $\nu : \mathcal{F} \to \mathbb{C}$ satisfying the following.

1. $\nu(\emptyset) = 0$;
2. If $\{A_n\}$ is a disjoint collection of measurable sets, then

$$\nu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \nu(A_n).$$

There is an analogous Radon-Nikodym theorems for complex measures.

Theorem 18.3.2. (Lebesgue-Radon-Nikodym theorem). Let ν be a complex measure and μ be a σ-finite measure on a measurable space $(\Omega, \mathcal{F})$. Then there exist unique complex measures ν_a, ν_s such that

$$\nu = \nu_a + \nu_s, \ \text{ where } \ \nu_a \ll \mu \ \text{ and } \ \nu_s \perp \mu.$$

Moreover, there exists an integrable function $f : \Omega \to \mathbb{C}$, uniquely defined up to μ a.e. equivalence, such that

$$\nu_a(A) = \int_A f d\mu$$

for every $A \in \mathcal{F}$.

Few Probable Questions

1. Define absolute measure of a signed measure ν. Show that it satisfies all the properties of a measure.
2. State and prove Hahn decomposition theorem.
3. Define Hahn decomposition of a set Ω. Is it unique? Justify your answer.
4. Show that a set function ν on a measurable space $(\Omega, \mathcal{F})$ is a finite signed measure iff there exist two finite measures μ_1 and μ_2 on $(\Omega, \mathcal{F})$ such that $\nu = \mu_1 - \mu_2$.

Unit 19

Course Structure

- Differentiation on absolute Continuity, Lebesgue differentiation Theorem

19.1 Absolutely continuous functions on $\mathbb{R}$

Definition 19.1.1. A function $f : \mathbb{R} \to \mathbb{R}$ is absolutely continuous (a.c.) if for all $\epsilon > 0$, there exists $\delta > 0$ such that if $I_j = [a_j, b_j]$, $j = 1, 2, \ldots, k$ ($k \in \mathbb{N}$) are disjoint and

$$\sum_{j=1}^{n}(b_j - a_j) < \delta, \text{ then } \sum_{j=1}^{n}|f(b_j) - f(a_j)| < \epsilon.$$

By the mean value theorem, it follows that if f is differentiable and f' is bounded, then f is a.c. Also, f is a.c. implies it is uniformly continuous.

Definition 19.1.2. A function $f : [a, b] \to \mathbb{R}$ is absolutely continuous if the function F, defined by

$$F(x) = \begin{cases} f(x), & \text{if } a \le x \le b \\ f(a), & \text{if } x < a \\ f(a), & \text{if } x > b \end{cases},$$

is absolutely continuous.

Example 19.1.3. The functions $f(x) = x$ is a.c. on $\mathbb{R}$. Any polynomial is a.c. on any bounded interval but not necessarily on all of $\mathbb{R}$. For example, $f(x) = x^2$ is a.c. on any bounded interval but not a.c. on $\mathbb{R}$, since it is not uniformly continuous on $\mathbb{R}$.

The following theorem is known as the fundamental theorem of Lebesgue integral calculus.

Theorem 19.1.4. A function $f : [a, b] \to \mathbb{R}$ is absolutely continuous iff there is a function $F : [a, b] \to \mathbb{R}$ such that F is Lebesgue measurable and integrable w.r.t. m and such that

$$f(x) = f(a) + \int_{[a,x]} F dm \tag{19.1.1}$$

for all $a \le x \le b$, where m is the Lebesgue measure.

Proof. First let (19.1.1) holds. Since $\int_{[a,b]} |F|dm < \infty$, for any $\epsilon > 0$, there exists a $\delta > 0$ such that

$$m(A) < \delta \Rightarrow \int_A |F|dm < \epsilon. \tag{19.1.2}$$

Thus, if $I_j = (a_j, b_j) \subset [a, b]$, $j = 1, 2, \dots, k$ are such that $\sum_{j=1}^{k}(b_j - a_j) < \delta$, then

$$\sum_{j=1}^{k} |f(b_j) - f(a_j)| \le \int_{\cup_{j=1}^{k} I_j} |F|dm < \epsilon,$$

since

$$m\left(\bigcup_{j=1}^{k} I_j\right) \le \sum_{j=1}^{k}(b_j - a_j) < \delta$$

and (19.1.2) holds. Thus, f is a.c.

Converse part is left as exercise. □

The expression (19.1.1) of an absolutely continuous f can be strengthened as follows:

Theorem 19.1.5. Let $f : \mathbb{R} \to \mathbb{R}$ satisfy (19.1.1). Then f is differentiable a.e. (m) and $f' = F$ a.e. (m).

Now, we recall that a measure μ on $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ is a Radon measure if $\mu(A) < \infty$ for every bounded Borel set A. In the following, we define the differential of a Radon measure on $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$.

Definition 19.1.6. A measure μ on $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ is differentiable at $x \in \mathbb{R}^k$ with derivative $\mu'(x)$ if for any $\epsilon > 0$, there is a $\delta > 0$ such that

$$\left|\frac{\mu(A)}{m(A)} - \mu'(x)\right| < \epsilon$$

for every open ball A such that $x \in A$ and $\text{diam}(A) < \delta$ ($\text{diam}(A) = \sup\{\|x - y\| : x, y \in A\}$.

Theorem 19.1.7. Let μ be a Radon measure on $(\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$. Then

1. μ is differentiable a.e. (m), μ' is Lebesgue measurable and greater equal to 0 a.e. (m) and for all bounded Borel sets $A \in \mathcal{B}(\mathbb{R}^k$,
$$\int_A \mu' dm \le \mu(A).$$

2. Let $\mu_a(A) = \int_A \mu' dm$, $A \in \mathcal{B}(\mathbb{R}^k$. Let μ_s be the unique measure on $\mathcal{B}(\mathbb{R}^k$ such that for all bounded Borel sets A,
$$\mu_s(A) = \mu(A) - \mu_a(A).$$
Then
$$\mu_s \perp m \quad \text{and} \quad \mu' = 0 \ \text{ a.e. } (m).$$

Unit 20

Course Structure

- Functions of Bounded variations, Riesz representation Theorem.

20.1 Functions of Bounded variations

Definition 20.1.1. Let $f : [a, b] \to \mathbb{R}$, where $-\infty < a < b < \infty$. Then for any partition $Q = \{a = x_0, x_1, \ldots, x_n = b\}$, where $x_0 < x_1 < \ldots < x_n$ for $n \in \mathbb{N}$, the positive, negative and total variations of f with respect to Q are respectively defined as

$$P(f,Q) \equiv \sum_{i=1}^{n} \left(f\left(x_i\right) - f\left(x_{i-1}\right)\right)^{+}$$
$$N(f,Q) \equiv \sum_{i=1}^{n} \left(f\left(x_i\right) - f\left(x_{i-1}\right)\right)^{-}$$
$$T(f,Q) \equiv \sum_{i=1}^{n} \left|f\left(x_i\right) - f\left(x_{i-1}\right)\right|.$$

It is easy to verify that

(i) if f is non-decreasing, then

$$P(f,Q) = T(f,Q) = f(b) - f(a) \quad \text{and} \quad N(f,Q) = 0$$

(ii) for any f,

$$P(f,Q) + N(f,Q) = T(f,Q).$$

Definition 20.1.2. Let $f = [a, b] \to \mathbb{R}$, where $-\infty < a < b < \infty$. The positive, negative and total variations of f over $[a, b]$ are respectively defined as

$$P(f,[a,b]) \equiv \sup_{Q} P(f,Q)$$
$$N(f,[a,b]) \equiv \sup_{Q} N(f,Q)$$
$$T(f,[a,b]) \equiv \sup_{Q} T(f,Q),$$

where the supremum in each case is taken over all finite partitions Q of $[a, b]$.

Definition 20.1.3. Let $f : [a, b] \to \mathbb{R}$, where $-\infty < a < b < \infty$. Then, f is said to be of bounded variation on $[a, b]$ if $T(f, [a, b]) < \infty$. The set of all such functions is denoted by $BV[a, b]$.

As noted earlier, if f is non-decreasing, then $T(f, Q) = f(b) - f(a)$ for each Q and hence $T(f, [a, b]) = f(b) - f(a)$. It follows that if $f = f_1 - f_2$, where both f_1 and f_2 are non-decreasing, then $f \in BV[a, b]$. A natural question is whether the converse is true. The answer is yes, as shown by the following result.

Theorem 20.1.4. Let $f \in BV[a, b]$. Let $f_1(x) = P(f, [a, x])$ and $f_2(x) = N(f, [a, x])$. Then f_1 and f_2 are nondecreasing in $[a, b]$ and for all $a \leq x \leq b$,

$$f(x) = f_1(x) - f_2(x).$$

Proof. From the definition, it follows that f_1 and f_2 are nondecreasing. Then

$$f(b) - f(a) = P(f, [a, b]) - N(f, [a, b]),$$

as this can be applied to $[a, x]$ for $a \leq x < b$. For each finite partition Q of $[a, b]$,

$$\begin{aligned} f(b) - f(a) &= \sum_{i=1}^{n} (f(x_i) - f(x_{i-1})) \\ &= P(f, Q) - N(f, Q). \end{aligned}$$

Thus, $P(f, Q) = f(b) - f(a) + N(f, Q)$. By taking supremum over all finite partitions Q, it follows that

$$P(f, [a, b]) = f(b) - f(a) + N(f, [a, b]).$$

If $f \in BV[a, b]$, this yields $f(b) - f(a) = P(f, [a, b]) - N(f, [a, b])$. $\square$

Remark 20.1.5. Since $T(f, Q) = P(f, Q) + N(f, Q) = 2P(f, Q) - (f(b) - f(a))$, it follows that if $f \in BV[a, b]$, then

$$\begin{aligned} T(f, [a, b]) &= 2P(f, [a, b]) - (f(b) - f(a)) \\ &= P(f, [a, b]) + N(f, [a, b]). \end{aligned}$$

References

1. R. F. Bass: Functional Analysis.
2. A. E. Taylor: Introduction to Functional Analysis.
3. E. Kreyszing: Introductory Functional Analysis with Applications.
4. B. V. Limaye: Functional Analysis.
5. A. N. Kolmogorov and S. V. Fomin: Elements of the Theory of Functions and Functional Analysis.
6. P. K. Jain: Functional Analysis.
7. K. B. Athreya and S. Lahiri: Measure Theory.
8. Stein and Sakarchy: Real Analysis