Mathematics 4
Further Mathematics for Physicists

Ph. D. Vitalii Konarovskyi

IPSP Summer 2020

April 6 – July 16

Contents

1 Main Classes of Sets (Lecture Notes)

1.1 Jordan Measure

Let A be a subset of $\mathbb{R}^d$. How can we define the volume of A? If A is a rectangle:

$$A = [a_1, b_1] \times \cdots \times [a_d, b_d] = \{x = (x_k)_{k=1}^d : a_k \leqslant x_k \leqslant b_k,\ k = 1, \ldots, n\},$$

then

$$V(A) = \prod_{k=1}^{n} (b_k - a_k).$$

What if A is more general as in Figure 1.1?

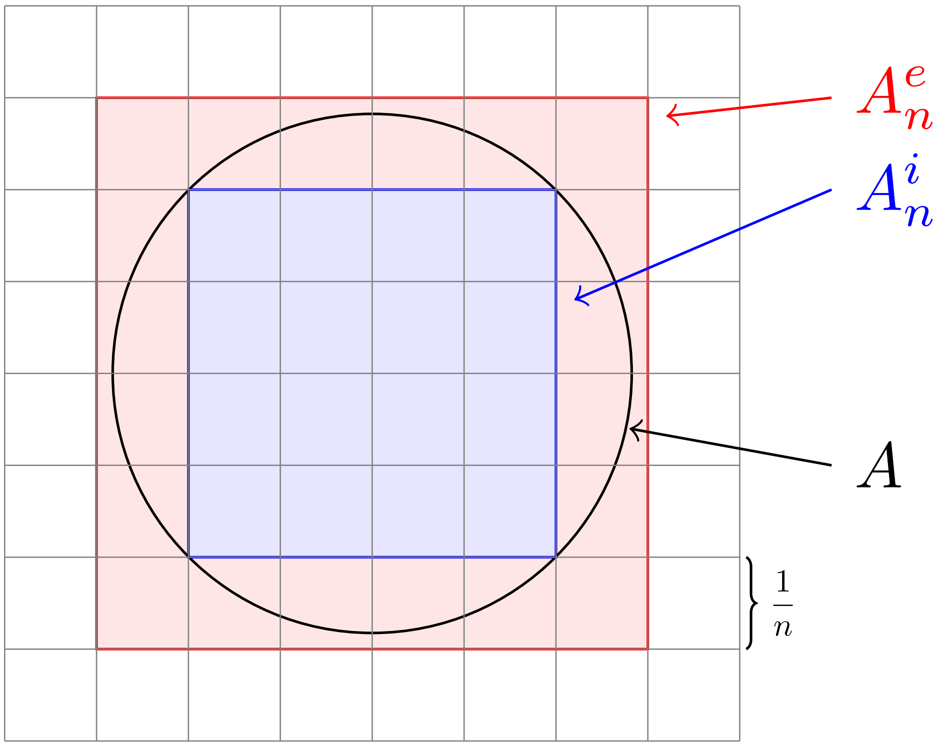

Figure 1.1: $A \subset \mathbb{R}^2$

If $\lim\limits_{n\to\infty} V(A_n^i) = \lim\limits_{n\to\infty} V(A_n^e)$, then we can say that the volume of A exists and is

$$V(A) = \lim_{n\to\infty} V(A_n^i).$$

Definition 1.1 *$V(A)$ is called the Jordan measure of A.*

Remark 1.2 *The Jordan measure was defined in Mathematics 3, Lecture 2 as*

$$V(A) = \mu(A) = \int\limits_I \mathbb{I}_A(x)\,dx = \int\limits_A dx,$$

where $I \supset A$ is a rectangle and

$$\mathbb{I}_A(x) = \begin{cases} 1, & x \in A, \\ 0, & x \notin A. \end{cases}$$

So, we can compute the volume of more general sets, but does this definition satisfy "intuitive" properties of volume? For example, let A and B be Jordan measurable.

1. $A \cup B$ is Jordan measurable and $V(A \cup B) = V(A) + V(B)$ if $A \cap B = \emptyset$
2. $A \setminus B$ is Jordan measurable and $V(A \setminus B) = V(A) - V(B)$ if $B \subset A$
3. $A \cap B$ is Jordan measurable

Let $A_1, A_2, \dots$ be Jordan measurable. Then

$$\bigcup_{n=1}^{\infty} A_n = \{x : \exists\, n \geqslant 1,\, x \in A_n\}$$

is not Jordan measurable in general.

Example 1.3 *Take $A = [0,1]^2 \cap \mathbb{Q}^2$, which is the set of all points from $[0,1]^2$ with rational coefficients. We know that A is countable, so $A = \{x_1, x_2, \dots\}$. Moreover, A is not Jordan measurable. However, one point sets $A_n = \{x_n\}$ are Jordan measurable and $V(A_n) = 0$.*

We find that $V(A_n) = 0$ but $V(A) = V\left(\bigcup_{n=1}^{\infty} A_n\right)$ does not exist as we cannot define it. Intuitively

$$V(A) \leqslant \sum_{n=1}^{\infty} V(A_n) = 0 \Rightarrow V(A) = 0.$$

This demonstrates that the Jordan measure is not well-defined for some sets which intuitively should have volume. Our goal is to define a volume, or measure in general, for a wider class of sets, which would satisfy the "intuitive" or expected properties. In particular, we expect that if we can define the measure for sets $A_1, A_2, \dots \subset \mathbb{R}^d$, then the volume must exist for any set obtained from $A_1, A_2, \dots$ by a countable number of operations like $\cap, \cup, \setminus$, and taking the complement.

1.2 Definitions of Main Classes of Sets

In this section, we will describe the classes of sets for which we can define a measure. Let X be a fixed, non-empty set. We denote by 2^X the family of all subsets of X.

Definition 1.4

- *A non-empty class of sets $H \subset 2^X$ is called a semiring if*
  1. $A, B \in H \Rightarrow A \cap B \in H$,
  2. $A, B \in H \Rightarrow \exists\, n \in \mathbb{N},\, \exists\, C_1, \dots, C_n \in H,\, C_j \cap C_k = \emptyset,\, j \neq k : A \setminus B = \bigcup_{k=1}^{n} C_k.$
- *A class H is called a semialgebra if H is a semiring and $X \in H$.*

Remark 1.5 *A semiring usually contains "simple" sets where a measure can be easily defined.*

Example 1.6 *Let $X = \mathbb{R}$.*

1. $H_1 = \{[a,b) : -\infty < a < b < \infty\} \cup \{\emptyset\}$ *is a semiring.*
2. $H_2 = \{[a,b) : -\infty < a < b < \infty\} \cup \{\emptyset, \mathbb{R}\} \cup \{(-\infty, b) : b < \infty\} \cup \{[a, \infty) : -\infty < a\}$ *is a semialgebra.*

Example 1.7 *Let $X = \mathbb{R}^2$.*

1. $H_1 = \{[a_1, b_1) \times [a_2, b_2) : -\infty < a_1 < b_1 < \infty,\ -\infty < a_2 < b_2 < \infty\} \cup \{\emptyset\}$ *is a semiring.*

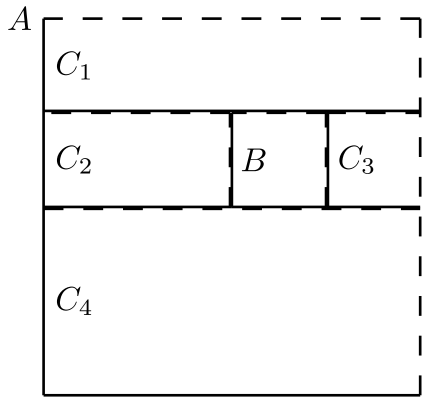

Figure 1.2

In this case $A \setminus B = C_1 \cup C_2 \cup C_3 \cup C_4$.

2. H_2 *can be defined in the same way as in Example 1.6 and it would be a semialgebra.*

One can see that the measure can be easily defined for sets like H_1 from Examples 1.6 and 1.7.

Definition 1.8

- *A non-empty class* $H \subset 2^X$ *is called a ring if*
  1. $A, B \in H \Rightarrow A \cup B \in H$,
  2. $A, B \in H \Rightarrow A \setminus B \in H$.
- *A class* H *is said to be an algebra if* H *is a ring and* $X \in H$.

Exercise 1.9 *Let H be a ring (algebra). Show that H is a semiring (semialgebra, respectively).*

Exercise 1.10 *Let H be a ring. Show that*

1. $\emptyset \in H$,
2. $A, B \in H \Rightarrow A \cap B \in H$,
3. $A_1, \ldots, A_n \in H \Rightarrow \bigcup_{k=1}^{n} A_k \in H,\ \bigcap_{k=1}^{n} A_k \in H$.

Proposition 1.11 *A non-empty class H is an algebra if and only if*

1. $A, B \in H \Rightarrow A \cup B \in H$
2. $A \in H \Rightarrow A^c = X \setminus A \in H$

Proof: Assume that H is an algebra. Then the first condition is trivially fulfilled by definition. We know that $A, X \in H$. Then by Definition 1.8 we have the second condition: $A^c = X \setminus A \in H$. Now we assume the converse. The first condition of Definition 1.8 is immediately satisfied. To check the second, take $A, B \in H$. We have

$$A \setminus B = A \cap B^c = (A \cap B^c)^{cc} = (A^c \cup B)^c.$$

Since we know that $A^c \in H$, then $A \setminus B \in H$. Remark that $X = A \cup A^c \in H$. $\square$

2 Generated Classes of Sets, The Borel σ-Algebra (Lecture Notes)

2.1 σ-Rings and σ-Algebras

Let X be a fixed set and let 2^X denote a class of all subsets of X. We recall that $H \subset 2^X$ is

1. a semiring if for all $A, B \in H$
   (a) $A \cap B \in H$
   (b) $A \setminus B = \bigcup_{k=1}^{n} C_k$, where $C_j \cap C_k = \emptyset$ for $j \neq k$, and $C_k \in H$ for $k = 1, \dots, n$
2. a semialgebra if it is a semiring and if $X \in H$
3. a ring if for all $A, B \in H$
   (a) $A \cup B \in H$
   (b) $A \setminus B \in H$

   (a ring is closed with respect to a finite number of operations $\cap, \cup, \setminus$)
4. an algebra if it is a ring and if $X \in H$ (an algebra is also closed with respect to the complement)

Definition 2.1

- *A non-empty class of sets $H \subset 2^X$ is called a σ-ring if*
  1. $A_1, A_2, \dots \in H \Rightarrow \bigcup_{n=1}^{\infty} A_n \in H$,
  2. $A, B \in H \Rightarrow A \setminus B \in H$.
- *A class H is called a σ-algebra if H is a σ-ring and $X \in H$.*

Proposition 2.2 *A non-empty class H is a σ-algebra if and only if*

1. $X \in H$
2. $A_1, A_2, \dots \in H \Rightarrow \bigcup_{n=1}^{\infty} A_n \in H$
3. $A \in H \Rightarrow A^c \in H$

Proof: The proof is similar to the proof of Proposition 1.11. □

Example 2.3 *Let $X = \mathbb{R}^2$ and let $H = \{A \subset \mathbb{R}^d : A \, \text{is Jordan measurable and} \, \mu(A) < \infty\}$. We know that if $A, B \in H$, that is, if A, B are Jordan measurable, then $A \cup B$ and $A \setminus B$ are also Jordan measurable, and $\mu(A \cup B) < \infty$, $\mu(A \setminus B) < \infty$. Hence $A \cup B$, $A \setminus B \in H$. This implies that H is a ring. However, note that H is not a σ-ring. Indeed, $\mathbb{Q}^2 = \bigcup_{n=1}^{\infty} A_n$ is a countably infinite union of Jordan measurable single point sets with $\mu(A_k) = 0$, but $\mathbb{Q}^2$ is not Jordan measurable. Additionally, H is neither an algebra nor σ-algebra, since $\mu(\mathbb{R}^2) \not< \infty \Rightarrow \mathbb{R}^2 \notin H$.*

Example 2.4 *Let $X = [0,1]^2$ and let $H = \{A \subset [0,1]^2 : A \, \text{is Jordan measurable}\}$. Then H is an algebra but not a σ-algebra.*

Exercise 2.5 *Let H be a σ-ring. Prove that $A_1, A_2, \dots \in H \Rightarrow \bigcap_{n=1}^{\infty} A_n \in H$.*

Remark 2.6

- *A σ-ring is a class closed with respect to a countable number of operations $\cap, \cup, \setminus$.*
- *A σ-algebra is additionally closed with respect to taking the complement.*

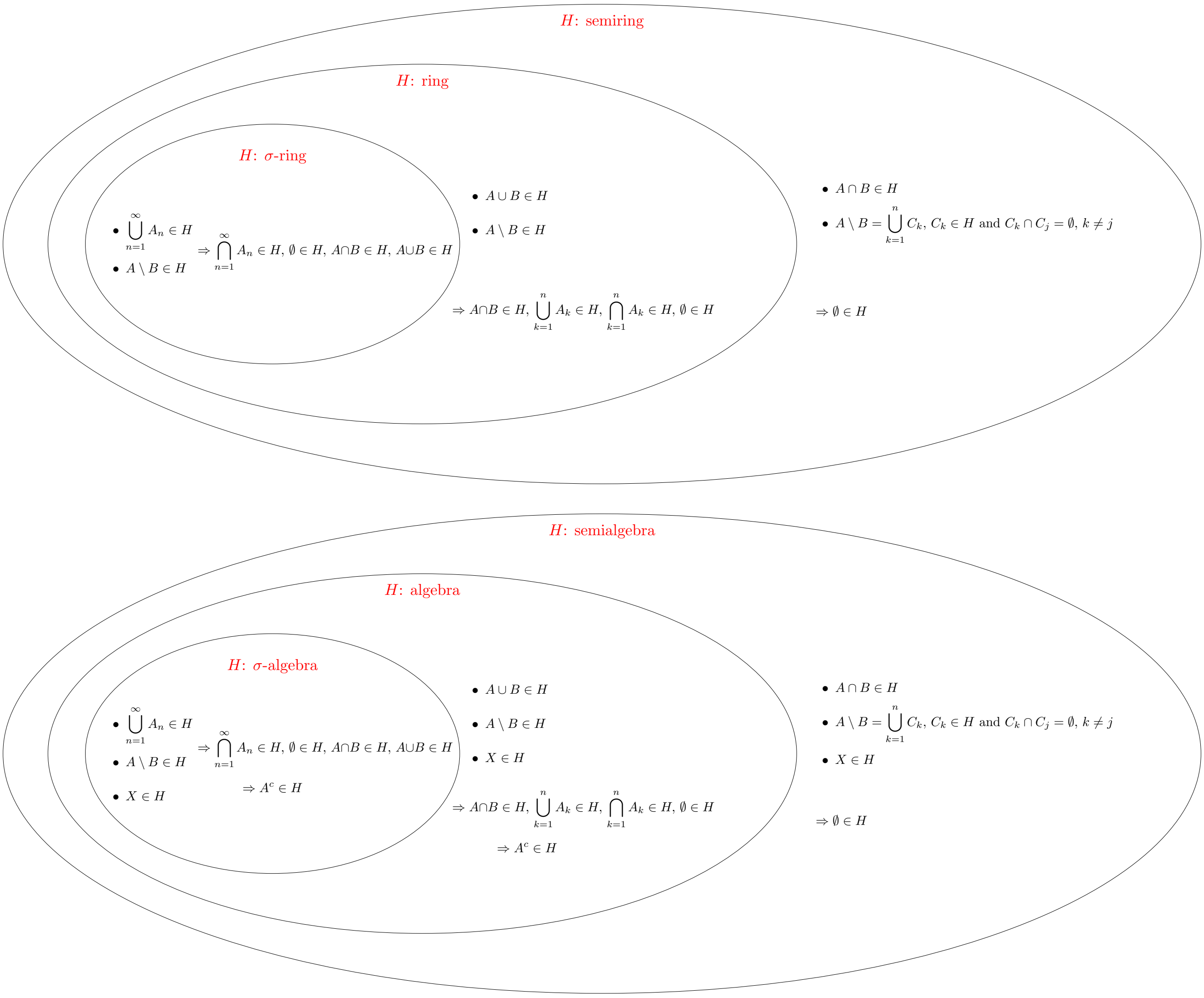

Figure 2.1

2.2 Generated Classes of Sets

Let H be a class of subsets of X.

Definition 2.7

- *The smallest σ-algebra which contains the class H is called the (smallest) σ-algebra generated by H and is denoted by $\sigma(H)$.*
- *The same definition is given for the ring $r(H)$, the algebra $a(H)$, and the σ-ring $\sigma r(H)$ generated by H.*

Example 2.8 *Take $X = \{a, b, c\}$ and $H = \{a, b\}$.*

1. *Then $\sigma(H) = \{\emptyset, X, \{a, b\}, \{c\}\}$. There are other σ-algebras containing H like 2^X, but they are not the smallest. Remark that $\sigma(H) = a(H)$ in this case.*

2. *Then $\sigma r(H) = \{\emptyset, \{a, b\}\} = r(H)$.*

Theorem 2.9 *The σ-algebra generated by H always exists.*

Proof: We construct

$$\sigma(H) = \bigcap_{H \subset \mathcal{A}} \mathcal{A},$$

where $\mathcal{A}$ is a σ-algebra containing H. In other words, $\sigma(H)$ is the class of all sets A such that A belongs to every σ-algebra containing H. Then $\sigma(H)$ is a σ-algebra. Indeed, if $A_1, A_2, \dots \in \sigma(H)$, then they belong to every σ-algebra containing H. That is, if $\mathcal{A}$ is a σ-algebra containing H, then $A_1, A_2, \dots \in \mathcal{A}$. Consequently $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$ for all σ-algebra $\mathcal{A}$ containing H. Hence $\bigcup_{n=1}^{\infty} A_n \in \sigma(H)$. Similarly, we can show that $A \in \sigma(H) \Rightarrow A^c \in \sigma(H)$, and $X \in \sigma(H)$. Proposition 2.2 implies that $\sigma(H)$ is a σ-algebra and it is trivial that it is the smallest one. □

Remark 2.10 *The same statement is true for $a(H)$, $r(H)$, and $\sigma r(H)$.*

Theorem 2.11 *Let H be a semiring. Then*

$$r(H) = \left\{ \bigcup_{k=1}^{n} A_k : A_1, \dots, A_n \in H,\ n \geqslant 1 \right\}.$$

Corollary 2.12 *Let H be a semialgebra. Then*

$$a(H) = \left\{ \bigcup_{k=1}^{n} A_k : A_1, \dots, A_n \in H,\ n \geqslant 1 \right\}.$$

Example 2.13 *If $H = \{[a, b) : -\infty < a < b < \infty\} \cup \{\emptyset\}$, then*

$$r(H) = \left\{ A = \bigcup_{k=1}^{n} [a_k, b_k) : -\infty < a_k < b_k < \infty,\ k = 1, \dots, n,\ n \geqslant 1 \right\}.$$

Exercise 2.14 *Let $H_1 \subset H_2 \subset \sigma(H_1)$. Show that $\sigma(H_1) = \sigma(H_2)$.*
Solution:
We first remark that $H_1 \subset H_2 \Rightarrow H_1 \subset \sigma(H_2)$, so $\sigma(H_2)$ is a σ-algebra containing H_1. This implies that $\sigma(H_1) \subset \sigma(H_2)$, because $\sigma(H_1)$ is the smallest σ-algebra which contains H_1. We also know that $H_2 \subset \sigma(H_1)$, so, similarly $\sigma(H_2) \subset \sigma(H_1)$. Hence $\sigma(H_1) = \sigma(H_2)$.

2.3 Borel Sets

In this section, we will assume that $X = \mathbb{R}^d$. Let

$$H = \{[a_1, b_1) \times \cdots \times [a_d, b_d) : -\infty < a_k < b_k < \infty\} \cup \{\emptyset\}.$$

We know from Lecture 1 that H is a semiring.

Definition 2.15 *The σ-algebra $\mathcal{B}(\mathbb{R}^d) := \sigma(H)$ is called the Borel σ-algebra. Sets from $\mathcal{B}(\mathbb{R}^d)$ are called Borel sets.*

Remark 2.16 *The Borel σ-algebra contains all rectangles as well as all sets which can be obtained from rectangles by a countable number of operations $\cap, \cup, \setminus$, and taking the complement.*

Example 2.17 *Let $X = \mathbb{R}$.*

1. $\{a\} \in \mathcal{B}(\mathbb{R}), \forall\, a \in \mathbb{R}$
   $\{a\} = \bigcap_{n=1}^{\infty} \left[a, a + \frac{1}{n}\right)$
2. $\mathbb{Q} \in \mathcal{B}(\mathbb{R})$
   $\mathbb{Q} = \bigcup_{a \in \mathbb{Q}} \{a\}$
3. $[a, b] \in \mathcal{B}(\mathbb{R})$
   $[a, b] = \bigcap_{n=1}^{\infty} \left[a, b + \frac{1}{n}\right)$
4. $(a, b) \in \mathcal{B}(\mathbb{R})$
   $(a, b) = \bigcup_{n=1}^{\infty} \left[a + \frac{1}{n}, b\right)$
5. *Any open set $G \subset \mathbb{R}$ belongs to $\mathcal{B}(\mathbb{R})$ as $G = \bigcup_{n=1}^{\infty} (a_n, b_n)$.*
6. *Any closed set F belongs to $\mathcal{B}(\mathbb{R})$ since F^c is open.*

Lemma 2.18 *Let $\tilde{H} = \{A \subset \mathbb{R}^d : A\,\text{is open}\}$. Then $\sigma(\tilde{H}) = \mathcal{B}(\mathbb{R}^d)$. In other words, the Borel σ-algebra is generated by all open subsets of $\mathbb{R}^d$.*

Proof: By Example 2.17, 5), which is true for any dimension d, we have $\tilde{H} \subset \mathcal{B}(\mathbb{R}^d)$. Hence $\sigma(\tilde{H}) \subset \mathcal{B}(\mathbb{R}^d)$. Next, we remark that

$$[a_1, b_1) \times \cdots \times [a_d, b_d) = \bigcap_{n=1}^{\infty} \left(\left(a_1 - \tfrac{1}{n}, b_1\right) \times \cdots \times \left(a_d - \tfrac{1}{n}, b_d\right)\right).$$

So $H \subset \sigma(\tilde{H}) \Rightarrow \sigma(H) \subset \sigma(\tilde{H})$. Hence $\mathcal{B}(\mathbb{R}^d) = \sigma(\tilde{H})$. $\square$

3 Properties of Measures (Lecture Notes)

3.1 Definition of a Measure and Basic Properties

Let X be a fundamental set and let $H \subset 2^X$ be a class of sets. The main object of measure theory is to find functions

$$\mu : H \mapsto (-\infty, \infty),$$

which satisfy certain requirements. Length, area, and volume are real examples of such functions. They lead to a class of functions which satisfy certain properties. For example, the area is nonnegative and the area of two nonintersecting sets equals the sum of the areas of those sets. We will generalize these properties to an abstract situation. We will assume that μ can take the value ∞. Moreover, we assume that

$$\infty + \infty = \infty, \quad a + \infty = \infty, \, \forall \, a \in \mathbb{R}, \, a < \infty.$$

Definition 3.1 *A function $\mu : H \mapsto (-\infty, \infty]$ is called*

1. *nonnegative if $\mu(A) \geqslant 0$, $\forall \, A \in H$*
2. *countably additive or σ-additive if $\forall \, A_n \in H$, $n \geqslant 1$, where $A_j \cap A_k = \emptyset$, $j \neq k$, we have*

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$$

Definition 3.2 *A measure is a nonnegative and σ-additive function on a semiring.*

Remark 3.3 *If μ is a measure on H then $\mu(\emptyset) = 0$. Indeed, if we take $A_1 = A \in H$ with $\mu(A) < \infty$ and $A_2 = A_3 = \cdots = \emptyset \in H$, then*

$$\mu(A) = \mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n) = \sum_{n=2}^{\infty} \mu(\emptyset) + \mu(A) \Rightarrow \mu(\emptyset) = 0.$$

Remark 3.4 *A measure is a also an additive function, that is, for all $A_k \in H$, $k = 1, \ldots, n$, where $A_j \cap A_k = \emptyset$, $j \neq k$, we have*

$$\mu\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} \mu(A_k).$$

This follows from Remark 3.3 because we can take $A_{n+1} = A_{n+2} = \cdots = \emptyset$. Then

$$\mu\left(\bigcup_{k=1}^{n} A_k\right) = \mu\left(\bigcup_{k=1}^{\infty} A_k\right) = \sum_{k=1}^{\infty} \mu(A_k) = \sum_{k=1}^{n} \mu(A_k) + \mu(A_{n+1}) + \cdots = \sum_{k=1}^{n} \mu(A_k).$$

Example 3.5 *Let $X = \mathbb{N} = \{1, 2, 3, \ldots\}$ and let $H = 2^X$. We set*

$$\mu(A) = \begin{cases} \text{number of elements of } A \text{ if } A \text{ is finite,} \\ \infty \text{ if } A \text{ is infinite.} \end{cases}$$

Then, for example, $\mu(\{1, 7, 8, 10\}) = 4$ and $\mu(\{\text{even numbers}\}) = \infty$. It is easy to see that μ is a measure.

Exercise 3.6 *Let* $X = \{x_1, x_2, \dots, x_n, \dots\}$, $H = 2^X$. *Take* $p_n \geqslant 0$, $n \geqslant 1$ *such that* $\sum_{n=1}^{\infty} p_n = 1$, *and set*

$$\mu(A) = \sum_{n: x_n \in A} p_n, \; A \in H.$$

For example, $\mu(\{x_1, x_{10}, x_{100}\}) = p_1 + p_{10} + p_{100}$. *Prove that* μ *is a measure on* H.

Theorem 3.7 *Let* R *be a ring and let* μ *be a measure on* R.

1. μ *is monotone on* R, *that is, for all* $A, B \in R$ *such that* $A \subset B$ *we have* $\mu(A) \leqslant \mu(B)$

2. $\forall\, A, B \in R$ *such that* $A \subset B$, $\mu(A) < \infty$ *we have* $\mu(B \setminus A) = \mu(B) - \mu(A)$

3. $\forall\, A, B \in R$ *such that* $\mu(A) < \infty$ *or* $\mu(B) < \infty$ *we have* $\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$

4. $\forall\, B_1, \dots, B_n, A \in R$ *such that* $A \subset \bigcup_{k=1}^{n} B_k$ *we have*

$$\mu(A) \leqslant \sum_{k=1}^{n} \mu(B_k)$$

5. μ *is* σ*-semiadditive, that is,* $\forall\, A_1, A_2, \dots \in R$ *such that* $\bigcup_{n=1}^{\infty} A_n \in R$ *we have*

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) \leqslant \sum_{n=1}^{\infty} \mu(A_n)$$

(we do not assume that $A_j \cap A_k = \emptyset$, $j \neq k$*)*

Proof:

1. Take $A, B \in R$ such that $A \subset B$. Then $B = A \cup (B \setminus A)$ and $A \cap (B \setminus A) = \emptyset$. By Remark 3.4

$$\mu(B) = \mu(A) + \mu(B \setminus A) \geqslant \mu(A). \tag{3.1}$$

2. If $\mu(A) < \infty$, then (3.1) implies

$$\mu(B \setminus A) = \mu(B) - \mu(A).$$

3. If $\mu(A) < \infty$ or $\mu(B) < \infty$, then by 1) $\mu(A \cap B) < \infty$. We can write

$$A \cup B = \big(A \setminus (A \cap B)\big) \cup B, \quad \big(A \setminus (A \cap B)\big) \cap B = \emptyset.$$

Then using Remark 3.4 and 2) we have

$$\mu(A \cup B) = \mu\big(A \setminus (A \cap B)\big) + \mu(B) = \mu(A) - \mu(A \cap B) + \mu(B).$$

4. Remark that

$$\bigcup_{k=1}^{n} B_k = B_1 \cup (B_2 \setminus B_1) \cup \big(B_3 \setminus (B_1 \cup B_2)\big) \cup \dots \cup \left(B_n \setminus \bigcup_{k=1}^{n-1} B_k\right).$$

Then using Remark 3.4 and 1) we have

$$\mu(A) \leqslant \mu\left(\bigcup_{k=1}^{n} B_k\right) = \sum_{k=1}^{n} \mu\left(B_k \setminus \bigcup_{l=1}^{k-1} B_l\right) \leqslant \sum_{k=1}^{n} \mu(B_k).$$

5. Using σ-additivity and 1) we have

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \mu\left(\bigcup_{n=1}^{\infty}\left(A_n \setminus \bigcup_{k=1}^{n-1} A_k\right)\right) = \sum_{n=1}^{\infty} \mu\left(A_n \setminus \bigcup_{k=1}^{n-1} A_k\right) \leqslant \sum_{n=1}^{\infty} \mu(A_n).$$

□

Exercise 3.8 *Let μ be a measure on a σ-ring H. Let $A_n \in H$ be such that $\mu(A_n) = 0$, $n \geqslant 1$. Show that*

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = 0.$$

3.2 Continuity of a Measure

Theorem 3.9 *Let R be a ring on which μ is a measure. Then for any increasing sequence $A_n \in R$, $n \geqslant 1$, where $A_n \subset A_{n+1}$, $\forall\, n \geqslant 1$, such that $\bigcup_{n=1}^{\infty} A_n \in R$, one has*

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n).$$

Proof:

I. If there exists n_0 such that $\mu(A_{n_0}) = \infty$, then for all $n \geqslant n_0$, we have $\mu(A_n) \geqslant \mu(A_{n_0}) = \infty$ and $\mu\left(\bigcup_{n=1}^{\infty} A_n\right) \geqslant \mu(A_{n_0}) = \infty$. Hence $\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n) = \infty$.

II. If $\mu(A_n) < \infty$, $\forall\, n \geqslant 1$, then

$$\begin{aligned}
\mu\left(\bigcup_{n=1}^{\infty} A_n\right) &= \mu\big(A_1 \cup (A_2 \setminus A_1) \cup (A_3 \setminus A_2) \cup \dots \cup (A_k \setminus A_{k-1}) \cup \dots\big) \\
&= \mu(A_1) + \sum_{k=1}^{\infty} \mu(A_k \setminus A_{k-1}) \\
&= \mu(A_1) + \lim_{n\to\infty} \sum_{k=1}^{n} \mu(A_k \setminus A_{k-1}) \\
&= \mu(A_1) + \lim_{n\to\infty} \big(\mu(A_2) - \mu(A_1) + \mu(A_3) - \mu(A_2) + \dots + \mu(A_n) - \mu(A_{n-1})\big) \\
&= \lim_{n\to\infty} \mu(A_n).
\end{aligned}$$

□

Theorem 3.10 *Let R be a ring and let μ be a measure on R. Then for any decreasing sequence $A_n \in R$, $n \geqslant 1$, where $A_n \supset A_{n+1}$, $\forall\, n \geqslant 1$, such that $\bigcap_{n=1}^{\infty} A_n \in R$ and $\mu(A_1) < \infty$, one has*

$$\mu\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n\to\infty} \mu(A_n).$$

Proof: We have

$$\mu\left(A_1 \setminus \bigcap_{n=2}^{\infty} A_n\right) = \mu\left(\bigcup_{n=2}^{\infty}(A_1 \setminus A_n)\right) = \lim_{n\to\infty} \mu(A_1 \setminus A_n) = \lim_{n\to\infty} (\mu(A_1) - \mu(A_n)).$$

Hence

$$\mu(A_1) - \mu\left(\bigcap_{n=1}^{\infty} A_n\right) = \mu\left(A_1 \setminus \bigcap_{n=2}^{\infty} A_n\right) = \lim_{n\to\infty} (\mu(A_1) - \mu(A_n)).$$

□

Remark 3.11 *The condition $\mu(A_1) < \infty$ is important in Theorem 3.10. Consider the measure from Example 3.5. Let*

$$A_n = \{n, n+1, \dots\},\ n \geqslant 1.$$

Obviously $A_n \supset A_{n+1}$, $\forall\, n \geqslant 1$, and $\bigcap_{n=1}^{\infty} A_n = \emptyset$, so $\mu\left(\bigcap_{n=1}^{\infty} A_n\right) = 0$. But $\lim\limits_{n\to\infty} \mu(A_n) = \infty$.

3.3 Examples of Measures

Theorem 3.12 *Let R be a ring of all Jordan measurable sets on $\mathbb{R}^d$ and let μ be the Jordan measure on R. Then the function μ is σ-additive on R, that is, it is a measure according to Definition 3.2.*

Corollary 3.13 *Let $X = \mathbb{R}$ and take the semiring $H = \{(a,b] : -\infty < a < b < \infty\} \cup \{\emptyset\}$. Then the function*

$$\lambda\big((a,b]\big) = b - a, \quad \lambda(\emptyset) = 0$$

is a measure on H.

Theorem 3.14 *Take $X = \mathbb{R}$ and $H = \{(a,b] : -\infty < a < b < \infty\} \cup \{\emptyset\}$. Let $F : \mathbb{R} \mapsto \mathbb{R}$ be a nonnegative right continuous function on $\mathbb{R}$. Define*

$$\lambda_F\big((a,b]\big) = F(b) - F(a),\ a < b, \quad \lambda_F(\emptyset) = 0.$$

Then the function is a measure on the semiring H.

4 Extensions of Measures (Lecture Notes)

4.1 Extending a Measure from a Semiring to a Generated Ring

Let X be a universal set and let $H \subset 2^X$. We recall that a nonnegative and σ-additive function μ defined on a semiring H is called a measure, that is, a measure $\mu : H \mapsto \mathbb{R}$ must satisfy the following properties:

1. $\mu(A) \geqslant 0, \forall\, A \in H$
2. $\forall\, A_n \in H,\, n \geqslant 1$, where $A_j \cap A_k = \emptyset,\, j \neq k$, we have $\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$

In this section, we will consider the extension of a measure from a semiring H to a ring. Recall that Theorem 2.11 implies that

$$r(H) = \left\{ \bigcup_{k=1}^{n} A_k : A_1, \ldots, A_n \in H,\, n \geqslant 1 \right\}.$$

Example 4.1 *Let* $H = \{[a,b) : a < b\} \cup \{\emptyset\}$. *Then*

$$r(H) = \left\{ \bigcup_{k=1}^{n} [a_k, b_k) : a_k < b_k,\, n \geqslant 1 \right\} \cup \{\emptyset\}.$$

For example, take the set $[2,5) \cup [7,10) \cup [9,11) = [2,5) \cup [7,11) \in r(H)$.

Theorem 4.2 *Let μ be a measure on a semiring H. The measure μ can be extended to a measure on $r(H)$ and this extension is unique. Moreover, if the measure μ is finite, then the extension is finite.*

Let $r(H) \ni A = \bigcup_{k=1}^{n} A_k,\, A_k \in H$. We first remark that there exists $C_1, C_2, \ldots, C_m \in H$ such that $C_j \cap C_k = \emptyset,\, j \neq k$ and

$$A = \bigcup_{k=1}^{n} A_k = \bigcup_{k=1}^{m} C_k. \tag{4.1}$$

Then the described extension is given by

$$\mu(A) := \sum_{k=1}^{m} \mu(C_k).$$

For example, take $A = \bigcup_{k=1}^{3} A_k = A_1 \cup (A_2 \setminus A_1) \cup \big(A_3 \setminus (A_1 \cup A_2)\big) = C_1 \cup \cdots \cup C_8$.

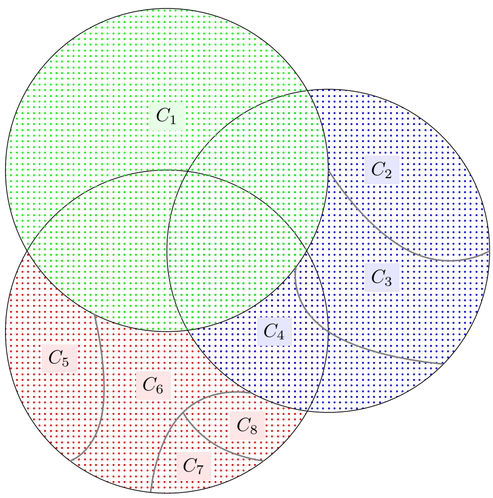

Figure 4.1

4.2 Outer Measure

Definition 4.3 *A function $\lambda^* : 2^X \mapsto (-\infty, \infty]$ is called an outer measure if*

1. $\lambda^*(\emptyset) = 0$, $\lambda^* \geqslant 0$ *(nonnegativity)*
2. $\forall\, A, A_n \in 2^X : A \subset \bigcup_{n=1}^{\infty} A_n$ *we have* $\lambda^*(A) \leqslant \sum_{n=1}^{\infty} \lambda^*(A_n)$ *(σ-semiadditivity)*

Definition 4.4 *Let μ be a measure on a ring $R \subset 2^X$. For any $A \in 2^X$ ($A \subset X$) set*

$$\mu^*(A) = \begin{cases} 0 \text{ if } A = \emptyset, \\ \infty \text{ if cover does not exist}, \\ \inf \left\{ \sum_{n=1}^{\infty} \mu(A_n) : A_n \in R,\, n \geqslant 1,\, A \subset \bigcup_{n=1}^{\infty} A_n \right\} \text{otherwise}. \end{cases}$$

Theorem 4.5 *μ^* is an outer measure on R.*

Proof: We only need to show that for any $A, A_n \in 2^X$, $n \geqslant 1$, $A \subset \bigcup_{n=1}^{\infty} A_n$ we have

$$\mu^*(A) \leqslant \sum_{n=1}^{\infty} \mu^*(A_n).$$

It is enough to show this only in the case $\mu^*(A_n) < \infty$, $n \geqslant 1$. Take $\epsilon > 0$. According to Definition 4.4, for all A_n there exists $B_{k,n} \in R$, $k \geqslant 1$ such that

$$A_n \subset \bigcup_{k=1}^{\infty} B_{k,n}, \quad \sum_{k=1}^{\infty} \mu(B_{k,n}) < \mu^*(A_n) + \frac{\epsilon}{2^n}.$$

Hence $A \subset \bigcup_{n=1}^{\infty} A_n \subset \bigcup_{n=1}^{\infty} \bigcup_{k=1}^{\infty} B_{k,n}$. By Definition 4.4

$$\mu^*(A) \leqslant \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \mu(B_{k,n}) < \sum_{n=1}^{\infty} \left(\mu^*(A_n) + \frac{\epsilon}{2^n} \right) = \sum_{n=1}^{\infty} \mu^*(A_n) + \epsilon.$$

Making $\epsilon \to 0^+$, we have

$$\mu^*(A) \leqslant \sum_{n=1}^{\infty} \mu^*(A_n).$$

□

Definition 4.6 *The function μ^* from Definition 4.4 is the outer measure generated by the measure μ.*

4.3 λ^*-Measurable Sets, Carathéodory Theorem

Definition 4.7 *Let λ^* be an outer measure on 2^X. A set A is called λ^*-measurable if $\forall\, B \subset X$ we have*

$$\lambda^*(B) = \lambda^*(B \cap A) + \lambda^*(B \setminus A). \tag{4.2}$$

Remark 4.8 *By the definition of an outer measure, the inequality*

$$\lambda^*(B) \leqslant \lambda^*(B \cap A) + \lambda^*(B \setminus A)$$

is always true since $B \subset (B \cap A) \cup (B \setminus A)$.

Theorem 4.9 (Carathéodory Theorem) *Let λ^* be an outer measure on 2^X and let $\mathcal{S}$ be the class of all λ^*-measurable sets. Then $\mathcal{S}$ is a σ-algebra and λ^* is a measure on $\mathcal{S}$.*

Definition 4.10 *A measure μ on a σ-algebra H is called complete if $\forall\, A \in H$ such that $\mu(A) = 0$ we have that any subset $C \subset A$ also belongs to H (in this case, $\mu(C) = 0$ by monotonicity).*

Proposition 4.11 *Under the assumptions of Theorem 4.9, the measure λ^* is complete on $\mathcal{S}$.*

Proof: Let $A \in \mathcal{S}$ be such that $\lambda^*(A) = 0$ and $C \subset A$. We need to show that $C \in \mathcal{S}$. We will check (4.2) for C. Let $B \in 2^X$. By the monotonicity of λ^*, we have

$$\lambda^*(B) \geqslant \lambda^*(B \cap C^c) \geqslant \lambda^*(B \cap A^c) = \lambda^*(B \cap A) + \lambda^*(B \cap \overline{A}) = \lambda^*(B),$$

since $0 \leqslant \lambda^*(B \cap A) \leqslant \lambda^*(A) = 0$. Similarly, $0 \leqslant \lambda^*(B \cap C) \leqslant \lambda^*(A) = 0$, so

$$\lambda^*(B) = \lambda^*(B \cap C^c) = \lambda^*(B \setminus C) + \lambda^*(B \cap C).$$

□

4.4 μ^*-Measurability of Sets from a Ring

Let R be a ring and let μ be a measure on R, with μ^* being the outer measure generated by μ. Let $\mathcal{S}$ be the class of all μ^*-measurable subsets of X. We also denote

$$\overline{\mu}(A) = \mu^*(A),\ A \in \mathcal{S}.$$

By Theorem 4.9 $\mathcal{S}$ is a σ-algebra and $\overline{\mu}$ is a measure on $\mathcal{S}$.

Theorem 4.12 *We have $R \subset \mathcal{S}$ and $\overline{\mu}$ is the extension of μ from R to $\mathcal{S}$, that is $\overline{\mu}(A) = \mu(A)$, $\forall\, A \in R$.*

Proof: We first show that $\forall\, A \in R$, we have $\mu^*(A) = \mu(A)$. Since $A \subset A \cup \emptyset \cup \emptyset \cup \cdots = \bigcup_{k=1}^n A_k$, then

$$\mu^*(A) \leqslant \sum_{k=1}^{\infty} \mu(A_k) = \mu(A).$$

Now let $A \subset \bigcup_{n=1}^{\infty} A_n$, $A_n \in R$, $n \geqslant 1$. Then

$$A = \bigcup_{n=1}^{\infty} (A \cap A_n).$$

By the monotonicity and σ-semiadditivity of μ, we have

$$\mu(A) \leqslant \sum_{n=1}^{\infty} \mu(A \cap A_n) \leqslant \sum_{n=1}^{\infty} \mu(A_n).$$

Hence $\mu(A) \leqslant \mu^*(A)$, and consequently $\mu(A) = \mu^*(A)$. Now we will show that $R \subset \mathcal{S}$. Take $A \in R$ and $\epsilon > 0$. We consider any set $B \subset X$, $\mu^*(B) < \infty$ and show that

$$\mu^*(B) \geqslant \mu^*(B \cap A) + \mu^*(B \setminus A).$$

According to Definition 4.4, $\exists A_n \in R$, $n \geqslant 1$ such that $B \subset \bigcup_{n=1}^{\infty} A_n$ and $\mu^*(B) + \epsilon > \sum_{n=1}^{\infty} \mu(A_n)$. So

$$\mu^*(B) + \epsilon > \sum_{n=1}^{\infty} \mu(A_n) = \sum_{n=1}^{\infty} \big(\mu(A_n \cap A) + \mu(A_n \setminus A)\big) \geqslant \mu^*(B \cap A) + \mu^*(B \setminus A),$$

since $B \cap A \subset \bigcup_{n=1}^{\infty} A_n \cap A$ and $B \setminus A \subset \bigcup_{n=1}^{\infty} A_n \setminus A$. This along with Remark 4.8 implies that

$$\mu^*(B) = \mu^*(B \cap A) + \mu^*(B \setminus A).$$

□

4.5 Lebesgue Measure

Let $X = \mathbb{R}$ and take the semiring $H = \{(a, b] : a < b\} \cup \{\emptyset\}$. Define

$$\lambda(\emptyset) = 0, \quad \lambda\big((a, b]\big) := b - a, \; a < b.$$

Then, by Corollary 3.13, λ is a measure on H. Additionally, by Theorem 4.2 there exists an extension of λ to the ring $r(H)$ generated by H. Next, let $\mathcal{S}$ be the class of all λ^*-measurable subsets of $X = \mathbb{R}$. Theorem 4.9 implies that $\mathcal{S}$ is a σ-algebra and λ^* is a measure on $\mathcal{S}$. Moreover, Theorem 4.12 implies that $H \subset r(H) \subset \mathcal{S}$. Since $\mathcal{B}(\mathbb{R})$ is the smallest σ-algebra which contains all sets from H, we have $\mathcal{B}(\mathbb{R}) \subset \mathcal{S}$. Hence, $H \subset r(H) \subset \mathcal{B}(\mathbb{R}) \subset \mathcal{S}$. We also remark that λ^* is the extension of λ from $r(H)$ to $\mathcal{S}$ by Theorem 4.12.

Definition 4.13

- *Sets from $\mathcal{S}$ are called Lebesgue measurable sets.*
- *The measure λ^* defined on $\mathcal{S}$ is called the Lebesgue measure on $\mathbb{R}$.*

Remark 4.14 *The extension of λ to $\mathcal{B}(\mathbb{R})$ is unique.*

Remark 4.15 *We can define the Lebesgue measure on $\mathbb{R}^d$ by taking*

$$H = \{(a_1, b_1] \times \cdots \times (a_d, b_d] : a_k < b_k, \; k = 1, \ldots, d\} \cup \{\emptyset\}$$

and

$$\lambda(\emptyset) = 0, \quad \lambda\big((a_1, b_1] \times \cdots \times (a_d, b_d]\big) = \prod_{k=1}^{d} (b_k - a_k).$$

Example 4.16

1. *Let $x \in \mathbb{R}$. Then since $\{x\} = \bigcap_{n=1}^{\infty} \left(x - \frac{1}{n}, x\right]$, by Theorem 3.10 we have*

$$\lambda\big(\{x\}\big) = \lim_{n\to\infty} \lambda\big(\big(x - \tfrac{1}{n}, x\big]\big) = \lim_{n\to\infty} \frac{1}{n} = 0.$$

2. $\lambda(\mathbb{Q}) = \lambda\big(\bigcup_{n=1}^{\infty} \{r_n\}\big) = \sum_{n=1}^{\infty} \lambda\big(\{r_n\}\big) = \sum_{n=1}^{\infty} 0 = 0$, *where $\mathbb{Q} = \{r_1, r_2, \ldots\}$ is the set of rational numbers.*

5 Measurable Functions (Lecture Notes)

5.1 Motivation of the Definition, Introduction to Lebesgue Integrals

Let $f : X \mapsto \mathbb{R}$ be a function, where $X = [0,1]$. Let us recall the definition of the Riemann integral. We define the Riemann sums

$$S_n = \sum_{k=1}^{n} f(\xi_k)\Delta x_k,\ \Delta x_k = x_k - x_{k-1},$$

and say that f is Riemann integrable if the limit

$$\lim_{|\Delta x|\to 0} \sum_{k=1}^{n} f(\xi_k)\Delta x_k,\ |\Delta x| := \max_k |\Delta x_k|$$

exists and does not depend on the choice of $\{\xi_k\}$. This limit is called the Riemann integral of f and is denoted by

$$\int_0^1 f(x)\,dx.$$

Example 5.1

1. *If f is a continuous function, then f is Riemann integrable.*
2. *Take*

$$f(x) = \begin{cases} 1,\ x \in \mathbb{Q} \cap [0,1], \\ 0,\ x \in [0,1] \setminus \mathbb{Q}. \end{cases}$$

This function is not Riemann integrable since the limit depends on the choice of $\{\xi_k\}$. Indeed, if $\xi_k \in [x_{k-1}, x_k]$, $k = 1, \dots, n$ are rational, then $\sum_{k=1}^n f(\xi_k)\Delta x_k = \sum_{k=1}^n 1\,\Delta x_k = 1$, but if they are irrational, then $\sum_{k=1}^n f(\xi_k)\Delta x_k = \sum_{k=1}^n 0\,\Delta x_k = 0$.

Let us consider another approach to defining the integral.

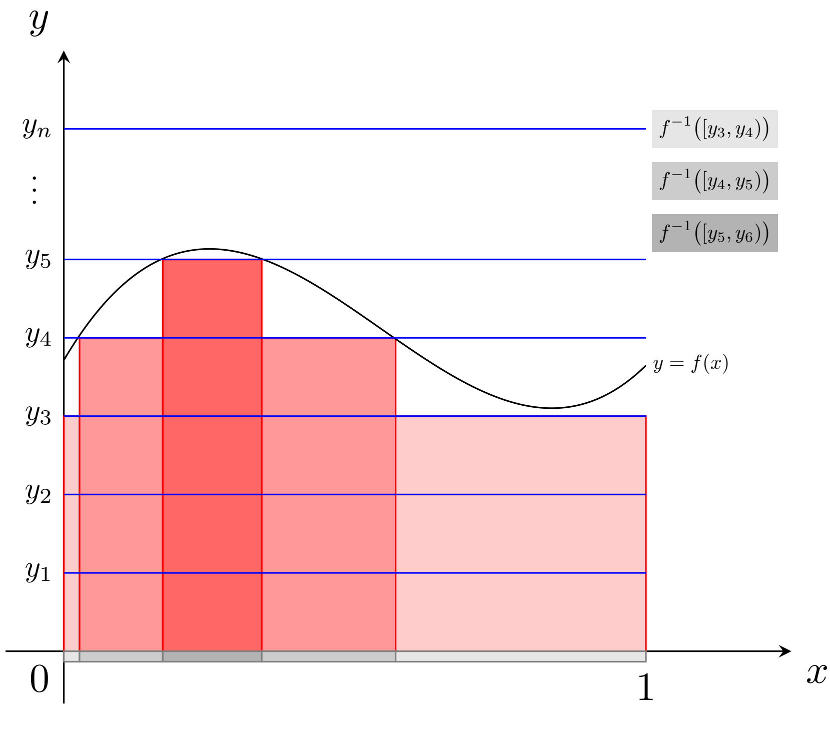

Figure 5.1

We can define the integral as

$$\lim_{|\Delta y| \to 0} \sum_{k=1}^{n} y_k \lambda\big(f^{-1}\big([y_k, y_{k+1})\big)\big) = \int\limits_a^b f(x)\,\lambda(dx).$$

Remark that $\int_a^b f(x)\,dx = \int_a^b f(x)\,\lambda(dx)$ if f is continuous, but this new definition is better.

Example 5.2 *We take*

$$f(x) = \begin{cases} 1,\ x \in \mathbb{Q} \cap [0,1], \\ 0,\ x \in [0,1] \setminus \mathbb{Q}, \end{cases}$$

and $y_k = \dfrac{k}{n}$. *Note that*

$$\begin{aligned} f^{-1}\big(\big[0, \tfrac{1}{n}\big)\big) &= [0,1] \setminus \mathbb{Q}, \\ f^{-1}\big(\big[\tfrac{n}{n}, \tfrac{n+1}{n}\big)\big) &= \mathbb{Q} \cap [0,1], \\ f^{-1}\big(\big[\tfrac{k}{n}, \tfrac{k+1}{n}\big)\big) &= \emptyset,\ k \neq 0, n. \end{aligned}$$

Hence

$$\sum_{k=0}^{n} \frac{k}{n} \lambda\big(f^{-1}\big(\big[\tfrac{k}{n}, \tfrac{k+1}{n}\big)\big)\big) = 0 \cdot \lambda\big([0,1] \setminus \mathbb{Q}\big) + \frac{1}{n}\lambda(\emptyset) + \cdots + \frac{n-1}{n}\lambda(\emptyset) + \frac{n}{n}\lambda\big(\mathbb{Q} \cap [0,1]\big) = 0,$$

and consequently

$$\int\limits_0^1 f(x)\,dx = \lim_{n \to \infty} 0 = 0.$$

With this approach to defining the integral, we need to be sure that we can compute the Lebesgue measure of sets

$$A_k = f^{-1}\big([y_k, y_{k+1})\big),$$

that is, the sets A_k, $k = 1, \ldots, n$ have to be Lebesgue (or Borel) measurable sets.

Remark 5.3 *Not all subsets of* $\mathbb{R}^d$ *are Lebesgue measurable.*

Consider the Banach-Tarski paradox. Given a solid ball in 3-dimensional space, there exists a decomposition of the ball into a finite number of disjoint subsets that can be put back together in a different way to yield two identical copies of the original ball. The Banach-Tarski paradox is a strong mathematical fact. However, we do not have any contradictions here since the pieces are not Lebesgue measurable:

$$V(B) = \sum_{k=1}^{n} V(A_k) = 2V(B),$$

because $V(A_k)$ do not exist.

5.2 Definition of Measurable Functions

Let X, X' be some sets and let $f : X \mapsto X'$ be a map.

Definition 5.4

1. *For $A \subset X$ the set $f(A) = \{f(x) : x \in A\}$ is called the image of A.*
2. *For a set $A' \subset X'$ the set $f^{-1}(A') = \{x \in X : f(x) \in A'\}$ is called the preimage of A'.*

Exercise 5.5 *Show that*

1. $f^{-1}\left(\bigcup_{k=1}^{n} A'_k\right) = \bigcup_{k=1}^{n} f^{-1}(A'_k)$,
2. $f^{-1}\left(\bigcap_{k=1}^{n} A'_k\right) = \bigcap_{k=1}^{n} f^{-1}(A'_k)$,
3. $f^{-1}(B' \setminus A') = f^{-1}(B') \setminus f^{-1}(A')$,

where $A'_k \subset X'$, $B', A' \subset X'$ and $n \in \mathbb{N} \cup \{\infty\}$.
Solution to 1):

$$f^{-1}\left(\bigcup_{k=1}^{n} A'_k\right) = \left\{x : f(x) \in \bigcup_{k=1}^{n} A'_k\right\} = \bigcup_{k=1}^{n}\{x : f(x) \in A'_k\} = \bigcup_{k=1}^{n} f^{-1}(A'_k)$$

Definition 5.6 *If X is a set and $\mathcal{F}$ is a σ-algebra on X, then $(X, \mathcal{F})$ is called a measurable space.*

Definition 5.7

1. *Let $(X, \mathcal{F})$ and $(X', \mathcal{F}')$ be measurable spaces and take $f : X \mapsto X'$. The function f is called $(\mathcal{F}, \mathcal{F}')$-measurable if $f^{-1}(A') \in \mathcal{F}$, $\forall\, A' \in \mathcal{F}'$.*
2. *In the case $X' = \mathbb{R}$, $\mathcal{F}' = \mathcal{B}(\mathbb{R})$, f is called $\mathcal{F}$-measurable.*
3. *If additionally $X = \mathbb{R}$, $\mathcal{F} = \mathcal{B}(\mathbb{R})$, that is, $f : \mathbb{R} \mapsto \mathbb{R}$ and $f^{-1}(A') \in \mathcal{B}(\mathbb{R})$, $\forall\, A' \in \mathcal{B}(\mathbb{R})$, then f is called Borel measurable.*

Example 5.8 *If $X = [0, 1]$, $\mathcal{F} = \{\emptyset, X\}$ and $X' = \mathbb{R}$, $\mathcal{F}' = \mathcal{B}(\mathbb{R})$, then only constant functions are $\mathcal{F}$-measurable. Indeed, we know that $A' = \{y\} \in \mathcal{F}' = \mathcal{B}(\mathbb{R})$. So $f^{-1}(A') \in \mathcal{F}$ means that*

$$f^{-1}(\{y\}) = \{x : f(x) = y\} = \emptyset \,\text{or}\, [0, 1].$$

Therefore $f(x) = c$, $\forall\, x \in [0, 1]$, where c is a constant.

Example 5.9 *Take $X = X' = \mathbb{R}$ and $\mathcal{F} = \mathcal{F}' = \mathcal{B}(\mathbb{R})$ with $f(x) = x$. Then f is Borel measurable since if $A' \in \mathcal{B}(\mathbb{R})$, then $f^{-1}(A') = A' \in \mathcal{B}(\mathbb{R})$.*

Remark 5.10 *The definition of measurability is very similar to that of continuity. Indeed, f is continuous if and only if the preimage of every open set is an open set, and for measurability we require that the preimage of any measurable set is measurable.*

6 Properties of Measurable Functions (Lecture Notes)

6.1 One Condition of Measurability

Let $(X, \mathcal{F})$ and $(X, \mathcal{F}')$ be measurable spaces. We recall that f is $(\mathcal{F}, \mathcal{F}')$-measurable if

$$\forall\, A' \in \mathcal{F}',\ f^{-1}(A') = \{x \in X : f(x) \in A'\} \in \mathcal{F}. \tag{6.1}$$

In general, property (6.1) is complicated to check, since the class $\mathcal{F}'$ can be too large. Theorem 6.1 says that it is enough to check (6.1) only for some subclass of $\mathcal{F}'$ in the case $\mathcal{F}' = \sigma(H)$.

Theorem 6.1 *Let $(X, \mathcal{F})$ and $(X, \mathcal{F}')$ be measurable spaces, where $\mathcal{F}' = \sigma(H)$, $H \subset 2^{X'}$. A map $f : X \mapsto X'$ is $(\mathcal{F}, \mathcal{F}')$-measurable if and only if $\forall\, A' \in H$, $f^{-1}(A') \in \mathcal{F}$.*

Proof: In the forward direction, the statement follows from the definition of measurability since

$$A' \in H \Rightarrow A' \in \mathcal{F}' \Rightarrow f^{-1}(A') \in \mathcal{F}.$$

To prove the converse, we set $Q := \{A' \in \mathcal{F}' : f^{-1}(A') \in \mathcal{F}\}$. Then $H \subset Q \subset \mathcal{F}' = \sigma(H)$. Let us show that Q is a σ-algebra.

1. $\emptyset \in Q$ because $f^{-1}(\emptyset) = \emptyset \in \mathcal{F}$.

2. If $A'_1, A'_2, \dots \in Q$, then $f^{-1}(A'_k) \in \mathcal{F}$. Consider $\bigcup\limits_{k=1}^{\infty} A'_k = A'$. Then

$$f^{-1}(A') = f^{-1}\left(\bigcup_{k=1}^{\infty} A'_k\right) = \bigcup_{k=1}^{\infty} f^{-1}(A'_k) \in \mathcal{F},$$

   because $\mathcal{F}$ is a σ-algebra.

3. If $A', B' \in Q$, then

$$f^{-1}(B' \setminus A') = f^{-1}(B') \setminus f^{-1}(A') \in \mathcal{F} \Rightarrow B' \setminus A' \in Q.$$

Hence $\sigma(H) \subset Q \Rightarrow \mathcal{F}' = \sigma(H) = Q$. □

Corollary 6.2 *Given $f : X \mapsto \mathbb{R}$, the following statements are equivalent.*

1. *f is $\mathcal{F}$-measurable*
2. *$\forall\, a \in \mathbb{R}$, $f^{-1}\big((-\infty, a)\big) = \{x \in X : f(x) < a\} \in \mathcal{F}$*
3. *$\forall\, a \in \mathbb{R}$, $f^{-1}\big((-\infty, a]\big) = \{x \in X : f(x) \leqslant a\} \in \mathcal{F}$*
4. *$\forall\, a \in \mathbb{R}$, $f^{-1}\big((a, \infty)\big) = \{x \in X : f(x) > a\} \in \mathcal{F}$*
5. *$\forall\, a \in \mathbb{R}$, $f^{-1}\big([a, \infty)\big) = \{x \in X : f(x) \geqslant a\} \in \mathcal{F}$*

Proof: We will only show that 1) and 2) are equivalent. We remark that for $H := \{(-\infty, a), a \in \mathbb{R}\}$, we have $\mathcal{B}(\mathbb{R}) = \sigma(H)$. 2) implies $\forall\, A' \in H$, $f^{-1}(A') \in \mathcal{F}$. Hence, by Theorem 6.1, f is $\mathcal{F}$-measurable (i.e. f is $\big(\mathcal{F}, \mathcal{B}(\mathbb{R})\big)$-measurable) if and only if it satisfies 2). □

Application of Corollary 6.2 ($X = X' = \mathbb{R}$, $\mathcal{F} = \mathcal{F}' = \mathcal{B}(\mathbb{R})$)

1. Every monotone function $f : \mathbb{R} \mapsto \mathbb{R}$ is Borel measurable.

2. Every continuous function $f : \mathbb{R} \mapsto \mathbb{R}$ is Borel measurable.

Proof:

1. Let f increase monotonically. Note that $f^{-1}\big((-\infty, a)\big)$ is always an interval for all a, because f increase monotonically. This implies $f^{-1}\big((-\infty, a)\big) \in \mathcal{B}(\mathbb{R})$.

2. We know that f is continuous if and only if the preimage $f^{-1}(G)$ of every open set G in $\mathbb{R}$ is open. Consequently, $f^{-1}\big((-\infty, a)\big)$ is open. Since every open set is Borel measurable, f is a Borel measurable function.

$\square$

Corollary 6.3 *If $f : \mathbb{R}^d \mapsto \mathbb{R}^m$ is continuous, then f is Borel measurable.*

Proof: Let $H := \{G \subset \mathbb{R}^m : G \,\text{is open}\}$. Then $\mathcal{B}(\mathbb{R}^m) = \sigma(H)$. Take $G \in H$. Then $f^{-1}(G)$ is open in $\mathbb{R}^d$ because it is continuous. Hence $f^{-1}(G) \in \mathcal{B}(\mathbb{R}^d)$ and, by Theorem 6.1, f is Borel measurable. $\square$

Exercise 6.4 *Let $f_k : X \mapsto \mathbb{R}$, $k = 1, \ldots, m$ be $\mathcal{F}$-measurable functions. We consider the function*

$$f = (f_1, \ldots, f_m) : X \mapsto \mathbb{R}^m.$$

Show that f is $\mathcal{F}$-measurable, that is, $\forall\, A' \in \mathcal{B}(\mathbb{R}^m)$, $f^{-1}(A') \in \mathcal{F}$. Take

$$H = \{[a_1, b_1) \times \cdots \times [a_m, b_m) : a_k < b_k\}$$

and use Theorem 6.1.

6.2 Composition of Measurable Maps

Theorem 6.5 *Let $(X, \mathcal{F}), (X', \mathcal{F}'), (X'', \mathcal{F}'')$ be measurable spaces. Let $f : X \mapsto X'$ and $g : X' \mapsto X''$ be $(\mathcal{F}, \mathcal{F}')$-measurable and $(\mathcal{F}', \mathcal{F}'')$-measurable respectively. Then $f \circ g$ is $(\mathcal{F}, \mathcal{F}'')$-measurable.*

Proof: Take $A'' \in \mathcal{F}''$. Then $A' := g^{-1}(A'') = \{y \in X' : g(y) \in A''\}$ because g is $(\mathcal{F}', \mathcal{F}'')$-measurable. Then

$$(g \circ f)^{-1}(A'') = \{x \in X : g\big(f(x)\big) \in A''\} = \{x \in X : f(x) \in A'\} = f^{-1}(A') \in \mathcal{F},$$

where $g\big(f(x)\big) \in A'' \Leftrightarrow f(x) \in A'$. $\square$

Corollary 6.6 *Let $(X, \mathcal{F})$ be a measurable space, $f_k : X \mapsto \mathbb{R}$, $k = 1, \ldots, m$ $\mathcal{F}$-measurable functions, and $F : \mathbb{R}^m \mapsto \mathbb{R}$ a Borel measurable function. Then $F(f_1, \ldots, f_m) : X \mapsto \mathbb{R}$ is $\mathcal{F}$-measurable.*

6.3 Properties of Measurable Functions

Theorem 6.7 *Let $(X, \mathcal{F})$ be a measurable space and let $f_1, f_2 : X \mapsto \mathbb{R}$ be $\mathcal{F}$-measurable functions. Then*

- cf_1
- $f_1 \pm f_2$
- $f_1 \cdot f_2$
- $\dfrac{f_1}{f_2}$, $f_2(x) \neq 0$, $x \in X$
- $\min\{f_1, f_2\}$
- $\max\{f_1, f_2\}$

are $\mathcal{F}$-measurable.

Proof: The statements follows from Corollary 6.6. For example, $f_1 + f_2$ is $\mathcal{F}$-measurable since we can take $f_1 + f_2 = F(f_1, f_2)$, where $F(u, v) = u + v$ is Borel measurable as a continuous function. $\square$

Theorem 6.8 *Let $(X, \mathcal{F})$ be a measurable space and let $f_n : X \mapsto \mathbb{R}$, $n \geqslant 1$ be a sequence of $\mathcal{F}$-measurable functions. Then*

- $g_1(x) := \sup\limits_{n \geqslant 1} f_n(x)$
- $g_2(x) := \inf\limits_{n \geqslant 1} f_n(x)$
- $g_3(x) := \varlimsup\limits_{n \to \infty} f_n(x)$
- $g_4(x) := \varliminf\limits_{n \to \infty} f_n(x)$

are $\mathcal{F}$-measurable. In particular, the function $f(x) := \lim\limits_{n \to \infty} f_n(x)$, $x \in X$, if the limit exists for all x, is also $\mathcal{F}$-measurable. The set $C := \big\{x \in X : \{f_n(x)\}_{n \geqslant 1} \, \text{converges in} \, \mathbb{R}\big\}$ belongs to $\mathcal{F}$.

Proof:

1. $\forall\, a \in \mathbb{R}$, $g_1^{-1}\big((-\infty, a]\big) = \{x : g_1(x) \leqslant a\} = \{x : \sup\limits_{n \geqslant 1} f_n(x) \leqslant a\} = \bigcap\limits_{n=1}^{\infty} \{f_n(x) \leqslant a\} \in \mathcal{F}$
2. $\forall\, a \in \mathbb{R}$, $g_2^{-1}\big([a, \infty)\big) = \{x : g_2(x) \geqslant a\} = \{x : \inf\limits_{n \geqslant 1} f_n(x) \geqslant a\} = \bigcap\limits_{n=1}^{\infty} \{f_n(x) \geqslant a\} \in \mathcal{F}$
3. $g_3(x) = \inf_{n \geqslant 1} \sup_{k \geqslant n} f_k(x)$ is $\mathcal{F}$-measurable because $\sup_{k \geqslant n} f_k(x)$ is $\mathcal{F}$-measurable by 1) and thus $\inf_{n \geqslant 1} \sup_{k \geqslant n} f_k(x)$ is $\mathcal{F}$-measurable by 2)
4. Similarly $g_4(x) = \sup\limits_{n \geqslant 1} \inf\limits_{k \geqslant n} f_k(x)$

Finally

$$C = \{x : g_3(x) = g_4(x)\} = \{x : g_3(x) - g_4(x) = 0\} = \big(g_3(x) - g_4(x)\big)^{-1}(\{0\}) \in \mathcal{F}$$

because $\{0\} \in \mathcal{B}(\mathbb{R})$. $\square$

7 Lebesgue Integrals (Lecture Notes)

7.1 Approximation by Simple Functions

Let $(X, \mathcal{F})$ be a measurable space and let λ be a measure on $\mathcal{F}$.

Definition 7.1 *A function $f : X \mapsto \mathbb{R}$ is called simple if the set $f(X)$ consists of a finite number of elements, that is, there exists distinct $a_1, \dots, a_m \in \mathbb{R}$ such that*

$$f(x) = \sum_{k=1}^{m} a_k \mathbb{I}_{A_k}(x), \tag{7.1}$$

where $A_k = \{x \in X : f(x) = a_k\} = f^{-1}(\{a_k\})$ and

$$\mathbb{I}_{A_k}(x) = \begin{cases} 0, \ x \notin A_k, \\ 1, \ x \in A_k. \end{cases}$$

Remark 7.2 *The sets $A_1, \dots, A_m \in \mathcal{F}$ if and only if the function f is measurable.*

Exercise 7.3 *Prove that the sum and product of two simple functions are simple functions.*

Theorem 7.4 *Let f be a nonnegative function. The function f is $\mathcal{F}$-measurable if and only if there exists a sequence $\{f_n\}_{n \geqslant 1}$ of simple $\mathcal{F}$-measurable functions such that $\forall\, x \in X$, $n \geqslant 1$, $f_n(x) \leqslant f_{n+1}(x)$ and $f(x) = \lim\limits_{n \to \infty} f_n(x)$.*

Proof: f is $\mathcal{F}$-measurable as a limit of $\mathcal{F}$-measurable functions by Theorem 6.8. Conversely, take an $\mathcal{F}$-measurable function f. For $n \in \mathbb{N}$ we consider numbers $\frac{k}{2^n}$, $k = 0, \dots, n2^n - 1$ and define

$$A_n^k := \left\{x \in X : \tfrac{k}{2^n} \leqslant f(x) \leqslant \tfrac{k+1}{2^n}\right\} \in \mathcal{F},$$

$$B_n := \{x \in X : f(x) \geqslant n\} \in \mathcal{F}.$$

Remark that $A_n^k = f^{-1}\left(\left[\frac{k}{2^n}, \frac{k+1}{2^n}\right)\right)$ and $B_n = f^{-1}\big([n, \infty)\big)$. Now take

$$f_n(x) = \sum_{k=0}^{n2^n - 1} \frac{k}{2^n} \mathbb{I}_{A_n^k}(x) + n \mathbb{I}_{B_n}(x).$$

$\square$

7.2 Definition of the Integral

Definition 7.5

I. Let f be a nonnegative $\mathcal{F}$-measurable simple function defined by (7.1) and take $A \in \mathcal{F}$. The value

$$\int_A f \, d\lambda := \int_A f(x) \lambda(dx) = \sum_{k=1}^{m} a_k \lambda(A \cap A_k)$$

is called the Lebesgue integral of f over A. We assume $a_k \lambda(A \cap A_k) = 0$ if $a_k = 0$, $\lambda(A \cap A_k) = \infty$.

II. Take $A \in \mathcal{F}$ and let $f : X \mapsto \mathbb{R}$ be a nonnegative $\mathcal{F}$-measurable function. The value

$$\int_A f\, d\lambda := \int_A f(x)\lambda(dx) := \sup_{p\in K(f)} \int_A p(x)\lambda(dx),$$

is called the Lebesgue integral of f over A, where $K(f)$ is the set of all simple functions $p : X \mapsto \mathbb{R}$ such that $0 \leqslant p(x) \leqslant f(x)$, $x \in X$.

Remark 7.6 (Alternate Definition to II) *Let $f \geqslant 0$ be $\mathcal{F}$-measurable and take $A \in \mathcal{F}$. Let $\{f_n\}$ be as described in Theorem 7.4. Then*

$$\int_A f(x)\lambda(dx) := \lim_{n\to\infty} \int_A f_n(x)\lambda(dx).$$

These two approaches define the same object.

Let $f : X \mapsto \mathbb{R}$ be any function. We consider its parts

$$f_+(x) = \max\{f(x), 0\},\ x \in X, \quad f_-(x) = -\min\{f(x), 0\},\ x \in X.$$

Then trivially

$$f(x) = f_+(x) - f_-(x),\ x \in X, \quad |f(x)| = f_+(x) + f_-(x),\ x \in X.$$

III. Take $A \in \mathcal{F}$ and let $f : X \mapsto \mathbb{R}$ be an $\mathcal{F}$-measurable function. If one of the integrals

$$\int_A f_+\, d\lambda, \quad \int_A f_-\, d\lambda \tag{7.2}$$

is finite, then

$$\int_A f\, d\lambda := \int_A f(x)\, dx := \int_A f(x)\lambda(dx) := \int_A f_+\, d\lambda - \int_A f_-\, d\lambda$$

is called the Lebesgue integral of f over A.

- *If both integrals in (7.2) are finite, then the function f is called Lebesgue integrable on A.*
- *The class of all Lebesgue integrable functions on A is denoted by $L(A, \lambda)$.*

8 Properties of Lebesgue Integrals (Lecture Notes)

8.1 Basic Properties

We assume that $f, g : X \mapsto \mathbb{R}$ are $\mathcal{F}$-measurable functions and $A \in \mathcal{F}$.

1. If $\lambda(A) = 0$, then $\int\limits_A f \, d\lambda = 0$.

2. If $\lambda(A) < \infty$ and $f(x) = c$, $x \in A$, then $f \in L(A, \lambda)$ and $\int\limits_A c \, d\lambda = c\lambda(A)$.

3. Let $0 \leqslant f(x) \leqslant g(x)$, $x \in A$. If $g \in L(A, \lambda)$, then $f \in L(A, \lambda)$ and $\int\limits_A f \, d\lambda \leqslant \int\limits_A g \, d\lambda$.

   Proof: This follows from Definition 7.5 II and the fact that $K(f) \subset K(g)$. So

$$\sup_{p \in K(f)} \int\limits_A p \, d\lambda \leqslant \sup_{p \in K(g)} \int\limits_A p \, d\lambda < \infty.$$

□

4. If $A \neq \emptyset$, $\lambda(A) < \infty$ and f is bounded on A, then $f \in L(A, \lambda)$ and

$$\inf_A f \cdot \lambda(A) \leqslant \int\limits_A f \, d\lambda \leqslant \sup_A f \cdot \lambda(A).$$

5. If $f \in L(A, \lambda)$, $c \in \mathbb{R}$, then $cf \in L(A, \lambda)$ and $\int\limits_A cf \, d\lambda = c \int\limits_A f \, d\lambda$.

6. If $f, g \in L(A, \lambda)$ and $f(x) \leqslant g(x)$, $\forall\, x \in A$, then $\int\limits_A f \, d\lambda \leqslant \int\limits_A g \, d\lambda$.

7. If $A, B \in \mathcal{F}$, $B \subset A$ and $f \in L(A, \lambda)$, then $f \in L(B, \lambda)$. If additionally $f \geqslant 0$, then

$$\int\limits_B f \, d\lambda \leqslant \int\limits_A f \, d\lambda.$$

8. If $A, B \in \mathcal{F}$, $A \cap B = \emptyset$ and $f \in L(A, \lambda)$, $f \in L(B, \lambda)$, then $f \in L(A \cup B, \lambda)$ and

$$\int\limits_{A \cup B} f \, d\lambda = \int\limits_A f \, d\lambda + \int\limits_B f \, d\lambda.$$

9. $f \in L(A, \lambda)$ if and only if $|f| \in L(A, \lambda)$.

   Proof: We write $f = f_+ - f_-$, $|f| = f_+ + f_-$. Remark that $f \in L(A, \lambda)$ if and only if

$$\int\limits_A f_+ \, d\lambda < \infty, \quad \int\limits_A f_- \, d\lambda < \infty.$$

Consider the sets

$$A_- := \{x \in A : f(x) < 0\} \in \mathcal{F}, \quad A_+ := \{x \in A : f(x) \geqslant 0\} \in \mathcal{F}.$$

Then $A_- \cap A_+ = \emptyset$. Hence

$$\int\limits_A |f|\, d\lambda = \int\limits_{A_-} |f|\, d\lambda + \int\limits_{A_+} |f|\, d\lambda = \int\limits_{A_-} f_-\, d\lambda + \int\limits_{A_+} f_+\, d\lambda \leqslant \int\limits_A f_-\, d\lambda + \int\limits_A f_+\, d\lambda < \infty.$$

This implies $|f| \in L(A, \lambda)$. Now assume $|f| \in L(A, \lambda)$. Since on A we have $0 \leqslant f_- \leqslant |f|$ and $0 \leqslant f_+ \leqslant |f|$, by 3) we also have

$$\int\limits_A f_-\, d\lambda < \infty, \quad \int\limits_A f_+\, d\lambda < \infty.$$

□

10. If $f \in L(A, \lambda)$ and $|g(x)| \leqslant f(x)$, $\forall\, x \in A$, then $g \in L(A, \lambda)$ and $\left|\int\limits_A g\, d\lambda\right| \leqslant \int\limits_A |f|\, d\lambda$.
11. If $f, g \in L(A, \lambda)$, then $f + g \in L(A, \lambda)$ and $\int\limits_A (f + g)\, d\lambda = \int\limits_A f\, d\lambda + \int\limits_A g\, d\lambda$.
12. If $f \in L(X, \lambda)$, then the function
$$\mu(A) := \int\limits_A f\, d\lambda,\ A \in \mathcal{F}$$
is σ-additive. In particular, if $f \geqslant 0$, then μ is a measure on $\mathcal{F}$.

Definition 8.1 *We say that $f = g\ \lambda-$a. e. (almost everywhere) on A if $\lambda\big(\{x \in A : f(x) \neq g(x)\}\big) = 0$.*

Example 8.2 *The functions $f(x) = \mathbb{I}_{\mathbb{Q}}(x)$, $x \in \mathbb{R}$ and $g(x) = 0$, $x \in \mathbb{R}$ are equal $\lambda-$a. e.*

Remark that the set $\{x \in A : f(x) \neq g(x)\} \in \mathcal{F}$ since

$$\{x \in A : f(x) \neq g(x)\} = \{x \in A : f(x) - g(x) \neq 0\} = (f - g)^{-1}(\mathbb{R} \setminus \{0\}) \in \mathcal{F}$$

and $f - g$ is $\mathcal{F}$-measurable as the difference of two measurable functions.

13. If $g = f\ \lambda-$a. e. on A and $f \in L(A, \lambda)$, then $g \in L(A, \lambda)$ and $\int\limits_A f\, d\lambda = \int\limits_A g\, d\lambda$.
14. If $f \in L(A, \lambda)$, $f \geqslant 0$ and $\int\limits_A f\, d\lambda = 0$, then $f = 0\ \lambda-$a. e. on A.

8.2 Convergence of Functions

Definition 8.3 *Let $f, f_n : X \mapsto \mathbb{R}$, $n \geqslant 1$ be $\mathcal{F}$-measurable functions. The sequence $\{f_n\}_{n \geqslant 1}$ converges to $f\ \lambda-$a. e. (a. e. with respect to λ) if there exists $\Phi \in \mathcal{F}$, $\lambda(\Phi) = 0$ such that*

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

In this case we write $f_n \to f\ \lambda-$a. e.

Exercise 8.4 *Let $f_n \to f\ \lambda-$a. e. and $f_n \to g\ \lambda-$a. e. Show that $f = g\ \lambda-$a. e.*

9 Limit Theorems for Lebesgue Integrals (Lecture Notes)

9.1 Convergence of Functions

Let $(X, \mathcal{F})$ be a fixed measurable space and let λ be a measure on $\mathcal{F}$. We take functions f, f_n, $n \geqslant 1$ from X to $\mathbb{R}$ that are $\mathcal{F}$-measurable.

Definition 9.1 *The sequence $\{f_n\}_{n\geqslant 1}$ converges to f $\lambda-$a. e. if there exists $\Phi \in \mathcal{F}$, $\lambda(\Phi) = 0$ such that*

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

In this case we write $f_n \to f$ $\lambda-$a. e.

Exercise 9.2 *Let $f_n \to f$ $\lambda-$a. e. and $f_n \to g$ $\lambda-$a. e. Show that $f = g$ $\lambda-$a. e.*

Definition 9.3 *The sequence $\{f_n\}_{n\geqslant 1}$ converges to f in measure if*

$$\forall\, \epsilon > 0,\ \lambda\big(\{x \in X : |f_n(x) - f(x)| \geqslant \epsilon\}\big) \to 0,\ n \to \infty$$

In this case we write $f_n \xrightarrow{\lambda} f$.

Theorem 9.4 *If $f_n \xrightarrow{\lambda} f$ and $f_n \xrightarrow{\lambda} g$, then $f = g$ $\lambda-$a. e.*

Proof: We first remark that

$$\{x : |f(x) - f_n(x) + f_n(x) - g(x)| \geqslant \epsilon\} \subset \big\{x : |f(x) - f_n(x)| \geqslant \tfrac{\epsilon}{2}\big\} \cup \big\{x : |f_n(x) - g(x)| \geqslant \tfrac{\epsilon}{2}\big\}.$$

Hence $\forall\, \epsilon > 0$

$$\begin{aligned}\lambda\big(\{x : |f(x) - g(x)| \geqslant \epsilon\}\big) &= \lambda\big(\{x : f(x) - f_n(x) + f_n(x) - g(x)| \geqslant \epsilon\}\big)\\ &\leqslant \lambda\big(\big\{x : |f(x) - f_n(x)| \geqslant \tfrac{\epsilon}{2}\big\}\big) + \lambda\big(\big\{x : |f_n(x) - g(x)| \geqslant \tfrac{\epsilon}{2}\big\}\big) \to 0,\ n \to \infty\end{aligned}$$

Thus $\forall\, \epsilon > 0$, $\lambda\big(\{x : |f(x) - g(x)| \geqslant \epsilon\}\big) = 0$. Next

$$\{x : f(x) \neq g(x)\} = \bigcup_{k=1}^{\infty} \big\{x : |f(x) - g(x)| \geqslant \tfrac{1}{k}\big\}.$$

By the σ-semiadditivity of the measure λ, $\lambda\big(\{x : f(x) \neq g(x)\}\big) = 0$. □

Note that convergence in measure does not imply convergence $\lambda-$a. e. It does not even imply convergence for some fixed point x. Likewise, convergence $\lambda-$a. e. does not imply convergence in measure.

Example 9.5 *Take the following sequence with the Lebesgue measure as λ, and $X = [0, 1]$, $\mathcal{F} = \mathcal{B}\big([0, 1]\big)$.*

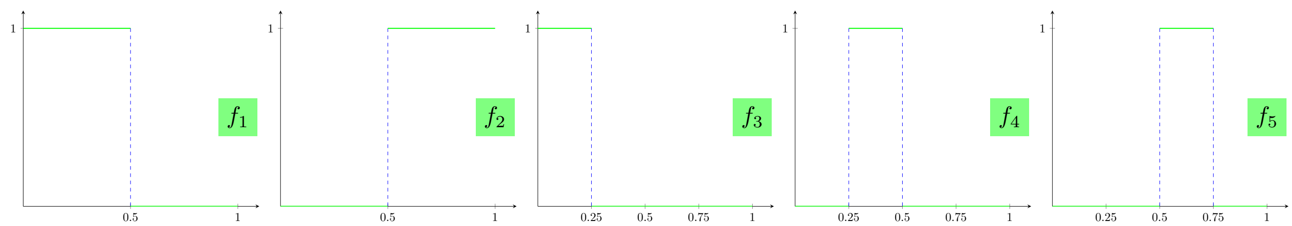

Figure 9.1

Then $f_n \xrightarrow{\lambda} f$ but $f_n \not\to f$ $\lambda-$a. e. Moreover, $\forall\, x \in [0, 1]$, $f_n(x) \not\to f(x)$.

Example 9.6 *We take* $X = \mathbb{R}$, $\mathcal{F} = \mathcal{B}(\mathbb{R})$ *and the Lebesgue measure as* λ. *Take* $f_n(x) = \mathbb{I}_{[n,\infty)}(x)$, $x \in \mathbb{R}$.

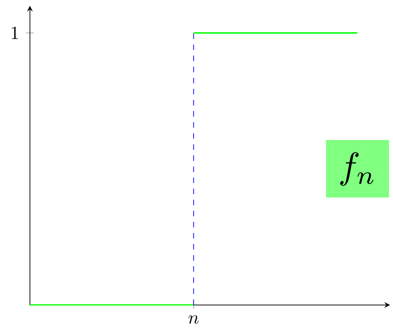

Figure 9.2

Then $\forall\, x \in \mathbb{R}$, $f_n(x) \to 0 \Rightarrow f \to 0\; \lambda-$a. e. *But* $f_n \stackrel{\lambda}{\not\to} f$ *since for* $\epsilon < 1$

$$\lambda\big(\{x : |f_n(x) - f(x)| > \epsilon\}\big) = \lambda\big([n,\infty)\big) = \infty.$$

Theorem 9.7 (Lebesgue) *If* $\lambda(X) < \infty$, *then* $f_n \to f\; \lambda-$a. e. $\Rightarrow f_n \xrightarrow{\lambda} f$.

Proof: Let $\epsilon > 0$ be fixed. We set

$$A_n := \{x : |f_n(x) - f(x)| \geqslant \epsilon\} \in \mathcal{F}, \quad B_n := \bigcup_{k=n}^{\infty} A_k \in \mathcal{F}.$$

We remark that B_n, $n \geqslant 1$ decreases. Set

$$B := \bigcap_{n=1}^{\infty} B_n = \overline{\lim_{n\to\infty}} A_n = \{x : \text{for an infinite number of indices}\, n,\ |f_n(x) - f(x)| > \epsilon\}.$$

Then $B \subset \{x : f_n(x) \not\to f(x)\}$ and consequently $\lambda(B) = 0$ by the convergence $f_n \to f\; \lambda-$a. e. Moreover $\lambda(B_1) \leqslant \lambda(X) < \infty$. Then by the continuity of the measure λ, $0 = \lambda(B) = \lim_{n\to\infty} \lambda(B_n)$. So

$$\lim_{n\to\infty} \lambda(A_n) \leqslant \lim_{n\to\infty} \lambda(B_n) = 0.$$

$\square$

Theorem 9.8 (Riesz) *If* $f_n \xrightarrow{\lambda} f$, *then there exists a subsequence* $\{f_{n_k}\}_{k\geqslant 1}$ *such that* $f_{n_k} \to f\; \lambda-$a. e.

Theorem 9.9 (Subsequence Criterion) *Let* $\lambda(X) < \infty$. *Then* $f_n \xrightarrow{\lambda} f$ *if and only if every subsequence* $\{f_{n_k}\}_{k\geqslant 1}$ *has a subsubsequence* $\{f_{n_{k_j}}\}_{j\geqslant 1}$ *such that* $f_{n_{k_j}} \to f\; \lambda-$a. e.

9.2 Monotone Convergence Theorem

Theorem 9.10 (Monotone Convergence Theorem) *Let* $A \in \mathcal{F}$, f, f_n, $n \geqslant 1$ *satisfy*

1. $0 \leqslant f_n(x) \leqslant f_{n+1}(x)$, $\forall\, n \geqslant 1$, $x \in A$
2. $f_n \to f\; \lambda-$a. e. *on* A

Then

$$\lim_{n\to\infty} \int_A f_n \, d\lambda = \int_A f \, d\lambda.$$

Proof: We first remark that by the monotonicity of f_n, we have

$$\int_A f_1\, d\lambda \leqslant \int_A f_2\, d\lambda \leqslant \cdots \leqslant \int_A f_n\, d\lambda \leqslant \cdots \int_A f\, d\lambda. \tag{9.1}$$

Since we have an increasing sequence of numbers, there exists

$$\alpha := \lim_{n\to\infty} \int_A f_n\, d\lambda \leqslant \infty.$$

We can assume that $\alpha < \infty$. Otherwise the equality

$$\int_A f\, d\lambda = \lim_{n\to\infty} \int_A f_n\, d\lambda$$

trivially follows from (9.1). Since $\alpha < \infty$, then $\int_A f_n\, d\lambda < \infty$, $\forall\, n \geqslant 1$. We take a simple $\mathcal{F}$-measurable function $p \in K(f)$ and $c \in (0,1)$, and set $A_n := \{x \in A : f_n(x) \geqslant cp(x)\} \in \mathcal{F}$. We know that

$$1.\ A_n \subset A_{n+1}, \quad 2.\ \bigcup_{n=1}^{\infty} A_n = A.$$

Take $x \in A_n \Rightarrow f_n(x) \geqslant cp(x) \Rightarrow f_{n+1}(x) \geqslant f_n(x) \geqslant cp(x) \Rightarrow x \in A_{n+1}$. This proves 1). Now since $A_n \subset A$, we have $\bigcup_{n=1}^{\infty} A_n \subset A$. Next take $x \in A$. Remark that $cp(x) < p(x) \leqslant f(x)$. Since $f_n(x) \to f(x)$, there exists n such that $cp(x) \leqslant f_n(x) \leqslant f(x)$. This implies $x \in A_n$ and hence $A \subset \bigcup_{n=1}^{\infty} A_n$. This proves 2). By properties 3), 5), and 7) of the Lebesgue integral

$$\int_A f_n\, d\lambda \geqslant \int_{A_n} f_n\, d\lambda \geqslant c \int_{A_n} p\, d\lambda.$$

Hence

$$c \int_{A_n} p\, d\lambda \leqslant \int_A f_n\, d\lambda \leqslant \alpha.$$

By the σ-additivity of the integral and the continuity of the measure λ

$$c \int_A p\, d\lambda = \lim_{n\to\infty} c \int_{A_n} p\, d\lambda \leqslant \alpha.$$

Since $c \int_A p\, d\lambda \leqslant \alpha$, $\forall\, p \in K(f)$, we have

$$c \int_A f\, d\lambda = c \sup_{p\in K(f)} \int_A p\, d\lambda = \sup_{p\in K(f)} c \int_A p\, d\lambda \leqslant \alpha.$$

So $c \int_A f\, d\lambda \leqslant \alpha$, where $c \in (0,1)$. Sending $c \to 1^-$, we get

$$\int_A f\, d\lambda \leqslant \alpha.$$

□

10 Limit Theorems, Change of Variables (Lecture Notes)

10.1 Monotone Convergence Theorem

Let X be a fundamental set with $\mathcal{F}$ being a σ-algebra on X. We take λ as a measure on X. We recall that

1. $f_n \to f\ \lambda\text{–a.e.} \Leftrightarrow \exists\,\Phi \in \mathcal{F},\ \lambda(\Phi) = 0 : \lim\limits_{n\to\infty} f_n(x) = f(x),\ \forall\, x \in X \setminus \Phi$

2. $f_n \xrightarrow{\lambda} f \Leftrightarrow \forall\,\epsilon > 0,\ \lambda\big(\{x \in X : |f_n(x) - f(x)| \geqslant \epsilon\}\big) \to 0,\ n \to \infty$

We also recall the Monotone Convergence Theorem.

Theorem 9.10 (Monotone Convergence Theorem) *Let $A \in \mathcal{F}$ and $f, f_n,\ n \geqslant 1$ satisfy*

$$0 \leqslant f_n(x) \leqslant f_{n+1}(x),\ \forall\, n \geqslant 1,\ x \in A, \quad f_n(x) \to f(x)\ \lambda\text{–a.e. on } A.$$

Then

$$\lim_{n\to\infty} \int\limits_A f_n \, d\lambda = \int\limits_A f \, d\lambda.$$

10.2 Fatou's Lemma

Lemma 10.1 (Fatou's Lemma) *Let $A \in \mathcal{F}$ and functions $f_n,\ n \geqslant 1$ satisfy $f_n(x) \geqslant 0,\ \forall\, x \in A$. Then*

$$\int\limits_A \varliminf_{n\to\infty} f_n(x)\lambda(dx) \leqslant \varliminf_{n\to\infty} \int\limits_A f_n \, d\lambda.$$

Proof: Consider $g_n(x) := \inf\limits_{k\geqslant n} f_k(x),\ x \in A,\ n \geqslant 1$. Then $0 \leqslant g_n(x) \leqslant g_{n+1}(x),\ \forall\, x \in A,\ n \geqslant 1$. Moreover

$$\lim_{n\to\infty} g_n(x) = \lim_{n\to\infty} \inf_{k\geqslant n} f_k(x) = \varliminf_{n\to\infty} f_n(x).$$

We also have $g_n(x) \leqslant f_n(x),\ \forall\, x \in A,\ n \geqslant 1$. Thus

$$\int\limits_A g_n(x) \leqslant \int\limits_A f_n(x).$$

By Theorem 9.10

$$\lim_{n\to\infty} \int\limits_A g_n \, d\lambda = \int\limits_A \lim_{n\to\infty} g_n \, d\lambda = \int\limits_A \varliminf_{n\to\infty} f_n \, d\lambda.$$

Hence

$$\varliminf_{n\to\infty} \int\limits_A f_n \, d\lambda \geqslant \varliminf_{n\to\infty} \int\limits_A g_n \, d\lambda = \int\limits_A \varliminf_{n\to\infty} f_n \, d\lambda.$$

□

Remark 10.2 *Fatou's lemma implies that if $f_n \geqslant 0$ on A, $f_n \to f\ \lambda\text{–a.e.}$ on A, and $\int_A f_n \, d\lambda \leqslant C$ for all $n \geqslant 1$, then $f \in L(A, \lambda)$ and $\int_A f \, d\lambda \leqslant C$.*

To see this, just apply Fatou's lemma for the set $A \setminus \Phi$, where $\Phi = \{x : f_n(x) \not\to f(x)\}$, and use properties 1) and 8) of the integral.

10.3 The Dominated Convergence Theorem

Theorem 10.3 (Dominated Convergence Theorem) *Let $A \in \mathcal{F}$ and a sequence f_n satisfy*

1. $f_n \to f$ $\lambda-$a.e. *on* A
2. $\exists\, g \in L(A, \lambda) : |f_n(x)| \leqslant g(x), \forall\, x \in A,\, n \geqslant 1$

Then $f, f_n \in L(A, \lambda)$, $n \geqslant 1$ *and*

$$\lim_{n\to\infty} \int\limits_A f_n \, d\lambda = \int\limits_A f \, d\lambda.$$

Proof: We remark that $-g(x) \leqslant f_n(x) \leqslant g(x)$, $\forall\, x \in A$, $n \geqslant 1$. Then $g + f_n \geqslant 0$ and $g - f_n \geqslant 0$, $\forall\, n \geqslant 1$. We can apply Fatou's lemma:

$$\varliminf_{n\to\infty} \int\limits_A (g + f_n) \, d\lambda \geqslant \int\limits_A (g + f) \, d\lambda,$$

$$\varliminf_{n\to\infty} \int\limits_A (g - f_n) \, d\lambda \geqslant \int\limits_A (g - f) \, d\lambda.$$

Hence

$$\int\limits_A g \, d\lambda + \varliminf_{n\to\infty} \int\limits_A f_n \, d\lambda \geqslant \int\limits_A g \, d\lambda + \int\limits_A f \, d\lambda$$

and

$$\int\limits_A g \, d\lambda - \varlimsup_{n\to\infty} \int\limits_A f_n \, d\lambda \geqslant \int\limits_A g \, d\lambda - \int\limits_A f \, d\lambda.$$

Hence

$$\int\limits_A f \, d\lambda \leqslant \varliminf_{n\to\infty} \int\limits_A f_n \, d\lambda \leqslant \varlimsup_{n\to\infty} \int\limits_A f_n \, d\lambda \leqslant \int\limits_A f \, d\lambda.$$

□

Corollary 10.4 *The clam of Theorem 10.3 remains true if condition 1) is replaced by*

$$1')\; f_n \xrightarrow{\lambda} f \text{ on } A \text{ i.e. } \forall\, \epsilon > 0,\; \lambda\big(\{x \in A : |f_n(x) - f(x)| \geqslant \epsilon\}\big) \to 0,\; n \to \infty.$$

Exercise 10.5 *Using Theorem 9.8, prove Corollary 10.4.*

10.4 Change of Variables

We consider two measurable spaces $(X, \mathcal{F})$ and $(X', \mathcal{F}')$. Let λ be a measure on X and let T be an $(\mathcal{F}, \mathcal{F}')$-measurable map. We define a new measure on X' which is a push forward of the measure λ:

$$T_{\#}\lambda(A') := \lambda\big(T^{-1}(A')\big) = \lambda\big(\{x \in X : T(x) \in A'\}\big), \forall\, A' \in \mathcal{F}'.$$

We will also use the notation $\lambda \circ T^{-1} := T_{\#}\lambda$.

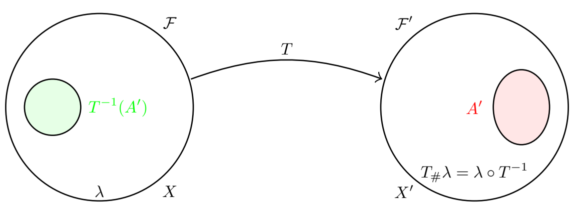

Figure 10.1

Exercise 10.6 *Check that $T_{\#}\lambda$ is a measure on $\mathcal{F}'$.*

Example 10.7 *Take $X = \mathbb{R}$ and $X' = [0, \infty)$. Let λ be the Lebesgue measure on $X = \mathbb{R}$. Take $T(x) = |x|$.*

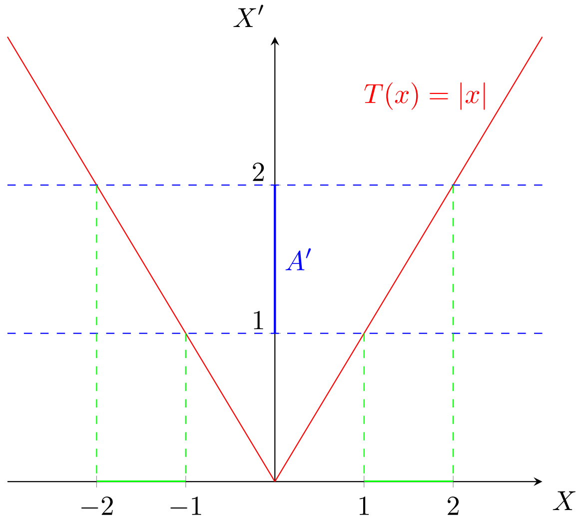

Figure 10.2

Then $\lambda \circ T^{-1}\big([1,2]\big) = \lambda\big([1,2] \cup [-2,-1]\big) = 2\lambda\big([1,2]\big) = 2$. It is easy to check that $\lambda \circ T^{-1}(A') = 2\lambda(A')$.

Theorem 10.8 (Change of Variables) *Let $f : X' \mapsto \mathbb{R}$ be $\mathcal{F}'$-measurable. Then*

$$\int_X f\big(T(x)\big)\lambda(dx) = \int_{X'} f(y)(\lambda \circ T^{-1})(dy)$$

holds if at least one of the integrals exists.

10.5 Comparison of Lebesgue and Riemann Integrals

Take $X = [a, b]$ and $\mathcal{F} = \mathcal{B}\big([a, b]\big) = \mathcal{B}(\mathbb{R}) \cap [a, b]$, and let λ be the Lebesgue measure on $[a, b]$. We will denote the Lebesgue integral over λ as

$$\int_a^b f\, d\lambda := \int_a^b f(x)\, dx := \int_{[a,b]} f\, d\lambda.$$

We denote by $R\big([a, b]\big)$ the set of all Riemann integrable functions $f : [a, b] \mapsto \mathbb{R}$ on $[a, b]$.

Theorem 10.9 *If $f \in R\big([a, b]\big)$, then $f \in L\big([a, b], \lambda\big)$ and*

$$\int_a^b f(x)\, dx = \int_a^b f\, d\lambda.$$

10.6 Lebesgue-Stieltjes Integral

Take $X = \mathbb{R}$ and $H = \{(a, b] : -\infty < a < b < \infty\} \cup \{\emptyset\}$. Let $F : \mathbb{R} \mapsto \mathbb{R}$ be a non-decreasing and right continuous function. We set

$$\lambda_F(\emptyset) := 0, \quad \lambda_F\big((a, b]\big) = F(b) - F(a),\ (a, b] \in H.$$

Theorem 10.10 *λ_F is a measure on H.*

Consequently λ_F can be extended to a measure on $r(H)$. We will denote this extension also by λ_F. Next let λ_F^* be the outer measure on 2^X generated by λ_F. Consider the class S_F of all λ_F^*-measurable sets from 2^X. By Theorem 4.9 S_F is a σ-algebra and λ_F^* is a measure on S_F. We denote this measure by λ_F. Next, by Theorem 4.12, $H \subset r(H) \subset S_F$. We can conclude that $\mathcal{B}(\mathbb{R}) = \sigma(H) \subset S_F$. Hence λ_F is defined on $\mathcal{B}(\mathbb{R})$.

Definition 10.11 *The integral*

$$\int\limits_A f\,d\lambda_F$$

is called the Lebesgue-Stieltjes integral on $\mathbb{R}$ and is denoted by

$$\int\limits_A f(x)\,dF(x) := \int\limits_A f\,d\lambda_F.$$

If $A = [a, b]$, then we write

$$\int\limits_a^b f(x)\,dF(x).$$

Exercise 10.12

1. *Let F be a continuously differentiable function and $F'(x) = f(x)$, $x \in \mathbb{R}$. Show that*

$$\int\limits_{-\infty}^{\infty} g(x)\,dF(x) = \int\limits_{-\infty}^{\infty} g(x)f(x)\,dx.$$

2. *Let $x_1 < \dots < x_n$ and $m_1, \dots, m_n \geqslant 0$. Define*

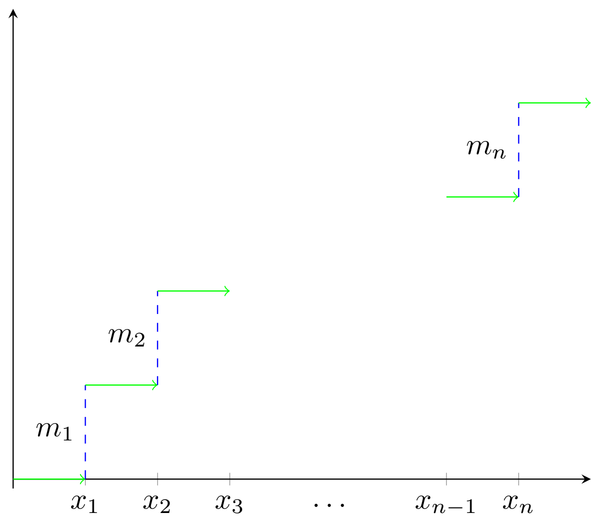

Figure 10.3

Show that

$$\int\limits_{-\infty}^{\infty} g\,dF = \sum_{k=1}^{n} g(x_k)m_k.$$

11 Metric Spaces (Lecture Notes)

11.1 Definition and Examples

Definition 11.1 *A metric space is a pair (X, d), where X is a set and d is a metric or distance function on X, that is, d is a function $d : X \times X \mapsto \mathbb{R}$ such that $\forall\, x, y, z \in X$*

(M1) $d(x, y) \geqslant 0$

(M2) $d(x, y) = 0 \Leftrightarrow x = y$

(M3) $d(x, y) = d(y, x)$ *(symmetry)*

(M4) $d(x, y) \leqslant d(x, z) + d(z, y)$ *(triangle inequality)*

Examples of Metric Spaces

1. Real line $\mathbb{R}$
   $X = \mathbb{R}$, $d(x, y) = |x - y|$, $x, y \in \mathbb{R}$

2. Euclidean space $\mathbb{R}^n$
   $X = \mathbb{R}^n$, $d(x, y) = \sqrt{\sum_{k=1}^{n}(\xi_k - \eta_k)^2}$, $x = (\xi_1, \ldots, \xi_n)$, $y = (\eta_1, \ldots, \eta_n)$

3. Sequence space l^∞
   $X = l^\infty := \{x = (\xi_k)_{k=1}^\infty : \xi_k \in \mathbb{R},\, x \text{ is bounded}\}$, $d(x, y) = \sup\limits_{k \in \mathbb{N}} |\xi_k - \eta_k|$, $x = (\xi_k)_{k=1}^\infty$, $y = (\eta_k)_{k=1}^\infty$

4. Space c
   $X = c = \{x = (\xi_k)_{k=1}^\infty : \xi_k \in \mathbb{R},\, x \text{ is convergent}\}$, $d(x, y) = \sup\limits_{k \in \mathbb{N}} |\xi_k - \eta_k|$
   We can say that c is a metric subspace of l^∞ because it is a subset of l^∞ and its metric is just a restriction of the metric on l^∞.

5. Space $B(A)$
   $X = B(A)$ is the set of all bounded functions $x : A \mapsto \mathbb{R}$, $A \subset \mathbb{R}$. We define

   $$d(x, y) = \sup_{t \in A} |x(t) - y(t)|,\ x, y \in B(A).$$

   Let us check that $\big(B(A), d\big)$ is a metric space.

   (M1) $d(x, y) \geqslant 0$ is trivial

   (M2) $d(x, y) = 0 \Leftrightarrow \sup\limits_{t \in A} |x(t) - y(t)| = 0 \Leftrightarrow x(t) = y(t),\ \forall\, t \in [a, b]$

   (M3) $d(x, y) = \sup\limits_{t \in A} |x(t) - y(t)| = \sup\limits_{t \in A} |y(t) - x(t)| = d(y, x)$

   (M4) $d(x, y) = \sup\limits_{t \in A} |x(t) - z(t) + z(t) - y(t)| \leqslant \sup\limits_{t \in A} |x(t) - z(t)| + \sup\limits_{t \in A} |z(t) - y(t)| = d(x, z) + d(z, y)$

6. Functional space $C[a, b]$
   $X = C[a, b]$ is the set of all continuous functions from $[a, b]$ to $\mathbb{R}$. We define

   $$d(x, y) = \max_{t \in [a, b]} |x(t) - y(t)|.$$

   Again we can say that $\big(C[a, b], d\big)$ is a metric subspace of $\big(B[a, b], d\big)$.

7. Space l^p, $p \geqslant 1$

   $X = l^p$ is the set of all sequences $x = (\xi_k)_{k=1}^{\infty}$ in $\mathbb{R}$ such that $\sum_{k=1}^{\infty} |\xi_k|^p < \infty$. We define

$$d(x,y) = \left(\sum_{k=1}^{\infty} |\xi_k - \eta_k|^p \right)^{\frac{1}{p}}, \tag{11.1}$$

   where $x = (\xi_k)_{k=1}^{\infty}$ and $y = (\eta_k)_{k=1}^{\infty}$. Let us check that (l^p, d) is a metric space. For this we need the following inequalities. We start from the Hölder inequality:

$$\sum_{k=1}^{\infty} |\xi_k \eta_k| \leqslant \left(\sum_{k=1}^{\infty} |\xi_k|^p \right)^{\frac{1}{p}} \left(\sum_{k=1}^{\infty} |\eta_k|^q \right)^{\frac{1}{q}}, \ \frac{1}{p} + \frac{1}{q} = 1, \ p > 1.$$

   In particular, if $p = 2$, then $q = 2$ and we have the Cauchy-Schwarz inequality:

$$\sum_{k=1}^{\infty} |\xi_k \eta_k| \leqslant \left(\sum_{k=1}^{\infty} |\xi_k|^2 \right)^{\frac{1}{2}} \left(\sum_{k=1}^{\infty} |\eta_k|^2 \right)^{\frac{1}{2}}.$$

   We also need the Minkowski inequality:

$$\left(\sum_{k=1}^{\infty} |\xi_k + \eta_k|^p \right)^{\frac{1}{p}} \leqslant \left(\sum_{k=1}^{\infty} |\xi_k|^p \right)^{\frac{1}{p}} + \left(\sum_{k=1}^{\infty} |\eta_k|^p \right)^{\frac{1}{p}}.$$

   Let us now show that d defined by (11.1) is a distance. Conditions (M1)-(M3) are trivial, so we will show (M4):

$$\begin{aligned} d(x,y) &= \left(\sum_{k=1}^{\infty} |\xi_k - \eta_k|^p \right)^{\frac{1}{p}} = \left(\sum_{k=1}^{\infty} |\xi_k - \zeta_k + \zeta_k - \eta_k|^p \right)^{\frac{1}{p}} \\ &\leqslant \left(\sum_{k=1}^{\infty} \left(|\xi_k - \zeta_k| + |\zeta_k - \eta_k| \right)^p \right)^{\frac{1}{p}} \\ &\leqslant \left(\sum_{k=1}^{\infty} |\xi_k - \zeta_k|^p \right)^{\frac{1}{p}} + \left(\sum_{k=1}^{\infty} |\zeta_k - \eta_k|^p \right)^{\frac{1}{p}} \\ &= d(x,z) + d(z,y), \end{aligned}$$

   where $z = (\zeta_k)_{k=1}^{\infty}$, and x and y are as before.

8. Space l_n^p, $p \geqslant 1$

   $X = l_n^p = \mathbb{R}^n$, $d(x,y) = \left(\sum_{k=1}^{n} |\xi_k - \eta_k|^p \right)^{\frac{1}{p}}$

9. Space $L^p[a,b]$, $p \geqslant 1$

   Let λ be the Lebesgue measure on $[a,b]$. We assume that two measurable functions $x, y : [a,b] \mapsto \mathbb{R}$ are equal to each other if $x = y$ λ−a. e. Then $X = L^p[a,b]$ is the space of all measurable functions x on $[a,b]$ (more precisely classes of equivalence) such that $\int_a^b |x(t)|^p \, dt < \infty$. We define

$$d(x,y) = \left(\int_a^b |x(t) - y(t)|^p \, dt \right)^{\frac{1}{p}}.$$

10. Discrete metric space

Let X be a set. We define

$$d(x,y)=\begin{cases}0, \; x=y\\ 1, \; x\neq y.\end{cases}$$

(X,d) is called a discrete metric space.

11.2 Open and Closed Sets

Let (X,d) be a metric space.

Definition 11.2

- $B_r(x_0)=\{x\in X : d(x,x_0)<r\}$ *is called an open ball with center* x_0 *and radius* r
- $\overline{B}_r(x_0)=\{x\in X : d(x,x_0)\leqslant r\}$ *is called a closed ball with center* x_0 *and radius* r

Definition 11.3

- *A set* G *is called open (in* X*) if* $\forall\, x\in G,\ \exists\, r>0 : B_r(x)\subset G$.
- *A set* F *is called closed (in* X*) if* $F^c=X\setminus F$ *is open.*

Exercise 11.4

1. *Prove that the union of any family of open sets is open.*
2. *Prove that the intersection of any finite family of open sets is open.*

Exercise 11.5 *Show that the set* $G=\left\{x\in C[0,1] : \left|f\left(\frac{1}{2}\right)\right|<1\right\}$ *is open in* $C[0,1]$.

12 Convergence in Metric Spaces (Lecture Notes)

12.1 Continuous Maps

Definition 12.1

- *Let* (X, d_X) *and* (Y, d_Y) *be metric spaces. A map* $T : X \mapsto Y$ *is said to be continuous at* x_0 *if*

$$\forall\, \epsilon > 0,\, \exists\, \delta > 0 : \forall\, x \in X,\, d_X(x, x_0) < \delta \Rightarrow d_Y(Tx_0, Tx) < \epsilon.$$

- *A function* T *is continuous on* X *if it is continuous at every point of* X*.*

Example 12.2 *The function* $T : l^\infty \mapsto \mathbb{R}^2$ *defined as*

$$Tx = (\xi_1, \xi_2),\; x = (\xi_k)_{k=1}^{\infty}$$

is continuous. Indeed take $x = (\xi_k)_{k=1}^{\infty} \in l^\infty$ *and* $\epsilon > 0$*. Then for all* $y = (\eta_k)_{k=1}^{\infty} \in l^\infty$ *such that*

$$d_{l^\infty}(x, y) = \sup_k |\xi_k - \eta_k| < \delta,$$

where δ *will be chosen later, we have*

$$d_{\mathbb{R}^2}(Tx,\, Ty) = d_{\mathbb{R}^2}\big((\xi_1, \xi_2), (\eta_1, \eta_2)\big) = \sqrt{(\xi_1 - \eta_1)^2 + (\xi_2 - \eta_2)^2} < \sqrt{\delta^2 + \delta^2} = \sqrt{2}\delta = \epsilon.$$

Hence $\delta = \frac{\epsilon}{\sqrt{2}}$*. So* T *is continuous at all* $x \in l^\infty$ *and is thus continuous on* l^∞*.*

Theorem 12.3 *A map* $T : X \mapsto Y$ *is continuous on* X *if and only if for all sets* G *open in* Y *the set*

$$f^{-1}(G) = \{x \in X : f(x) \in G\}$$

is open in X*.*

Definition 12.4

- *A point* x_0 *is called a limit point of a set* $M \subset X$ *if*

$$\forall\, \epsilon > 0,\, \exists\, x \in M,\, x \neq x_0 : x \in B_\epsilon(x_0).$$

- *The set* $\overline{M}$ *which contains all points of* M *and all limit points of* M *is called the closure of* M*.*

Example 12.5 *Take* $X = \mathbb{R}^2$ *and* $M = \mathbb{Q}^2 = \{(\xi_1, \xi_2) \in \mathbb{R}^2 : \xi_1, \xi_2 \in \mathbb{Q}\}$*. Then* $\overline{\mathbb{Q}^2} = \mathbb{R}^2$ *since every point of* $\mathbb{R}^2$ *is a limit point of* $\mathbb{Q}^2$*:*

$$\forall\, \epsilon > 0,\, \exists\, x \in \mathbb{Q}^2,\, x \neq x_0 : x \in B_\epsilon(x_0),\, \forall\, x_0 \in \mathbb{R}^2.$$

Exercise 12.6 *Propose a metric space* X *and a ball* $B_r(x_0) \in X$ *such that*

$$\overline{B_r(x_0)} \neq \overline{B_r}(x_0) = \{x \in X : d(x, x_0) \leqslant r\}.$$

Definition 12.7

- *A subset M of X is called dense in X if $\overline{M} = X$.*
- *X is called separable if there exists a countable subset $M \subset X$ which is dense in X.*

Example 12.8 *According to Example 12.5 the metric space $\mathbb{R}^2$ is separable.*

Remark 12.9 *A metric space is separable if there exists a countable set $M \subset X$ such that every ball $B_r(x)$, $r > 0$, $x \in X$ contains points from M, that is,*

$$\forall\, x \in X,\, r > 0,\, B_r(x) \cap M \neq \emptyset.$$

Remark 12.10 *The spaces $\mathbb{R}$, $\mathbb{R}^n$, c, $C[a,b]$, l^p, l^p_n, L^p are separable while the spaces $B[a,b]$ and l^∞ are not.*

Example 12.11 *l^p is separable. To show this, take*

$$M = \{x \in l^p : x = (\xi_1, \ldots, \xi_n, 0, 0, \ldots),\, \xi_k \in \mathbb{Q},\, k = 1, \ldots, n,\, n \geqslant 1\}.$$

Remark that M is countable. Indeed, we can identify

$$M_n = \{x \in l^p : x = (\xi_1, \ldots, \xi_n, 0, 0, \ldots),\, \xi_k \in \mathbb{Q}\}$$

with $\mathbb{Q}^n$ that is countable. Consequently $M = \bigcup_{n=1}^{\infty} M_n$ is countable. Let us show that $\overline{M} = X$. By Remark 12.9 we need to take arbitrary $x \in l^p$ and $r > 0$, and find $y \in M : y \in B_r(x) \Leftrightarrow d(x,y) < r$. Since $x \in l^\infty$

$$\sum_{k=1}^{\infty} |\xi_k|^p < \infty.$$

There exists $n \geqslant 1$ such that

$$\sum_{k=n+1}^{\infty} |\xi_k|^p < \delta_1 = \frac{\epsilon^p}{2}.$$

Next we choose $\eta_k \in \mathbb{Q}$, $k = 1, \ldots, n$ such that

$$|\xi_k - \eta_k| < \delta_2 = \frac{\epsilon}{\sqrt[p]{2n}},\ k = 1, \ldots, n.$$

Take $y = (\eta_1, \ldots, \eta_n, 0, 0, \ldots) \in M$. Then

$$d^p(x,y) = \sum_{k=1}^{\infty} |\xi_k - \eta_k|^p = \sum_{k=1}^{n} |\xi_k - \eta_k|^p + \sum_{k=n+1}^{\infty} |\xi_k|^p < n\delta_2^p + \delta_1 = \frac{\epsilon^p}{2} + \frac{\epsilon^p}{2} = \epsilon^p.$$

12.2 Convergence, Cauchy Sequences, Completeness

Definition 12.12

- *A sequence $\{x_n\}_{n \geqslant 1}$ in a metric space (X,d) is said to be convergent if there exists $x \in X$ such that*

$$\lim_{n \to \infty} d(x_n, x) = 0.$$

- *x is called the limit of $\{x_n\}_{n \geqslant 1}$ and we write $\lim_{n \to \infty} x_n = x$ or $x_n \to x$.*

Remark 12.13 *$x_n \to x$ if and only if $\forall\, \epsilon > 0,\ \exists\, N : d(x_n, x) < \epsilon,\ \forall\, n \geqslant N$.*

A set M is bounded if it is contained in a ball $B_r(x_0)$, that is, $\exists\, x_0 \in X,\ r > 0 : M \subset B_r(x_0)$.

Lemma 12.14 *Let (X, d) be a metric space.*

1. *A convergent sequence in X is bounded and its limit is unique.*
2. *If $x_n \to x$ and $y_n \to y$ in X, then $d(x_n, y_n) \to d(x, y)$.*

13 Completeness of Metric Spaces (Lecture Notes)

13.1 Cauchy Sequences

Recall the definition of convergence in a metric space.

Theorem 12.12

- *A sequence $\{x_n\}_{n\geqslant 1}$ in a metric space (X, d) is said to be convergent if there exists $x \in X$ such that*

$$\lim_{n\to\infty} d(x_n, x) = 0.$$

- *x is called the limit of $\{x_n\}_{n\geqslant 1}$ and we write $\lim\limits_{n\to\infty} x_n = x$ or $x_n \to x$.*

Definition 13.1

- *A sequence $\{x_n\}_{n\geqslant 1}$ is said to be a Cauchy sequence if $d(x_n, x_m) \to 0$, $n, m \to \infty$, i.e.*

$$\forall\, \epsilon > 0,\, \exists\, N : d(x_n, x_m) < \epsilon,\, \forall\, n, m \geqslant N.$$

- *The space X is said to be complete if every Cauchy sequence in X converges, that is, it has a limit which is an element of X.*

Example 13.2

1. *Spaces 1) - 9) from Lecture 11 are complete.*

2. *The metric space (X, d) where $X = \mathbb{Q}$ and $d(x, y) = |x - y|$, $x, y \in \mathbb{Q}$ is incomplete. Take*

$$x_n = \sum_{k=0}^{n} \frac{1}{k!} \in \mathbb{Q}.$$

We know that $\sum_{k=0}^{\infty} \frac{1}{k!} = e \notin \mathbb{Q}$. The sequence $\{x_n\}_{n\geqslant 1}$ is a Cauchy sequence. Indeed, for $n < m$

$$d(x_n, x_m) = |x_n - x_m| = \sum_{k=n+1}^{m} \frac{1}{k!} \leqslant \sum_{k=n+1}^{\infty} \frac{1}{k!} \to 0,\ n, m \to \infty.$$

But $\{x_n\}_{n\geqslant 1}$ is not convergent in $X = \mathbb{Q}$ because there exists no $x \in \mathbb{Q}$ such that $x_n \to x$ in $X = \mathbb{Q}$.

3. *Take $X = (0,1)^2 = \{(\xi_1, \xi_2) : \xi_1, \xi_2 \in (0,1)\}$ and $d(x, y) = \sqrt{(\xi_1 - \eta_1)^2 + (\xi_2 - \eta_2)^2}$. This metric space is incomplete. Indeed, take $x_n = \left(\frac{1}{n}, \frac{1}{n}\right) \in X$, $n \geqslant 2$. Then*

$$d(x_n, x_m) = \sqrt{\left(\frac{1}{n} - \frac{1}{m}\right)^2 + \left(\frac{1}{n} - \frac{1}{m}\right)^2} = \sqrt{2}\left|\frac{1}{n} - \frac{1}{m}\right| \to 0,\ n, m \to \infty.$$

Hence $\{x_n\}_{n\geqslant 2}$ is a Cauchy sequence but $\nexists\, x \in X : x_n \to x$, because $x_n \to (0,0) \notin X$.

4. *Let $X = C[0,1]$ and take*

$$d(x, y) = \int_0^1 |x(t) - g(t)|\, dt.$$

(X, d) is a metric space but it is not complete.

Take

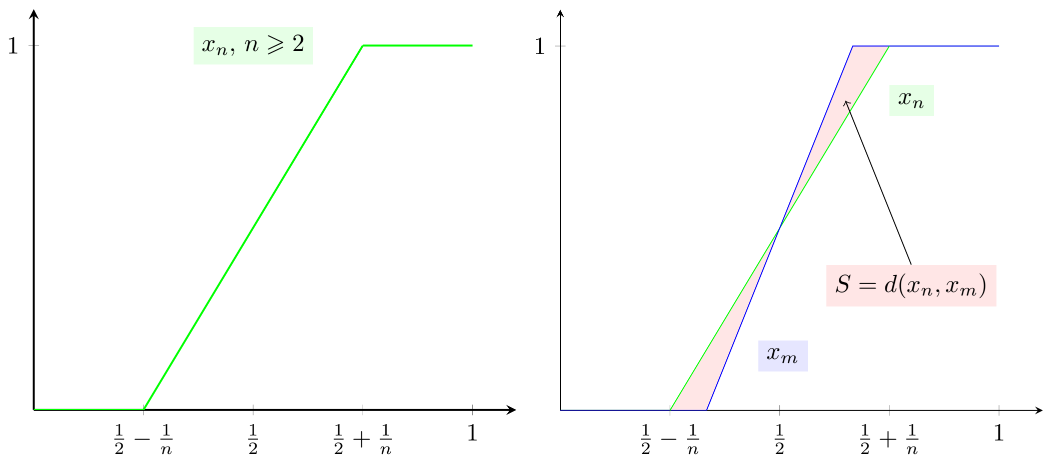

Figure 13.1

Hence $\{x_n\}_{n\geqslant 1}$ *is a Cauchy sequence but it does not converge in* $C[0,1]$*:*

$$x_n \to \begin{cases} 1, & x \geqslant \frac{1}{2} \\ 0, & x < \frac{1}{2} \end{cases} \notin C[0,1].$$

13.2 Some Properties

Theorem 13.3 *Every convergent sequence in a metric space is a Cauchy sequence.*

Proof: Let $\{x_n\}_{n\geqslant 1}$ converge to x. Then

$$0 \leqslant d(x_n, x_m) \leqslant d(x_n, x) + d(x_m, x) \to 0,\ n, m \to \infty.$$

This implies that $\{x_n\}_{n\geqslant 1}$ is a Cauchy sequence. □

Exercise 13.4 *Show that a Cauchy sequence is bounded, that is,* $\{x_n\}_{n\geqslant 1}$ *is bounded if* $\exists\, y \in X,\ r > 0$ *such that* $x_n \in B_r(y)$*.*

Example 13.5 *Let us prove that* l^p *is a complete metric space. Recall that for* l^p

$$X = \left\{ x = (\xi_k)_{k=1}^{\infty} : \sum_{k=1}^{\infty} |\xi_k|^p < \infty \right\}, \quad d(x,y) = \left(\sum_{k=1}^{\infty} |\xi_k - \eta_k|^p \right)^{\frac{1}{p}}.$$

Take a Cauchy sequence $x_n = (\xi_k^n)_{k=1}^{\infty} \in l^p$*.*

1. *First we need to show that* $\{\xi_l^n\}_{n\geqslant 1}$ *is a Cauchy sequence in* $\mathbb{R}$ *for all* l*. Indeed*

$$|\xi_l^n - \xi_l^m| = \left(|\xi_l^n - \xi_l^m|^p\right)^{\frac{1}{p}} \leqslant \left(\sum_{k=1}^{\infty} |\xi_k^n - \xi_k^m|^p \right)^{\frac{1}{p}} = d(x^n, x^m) \to 0,\ n, m \to \infty.$$

   So $\{\xi_l^n\}_{n\geqslant 1}$ *is a Cauchy sequence in* $\mathbb{R}$*, and since* $\mathbb{R}$ *is complete, there exists* $\xi_l \in \mathbb{R}$ *such that* $\xi_l^n \to \xi_l,\ n \to \infty$*.*

2. *Next we need to show that $x = (\xi_k)_{k=1}^{\infty} \in l^p$ and $x_n \to x$ in l^p. Take $\epsilon > 0$. By the fact that $\{x_n\}_{n\geqslant 1}$ is a Cauchy sequence, $\exists\, N \geqslant 1 : d(x_n, x_m) < \frac{\epsilon}{2}, \forall\, n, m \geqslant N$. So*

$$\left(\sum_{k=1}^{\infty} |\xi_k^n - \xi_k^m|^p\right)^{\frac{1}{p}} < \tfrac{\epsilon}{2} \Rightarrow \sum_{k=1}^{\infty} |\xi_k^n - \xi_k^m|^p < \frac{\epsilon^p}{2^p}.$$

By Fatou's lemma

$$\sum_{k=1}^{\infty} \lim_{m\to\infty} |\xi_k^n - \xi_k^m|^p = \sum_{k=1}^{\infty} |\xi_k^n - \xi_k|^p \leqslant \frac{\epsilon^p}{2^p} < \epsilon^p.$$

So

$$\left(\sum_{k=1}^{\infty} |\xi_k^n - \xi_k|^p\right)^{\frac{1}{p}} < \epsilon, \forall\, n \geqslant N.$$

We need only to show that $x = (\xi_k)_{k=1}^{\infty} \in l^p$. By Fatou's lemma

$$\sum_{k=1}^{\infty} |\xi_k|^p = \sum_{k=1}^{\infty} \lim_{n\to\infty} |\xi_k^n|^p \leqslant \underline{\lim}_{n\to\infty} \sum_{k=1}^{\infty} |\xi_k^n|^p = \underline{\lim}_{n\to\infty} d(0, x_n) < \infty,$$

because $\{x_n\}_{n\geqslant 1}$ is bounded.

Theorem 13.6 *Let $M \subset X$ be non-empty.*

1. *$x \in \overline{M} \Leftrightarrow \exists\, x_n \in M,\, n \geqslant 1 : x_n \to x$.*

2. *M is closed if and only if for all $\{x_n\}_{n\geqslant 1} \in M$ such that $x_n \to x$ in X we have that $x \in M$.*

Theorem 13.7 *Let (X, d) be a complete metric space and take $M \subset X$. The metric subspace (M, d) is complete if and only if M is a closed subset of X.*

Proof: Let (M, d) be complete. We will prove that M is closed in X by using Theorem 13.6 2). Take a sequence $\{x_n\}_{n\geqslant 1} \in M$ such that $x_n \to x$ in X. Then by Theorem 13.3 $\{x_n\}_{n\geqslant 1}$ is a Cauchy sequence in X, that is $d(x_n, x_m) \to 0$, $n, m \to \infty$. Then $\{x_n\}_{n\geqslant 1}$ is a Cauchy sequence in (M, d). Since M is complete, there exists $y \in M$ such that $x_n \to y$ in M, that is, $d(x_n, y) \to 0$, $n \to \infty$. Then $x_n \to y$ in X. Since the limit is unique by Lemma 12.14, $x = y \in M$. Now let M be closed in X and let (X, d) be complete. Take a Cauchy sequence $\{x_n\}_{n\geqslant 1}$ in M. $\{x_n\}_{n\geqslant 1}$ is also a Cauchy sequence in X, so by the completeness of X, there exists $x \in X$ such that $x_n \to x$, $n \to \infty$ in X. Then by Theorem 13.6 2) $x \in M$. So $x_n \to x$ in M, $n \to \infty$. □

Definition 13.8 (Isometric Spaces)

1. *A map $T : X \to \tilde{X}$ is said to be isometric if T preserves distances, that is, $\tilde{d}(Tx, Ty) = d(x, y)$ for all $x, y \in X$.*

2. *The space X is said to be isometric with the space $\tilde{X}$ if there exists a bijective isometry of X onto $\tilde{X}$. $X, \tilde{X}$ are called isometric spaces.*

Definition 13.9 *For a metric space (X, d) there exists a complete metric space $(\hat{X}, \hat{d})$ which has a subspace W that is isometric with X and dense in $\hat{X}$. This metric space $\hat{X}$ is unique except for isometries.*

14 Normed and Banach Spaces (Lecture Notes)

14.1 Vector Spaces

Let $K = \mathbb{R}$ or $K = \mathbb{C}$ be a fixed field of scalars.

Definition 14.1 *A vector space over a field of scalars K is a non-empty set X of elements called vectors, together with the operations of addition "+" and multiplication "·", satisfying the following conditions for any $\alpha, \beta \in K$ and $x, y, z \in X$:*

1. $x + y = y + x$
2. $(x + y) + z = x + (y + z)$
3. *There exists a vector* $0 \in X$ *such that* $\forall\, x \in X$, $0 + x = x$
4. $\forall\, x \in X$, $\exists\, y \in X$ *denoted by* $-x$ *such that* $x + y = 0$
5. $1 \cdot x = x$
6. $\alpha(x + y) = \alpha x + \alpha y$, $(\alpha + \beta)x = \alpha x + \beta x$

Recall that $Y \subset X$ is called a vector subspace of X if Y is closed with respect to "+" and "·", that is, $\forall\, x, y \in Y$, $\forall\, \alpha, \beta \in K$, $\alpha x + \beta y \in Y$.

Example 14.2 *The following sets together with "+" and "·" are vector spaces.*

1. $K^n = \{(\xi_1, \dots, \xi_n) : \xi_k \in K,\ k = 1, \dots, n\}$

$$(\xi_1, \dots, \xi_n) + (\eta_1, \dots, \eta_n) = (\xi_1 + \eta_1, \dots, \xi_n + \eta_n), \quad \alpha(\xi_1, \dots, \xi_n) = (\alpha\xi_1, \dots, \alpha\xi_n)$$

2. $C[a, b] = \{x : [a, b] \mapsto \mathbb{R} : x \,\text{is continuous on}\, [a, b]\}$

$$(x + y)(t) = x(t) + y(t), \quad (\alpha x)(t) = \alpha x(t)$$

3. $l^p = \{x = (\xi_1, \xi_2, \dots) : \xi_k \in \mathbb{R},\ \sum_{k=1}^{\infty} |\xi_k|^p < \infty\}$, $l^\infty = \{x = (\xi_1, \xi_2, \dots) : \xi_k \in \mathbb{R},\ \sup_k |\xi_k| < \infty\}$

$$x + y = (\xi_1 + \eta_1, \xi_2 + \eta_2, \dots), \quad \alpha x = (\alpha\xi_1, \alpha\xi_2, \dots)$$

4. $L^p[a, b] = \{x : [a, b] \mapsto \mathbb{R} : x \,\text{is measurable and}\, \int_a^b |x(t)|^p \, dt < \infty\}$

$$(x + y)(t) = x(t) + y(t), \quad (\alpha x)(t) = \alpha x(t)$$

We identify $x, y \in L^p$ *if* $x = y\ \lambda\text{−a. e.}$, *where* λ *is the Lebesgue measure on* $[a, b]$.

14.2 Normed and Banach Spaces

Definition 14.3

- *A norm on a vector space X is a real-valued function on X whose value at $x \in X$ is denoted by $\|x\|$ and which satisfies the following properties:*

  (N1) $\|x\| \geqslant 0, \forall\, x \in X$

  (N2) $\|x\| = 0 \Leftrightarrow x = 0$

  (N3) $\|\alpha x\| = |\alpha|\|x\|, \forall\, \alpha \in K,\, x \in X$

  (N4) $\|x + y\| \leqslant \|x\| + \|y\|, \forall\, x, y \in X$

- *A normed space X is a vector space with a norm defined on it.*

Let $(X, \|\cdot\|)$ be a normed vector space. The norm $\|\cdot\|$ defines the metric d on X given by

$$d(x, y) = \|x - y\|,\; x, y \in X.$$

The metric d is called the metric induced by the norm $\|\cdot\|$. We will also consider every normed space $(X, \|\cdot\|)$ as a metric space with the metric induced by the norm. So $\{x_n\}_{n\geqslant 1}$ converges in X if

$$\|x_n - x\| \to 0,\; n \to \infty.$$

Similarly $\{x_n\}_{n\geqslant 1}$ is a Cauchy sequence if

$$\|x_n - x_m\| \to 0,\; n, m \to \infty.$$

Definition 14.4 *A normed space $(X, \|\cdot\|)$ is called a Banach space if it is complete with respect to the metric induced by the norm $\|\cdot\|$.*

Exercise 14.5 *Show that a norm satisfies the inequality*

$$\big|\|x\| - \|y\|\big| \leqslant \|x - y\|.$$

This inequality implies that the map $X \ni x \mapsto \|x\| \in \mathbb{R}$ is continuous.

Example 14.6 *The following sets are Banach spaces.*

1. *Euclidean space $\mathbb{R}^n$ and unitary space $\mathbb{C}^n$*

$$\|x\| = \left(\sum_{k=1}^{\infty} |\xi_k|^2\right)^{\frac{1}{2}}$$

2. *Sequence spaces l^∞ and l^p*

$$\|x\| = \sup_{k\geqslant 1} |\xi_k|,\; x \in l^\infty, \quad \|x\| = \left(\sum_{k=1}^{\infty} |\xi_k|^p\right)^{\frac{1}{p}},\; x \in l^p$$

3. *Space c*

$$\|x\| = \sup_{k\geqslant 1} |\xi_k|$$

   Remark that c is a subspace of l^∞.

4. *Space $B(A)$*

$$\|x\| = \sup_{t\in A} |x(t)|$$

5. *Space $C[a,b]$*

$$\|x\| = \max_{t\in[a,b]} |x(t)|$$

6. *Spaces l_n^p, $p \geqslant 1$ and l_n^∞*

$$\|x\| = \left(\sum_{k=1}^{n} |\xi_k|^p\right)^{\frac{1}{p}}, \; x \in l_n^p, \quad \|x\| = \max_{k=1,\dots,n} |\xi_k|, \; x \in l_n^\infty$$

7. *Space $L^p[a,b]$, $p \geqslant 1$*

$$\|x\| = \left(\int_a^b |x(t)|^p \, dt\right)^{\frac{1}{p}}$$

Example 14.7 $(C[a,b], \|\cdot\|)$ *with norm*

$$\|x\| = \int_a^b |x(t)| \, dt$$

is incomplete because the metric space $C[a,b]$ with $d(x,y) = \int_a^b |x(t) - y(t)| \, dt$ is not complete.

14.3 Finite Dimensional Normed Spaces

Definition 14.8

- *Vectors $x_1, \dots, x_n \in X$ are called linearly independent if the equality*

$$\alpha_1 x_1 + \cdots + \alpha_n x_n = 0$$

  only holds if $\alpha_1 = \cdots = \alpha_n = 0$.

- *$M \subset X$ is linearly independent if every non-empty finite subset of M is linearly independent.*

- *A vector space X is finite dimensional if $\exists\, n \geqslant 1$ such that X contains a linearly independent set of vectors and every set containing more than n vectors is linearly dependent. The number $n = \dim X$ is called the dimension of X. If n does not exist, then X is infinite-dimensional.*

- *If $n = \dim X$, then any family of vectors $\{e_1, \dots, e_n\}$ that is linearly independent is called a basis for X. If $\{e_1, \dots, e_n\}$ is a basis, then for every vector $x \in X$ there exists a unique set of scalars $\alpha_1, \dots, \alpha_n$ such that*

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

- *We say that $Y \subset X$ is a subspace of a normed space X if Y is a vector subspace of X and the norm on Y is a restriction of the norm on X. Y is a closed subspace of X if additionally Y is a closed subset of X.*

14.4 Schauder Basis

In a normed space we can use series. Take $x_n \in X$, $n \geqslant 1$. We define the partial sum

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

We say that the series $\sum_{n=1}^{\infty} x_n$ converges if $\{S_n\}_{n\geqslant 1}$ is convergent, that is, there exists $S \in X$ such that $S_n \to S$, $n \to \infty$. The element S is called the sum of the series $\sum_{n=1}^{\infty} x_n$. A series $\sum_{n=1}^{\infty} x_n$ is absolutely convergent if $\sum_{n=1}^{\infty} \|x_n\|$ converges in $\mathbb{R}$.

Exercise 14.9 *Show that absolute convergence implies convergence in X if and only if X is a Banach space.*

Definition 14.10 *If a normed space X contains a sequence $\{e_n\}_{n\geqslant 1}$ with the property that for every $x \in X$ there exists a unique sequence of scalars $\{\alpha_n\}_{n\geqslant 1}$ such that*

$$x = \sum_{k=1}^{\infty} \alpha_k e_k,$$

then $\{e_n\}_{n\geqslant 1}$ is called a Schauder basis for X.

Exercise 14.11 *Show that if a normed space has a Schauder basis then X is separable. The inverse statement is not true in general.*

Example 14.12 $\{e_n = (0, 0, \dots, 0, 1, 0, \dots),\ n \geqslant 1\}$*, where 1 is in the nth position is a Schauder basis for l^p, $p \geqslant 1$.*

15 Linear Operators (Lecture Notes)

15.1 Basic Definition

Let X, Y be vector spaces over the same scalar field K.

Definition 15.1

- *A linear operator T is a map from $\mathcal{D}(T) \subset X$ to Y such that*
  1. *the domain $\mathcal{D}(T)$ is a vector subspace of X,*
  2. *$\forall\, x, y \in \mathcal{D}(T)$ and for any scalar α, $T(x+y) = Tx + Ty$ and $T(\alpha x) = \alpha T(x)$.*
- *If $Y = K$, then T is called a linear functional.*

Example 15.2

1. *Consider $X = \mathbb{R}^n$ and $Y = \mathbb{R}^m$. Let $A = (a_{ij})_{i,j=1}^{m,n}$ be an $m \times n$ matrix. We define*

$$Tx = Ax,\ x \in \mathbb{R}^n,$$

   that is, for $x = (\xi_1, \ldots, \xi_n)$ we have

$$Tx = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix} = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_m \end{pmatrix}.$$

   Then $\mathcal{D}(T) = \mathbb{R}^n$ and T is a linear operator.

2. *Consider $X = C[a,b]$ and $Y = C[a,b]$. We define*

$$(Tx)(t) = \int\limits_{\alpha}^{t} x(s)\, ds,\ t \in [a,b].$$

   Then $\mathcal{D}(T) = C[a,b]$.

3. *Consider $X = C[a,b]$ and $Y = C[a,b]$. We define*

$$(Tx)(t) = x'(t),\ t \in [a,b].$$

   Then $C[a,b] \supset \mathcal{D}(T) = C^1[a,b]$, the set of all continuously differentiable functions on $[a,b]$.

4. *Consider $X = L^p[a,b]$ and $Y = L^q[a,b]$. Fix $\varphi : [a,b] \mapsto \mathbb{R}$ that is Lebesgue measurable. We define*

$$(Tx)(t) = \varphi(t)x(t).$$

   Then $\mathcal{D}(T) = \{x \in L^p : \int_a^b |\varphi(t)x(t)|^q\, dt < \infty\}$.

5. *Consider $X = l^\infty$ and $Y = \mathbb{R}$. Take*

$$Tx = \lim_{k\to\infty} \xi_k,\ x = (\xi_k)_{k=1}^{\infty}.$$

   Then $\mathcal{D}(T) = c \subset l^\infty$ and T is a linear functional.

15.2 Bounded and Continuous Linear Operators

Let X and Y be normed spaces over the same scalar field.

Definition 15.3

- *A linear operator* $T : \mathcal{D}(T) \mapsto Y$, $\mathcal{D}(T) \subset X$ *is said to be bounded if there exists* $C > 0$ *such that*

$$\|Tx\| \leqslant C\|x\|. \tag{15.1}$$

- *The number*

$$\|T\| = \sup_{\substack{x\in\mathcal{D}(T)\\ x\neq 0}} \frac{\|Tx\|}{\|x\|}$$

  is called the norm of T.

Exercise 15.4

1. *Show that* $\|T\|$ *is the smallest constant* C *satisfying (15.1), that is,*

$$\|T\| = \min\{C : \|Tx\| \leqslant C\|x\|, \forall\, x \in \mathcal{D}(T)\}.$$

2. *Show that* $\|T\| = \sup\limits_{\substack{x\in\mathcal{D}(T)\\ \|x\|=1}} \|Tx\|$.

Example 15.5

1. *Consider* $X = Y = C[0,1]$. *We define*

$$(Tx)(t) = \int_0^t x(s)\, ds,\ x \in C[0,1] = \mathcal{D}(T).$$

We claim that T *is bounded. To show this we compute*

$$\begin{aligned}\|Tx\| &= \max_{t\in[0,1]} \left|\int_0^t x(s)\, ds\right| \leqslant \max_{t\in[0,1]} \int_0^t |x(s)|\, ds \\ &\leqslant \max_{t\in[0,1]} \int_0^t \max_{s\in[0,1]} |x(s)|\, ds = \|x\| \max_{t\in[0,1]} \int_0^t ds = \|x\| \max_{t\in[0,1]} t = \|x\|,\end{aligned}$$

where we use $\max\limits_{s\in[0,1]} |x(s)| = \|x\|$, $x \in C[0,1]$. *So* $\|T\| \leqslant 1$. *Let us show that* $\|T\| = 1$. *Take* $x = 1$. *Then* $\|x\| = 1$ *and moreover,* $(Tx)(t) = \int_0^t 1\, ds = t$. *So*

$$\|Tx\| = 1 \Rightarrow \|T\| \geqslant \frac{\|Tx\|}{\|x\|} = 1.$$

This implies $\|T\| = 1$.

2. *Take again* $X = Y = C[0,1]$. *We consider*

$$(Tx)(t) = x'(t),\ \mathcal{D}(T) = C^1[0,1].$$

We claim that T *is unbounded. Take* $x_n(t) = t^n$, $t \in [0,1]$, $n \geqslant 1$. *We compute*

$$\|x_n\| = \max_{t\in[0,1]} |t^n| = 1, \quad \|Tx_n\| = \max_{t\in[0,1]} |nt^{n-1}| = n.$$

Then

$$\|T\| \geqslant \frac{\|Tx_n\|}{\|x_n\|} = n,\ \forall\, n \geqslant 1.$$

We see that there does not exist C *such that for all* n *we have* $n \leqslant C$, *so* T *is unbounded.*

Theorem 15.6 *Let* X *be a finite-dimensional normed space and* T *a linear operator on* X. *Then* T *is bounded.*

Let us recall that $T : \mathcal{D}(T) \mapsto Y$ is continuous at $x_0 \in \mathcal{D}(T)$ if

$$\forall\, \epsilon > 0,\ \exists\, \delta > 0 : \forall\, x \in \mathcal{D}(T),\ \|x - x_0\| < \delta \Rightarrow \|Tx - Tx_0\| < \epsilon.$$

Theorem 15.7 *Let* $T : \mathcal{D}(T) \mapsto Y$ *be a linear operator.*

1. T *is continuous if and only if* T *is bounded.*
2. *If* T *is continuous at a single point, then it is continuous at every point.*

Proof: We prove the first statement. For $T = 0$, the statement is trivial. So we take $T \neq 0 \Rightarrow \|T\| \neq 0$. Let the boundedness of T be given. Take $x_0 \in \mathcal{D}(T)$ and show that T is continuous:

$$\forall\, \epsilon > 0,\ \exists\, \delta > 0 : \forall\, x \in \mathcal{D}(T),\ \|x - x_0\| < \delta \Rightarrow \|Tx - Tx_0\| < \epsilon.$$

Let $\epsilon > 0$ be given. We take $\delta = \frac{\epsilon}{\|T\|}$ and $x \in \mathcal{D}(T) : \|x - x_0\| < \delta$. Then

$$\|Tx - Tx_0\| = \|T(x - x_0)\| \leqslant \|T\|\|x - x_0\| < \|T\|\delta = \|T\|\frac{\epsilon}{\|T\|} = \epsilon.$$

Since x_0 was arbitrary, T is continuous. Now let T be continuous at $x_0 \in \mathcal{D}(T)$. Then if we choose $\epsilon = 1$, we can find δ such that

$$\|x - x_0\| < \delta \Rightarrow \|Tx - Tx_0\| < \epsilon = 1.$$

Now take any $y \neq 0$ from $\mathcal{D}(T)$ and set $x = x_0 + \frac{\delta}{2\|y\|}y$. Then

$$\|x - x_0\| = \frac{\delta}{2} < \delta \Rightarrow \|Tx - Tx_0\| < \epsilon = 1.$$

Then we also have

$$1 > \|Tx - Tx_0\| = \|T(x - x_0)\| = \left\|\left(T\frac{\delta}{2\|y\|}y\right)\right\| = \frac{\delta}{2\|y\|}\|Ty\|.$$

Thus

$$\frac{\delta}{2\|y\|}\|Ty\| < 1 \Rightarrow \|Ty\| < \frac{2}{\delta}\|y\|.$$

Since $y \in \mathcal{D}(T)$ was arbitrary, this implies that T is bounded, as we can take $C = \frac{2}{\delta}$. Remark that we only used the continuity of T at x_0. So we conclude that if T is continuous at x_0, it must be bounded, and if it is bounded, it must then be continuous on $\mathcal{D}(T)$, proving the second statement. $\square$

Corollary 15.8 *Let T be a bounded linear operator.*

1. *For $x_n, x \in \mathcal{D}(T)$, we have $x_n \to x \Rightarrow Tx_n \to Tx$.*
2. *The null set $\ker(T) = \{x : Tx = 0\}$ is closed in X.*

Exercise 15.9 *Prove Corollary 15.8.*

Theorem 15.10 *Let $T : \mathcal{D}(T) \mapsto Y$ be a bounded linear operator and Y a Banach space. Then T has an extension $\tilde{T} : \overline{\mathcal{D}(T)} \mapsto Y$, where $\tilde{T}$ is a bounded linear operator and $\|\tilde{T}\| = \|T\|$.*

Proof: We only show how $\tilde{T}$ can be constructed. Take $x \in \overline{\mathcal{D}(T)}$. Then there exists a sequence $x_n \in \mathcal{D}(T)$ such that $x_n \to x$. Since T is linear and bounded, then

$$\|Tx_n - Tx_m\| \leqslant \|T(x_n - x_m)\| \leqslant \|T\|\|x_n - x_m\| \to 0,\ n, m \to \infty.$$

So $\{Tx_n\}_{n \geqslant 1}$ is a Cauchy sequence in Y. Since Y is a Banach space, there exists $y \in Y$ such that $Tx_n \to y$, $n \to \infty$. Set $\tilde{T}x := y$. Now we show that $\tilde{T}x$ is well-defined. If z_n, $n \geqslant 1$ is another sequence from $\mathcal{D}(T)$ converging to x, then $Tz_n \to y'$. Consider the sequence $v_n : x_1, z_1, x_2, z_2, x_3, z_3, \ldots$. This sequence converges to x and $Tv_n \to y''$. But $Tv_{2k+1} \to y = y''$ and $Tv_{2k} \to y' = y''$. This implies $y = y'$. $\square$

16 Dual Spaces (Lecture Notes)

16.1 Normed Spaces of Operators

Let X and Y be normed spaces. Consider bounded linear operators $T : X \mapsto Y$ such that $\|Tx\| \leqslant C\|x\|$. Denote $B(X,Y)$ the set of all such bounded linear operators. $B(X,Y)$ is a vector space if we define

$$(T_1 + T_2)(x) = T_1 x + T_2 x, \quad (\alpha T)(x) = \alpha T x,$$

where $T_1, T_2, T \in B(X,Y)$, $\alpha \in K$, and $x \in X$.

Theorem 16.1 *The vector space $B(X,Y)$ is a normed space with norm defined by*

$$\|T\| = \sup_{x \neq 0} \frac{\|Tx\|}{\|x\|} = \sup_{\|x\|=1} \|Tx\|.$$

Exercise 16.2 *Prove Theorem 16.1, that is, show that $\|\cdot\| : B(X,Y) \mapsto \mathbb{R}$ is a norm on $B(X,Y)$.*

Theorem 16.3 *If Y is a Banach space, then $B(X,Y)$ is a Banach space.*

Proof: Let $T_n \in B(X,Y)$, $n \geqslant 1$ be a Cauchy sequence in $B(X,Y)$. We want to show that there exists $T \in B(X,Y)$ such that $T_n \to T$ in $B(X,Y)$. Take $x \in X$ and define $Tx = \lim_{n\to\infty} T_n x$. Consider the sequence $T_n x$, $n \geqslant 1$ in Y. Claim that $T_n x$, $n \geqslant 1$ is a Cauchy sequence in Y:

$$\|T_n x - T_m x\| = \|(T_n - T_m)x\| = \|T_n - T_m\|\|x\| \to 0,\ n, m \to \infty.$$

Since Y is complete, there exists $y \in Y$ such that $T_n x \to y := Tx$. We have obtained the map $T : X \mapsto Y$. Now we show that T is linear:

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

Take $\epsilon > 0$. Then there exists $N : \forall\, n, m \geqslant N$, $\|T_n - T_m\| < \frac{\epsilon}{2}$. For $n \geqslant N$

$$\|T_n x - Tx\| = \left\|T_n x - \lim_{m\to\infty} T_m x\right\| = \lim_{m\to\infty} \|T_n x - T_m x\| \leqslant \lim_{m\to\infty} \|T_n - T_m\|\|x\| \leqslant \frac{\epsilon}{2}\|x\| < \epsilon\|x\|.$$

Thus T is bounded so $T \in B(X,Y)$. Furthermore, this implies $\|T_n - T\| \leqslant \frac{\epsilon}{2} < \epsilon$, $\forall\, n \geqslant N$, which means that $T_n \to T$, $n \to \infty$, that is, it converges in $B(X,Y)$. Therefore $B(X,Y)$ is a Banach space. □

16.2 Dual Spaces

Let X be a normed space and take $Y = K$.

Definition 16.4 *The set of all bounded linear functionals on X with norm*

$$\|f\| = \sup_{x \neq 0} \frac{|f(x)|}{\|x\|} = \sup_{\|x\|=1} |f(x)|$$

is called the dual space of X and is denoted by $X' = B(X,K)$.

Theorem 16.5 *The dual space X' of a normed space X is a Banach space, whether or not X is.*

Definition 16.6

- *An isomorphism of a normed space X onto another normed space $\tilde{X}$ is a bijective linear operator $T : X \mapsto \tilde{X}$ that preserves the norm, that is, $\|Tx\| = \|x\|$ for all $x \in X$.*
- *If there exists an isomorphism of X onto $\tilde{X}$, then X and $\tilde{X}$ are called isomorphic normed spaces.*

Example 16.7

1. *We have $(l_n^p)' = l_n^q$, $\frac{1}{p} + \frac{1}{q} = 1$, $1 < p < \infty$. Let $f \in (l_n^p)'$ be a bounded linear functional. Take a basis $e_1 = (1, 0, \ldots, 0), \ldots, e_n = (0, \ldots, 0, 1) \in l_n^p$. Then*

$$x = \sum_{k=1}^{n} \xi_k e_k \in l_n^p, \quad f(x) = f\left(\sum_{k=1}^{n} \xi_k e_k\right) = \sum_{k=1}^{n} \xi_k f(e_k) = \sum_{k=1}^{n} \gamma_k \xi_k = \langle u, x \rangle,$$

where $u = (\gamma_1, \ldots, \gamma_n)$, $\gamma_k = f(e_k)$, $k = 1, \ldots, n$. Next we compute the norm of f using the Hölder inequality:

$$|f(x)| = \left|\sum_{k=1}^{n} \gamma_k \xi_k\right| \leqslant \sum_{k=1}^{n} |\gamma_k \xi_k| \leqslant \left(\sum_{k=1}^{n} |\gamma_k|^q\right)^{\frac{1}{q}} \left(\sum_{k=1}^{n} |\xi_k|^p\right)^{\frac{1}{p}} = \|u\|_q \|x\|_p, \; \forall\, x \in l_n^p.$$

This implies $\|f\| \leqslant \|u\|_q$. Now take $x = \big(\pm|\gamma_1|^{q-1}, \ldots, \pm|\gamma_n|^{q-1}\big)$, where we take $+$ if $\gamma_k > 0$ and $-$ if $\gamma_k < 0$. Then

$$|f(x)| = \sum_{k=1}^{n} \gamma_k \big(\pm|\gamma_k|^{q-1}\big) = \sum_{k=1}^{n} |\gamma_k|^q$$

and

$$\|x\| = \left(\sum_{k=1}^{n} |\gamma_k|^{(q-1)p}\right)^{\frac{1}{p}} = \left(\sum_{k=1}^{n} |\gamma_k|^q\right)^{1-\frac{1}{q}}.$$

So

$$|f(x)| = \sum_{k=1}^{n} |\gamma_k|^q = \left(\sum_{k=1}^{n} |\gamma_k|^q\right)^{\frac{1}{q}} \left(\sum_{k=1}^{n} |\gamma_k|^q\right)^{1-\frac{1}{q}} = \|u\|\|x\|.$$

Hence $\|f\| = \|u\|$. Consequently, the map $f \mapsto \big(f(e_k)\big)_{k=1}^{n} =: u$ is an isomorphism of $(l_n^p)'$ onto l_n^q and $\|f\| = \|u\|_q$. In other words, any bounded linear functional f can be written in the form

$$f(x) = \sum_{k=1}^{n} \gamma_k \xi_k =: \langle u, x \rangle,$$

where $u = (\gamma_k)_{k=1}^{n} \in l_n^q$ and $\|f\| = \|u\|_q$.

2. $(l_n^1)' = l_n^\infty$ *and* $(l_n^\infty)' = l_n^1$.
3. $(l^p)' = l^q$, $\frac{1}{p} + \frac{1}{q} = 1$, $1 < p < \infty$
4. $(l^1)' = l^\infty$
5. $c' = (c_0)' = l^1$
6. $(L^p[a, b])' = L^q[a, b]$ *and* $(L'[a, b])' = L^\infty[a, b]$.
7. $(C[a, b])' =$ *"functions of bounded variation"*

16.3 Dual Space to $C[a,b]$

Definition 16.8 *A function $w : [a,b] \mapsto \mathbb{R}$ is said to be of bounded variation on $[a,b]$ if its total variation*

$$\mathrm{Var}(w) = \sup \sum_{j=1}^{n} |w(t_j) - w(t_{j-1})|$$

is finite, where the supremum is taken over all partitions $a = t_0 < t_1 < \cdots < t_n = b$.

Example 16.9 *If w is non-decreasing, then w has bounded variation. Indeed*

$$\mathrm{Var}(w) = \sup \sum_{j=1}^{n} |w(t_j) - w(t_{j-1})| = \sup \sum_{j=1}^{n} \big(w(t_j) - w(t_{j-1})\big) = w(b) - w(a).$$

Remark 16.10 *A function w has bounded variation if it can be written as a difference of two non-decreasing functions, that is $\exists\, w_1, w_2 : [a,b] \mapsto \mathbb{R}$ that are non-decreasing such that $w = w_1 - w_2$.*

Let $BV[a,b]$ be the set of all functions on $[a,b]$ of bounded variation. It is obvious that $BV[a,b]$ is a vector space over $K = \mathbb{R}$. Define the norm on this space as $\|w\| = |w(a)| + \mathrm{Var}(w)$.

Lemma 16.11 *$BV[a,b]$ is a Banach space.*

If $x \in C[a,b]$ and $w \in BV[a,b]$, then one can check that the Riemann-Stieltjes integral

$$\int_a^b x(t)\, dw(t) = \lim_{\lambda \to 0} \sum_{k=1}^{n} x(\xi_k)\big(w(t_k) - w(t_{k-1})\big)$$

exists, where $\lambda = \max_k |t_k - t_{k-1}|$, $\xi_k \in [t_{k-1}, t_k]$ and $a = t_0 < t_1 < \cdots < t_n = b$. Remark that if $w \in C^1[a,b]$, then $w \in BV[a,b]$ and

$$\int_a^b x(t)\, dw(t) = \int_a^b x(t) w'(t)\, dt.$$

Theorem 16.12 *Every $f \in (C[a,b])'$ can be expressed as a Riemann-Stieltjes integral:*

$$f(x) = \int_a^b x(t)\, dw(t)$$

with $\|f\| = \mathrm{Var}(w)$.

In Theorem 16.12 the function can be made unique if we additionally require that w is right continuous and $w(0) = 0$. So $(C[a,b])' = BV_0[a,b]$, where $BV_0[a,b] \subset BV[a,b]$ contains all right continuous functions of bounded variation.

17 Hilbert Spaces (Lecture Notes)

17.1 Definitions of Inner Product and Hilbert Spaces

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

Definition 17.1

- *An inner product on X is a map $\langle \cdot, \cdot \rangle : X \times X \mapsto K$ such that for all $x, y, z \in X$ and $\alpha \in K$*

  (IP1) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$

  (IP2) $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$

  (IP3) $\langle x, y \rangle = \overline{\langle y, x \rangle}$

  (IP4) $\langle x, x \rangle \geqslant 0$, $\langle x, x \rangle = 0 \Leftrightarrow x = 0$

- *A vector space X with an inner product on it is called an inner product space.*

Remark that $\langle x, y+z \rangle = \overline{\langle y + z, x \rangle} = \overline{\langle y, x \rangle + \langle z, x \rangle} = \overline{\langle y, x \rangle} + \overline{\langle z, x \rangle} = \langle x, y \rangle + \langle x, z \rangle$ and $\langle x, \alpha y \rangle = \overline{\alpha} \langle x, y \rangle$.

Example 17.2

1. *Euclidean space* $\mathbb{R}^n$
$$\langle x, y \rangle = \xi_1 \eta_1 + \cdots + \xi_n \eta_n$$

2. *Unitary space* $\mathbb{C}^n$
$$\langle x, y \rangle = \xi_1 \overline{\eta_1} + \cdots + \xi_n \overline{\eta}_n$$

3. *Space* $l^2 = \{x = (\xi_k)_{k=1}^{\infty} : \xi_k \in K,\ k \geqslant 1,\ \sum_{k=1}^{\infty} |\xi_k|^2 < \infty\}$
$$\langle x, y \rangle = \sum_{k=1}^{\infty} \xi_k \overline{\eta}_k$$

4. *Space* $L^2[a, b] = \{x : [a, b] \mapsto K : \int_a^b |x(t)|^2 \, dt < \infty\}$
$$\langle x, y \rangle = \int_a^b x(t) \overline{y(t)} \, dt$$

5. *Space* $L^2(\mathbb{R}) = \{x : \mathbb{R} \mapsto K : \int_{-\infty}^{\infty} |x(t)|^2 \, dt < \infty\}$
$$\langle x, y \rangle = \int_{-\infty}^{\infty} x(t) \overline{y(t)}, dt$$

17.2 Properties of Inner Product Spaces

Define $\|x\| = \sqrt{\langle x, x\rangle}$, $x \in X$. $\|\cdot\|$ satisfies properties (N1) - (N3) of Definition 14.3. The space X with norm $\|\cdot\|$ induced by the inner product is a normed space.

Lemma 17.3 (Cauchy-Schwarz and Triangle Inequalities)

1. *For all $x, y \in X$ we have $|\langle x, y\rangle| \leqslant \|x\|\|y\|$, and equality holds if and only if x and y are linearly dependent.*

2. *For all $x, y \in X$ we have $\|x + y\| \leqslant \|x\| + \|y\|$.*

Exercise 17.4 *Check that a norm $\|x\| = \sqrt{\langle x, x\rangle}$ on an inner product space satisfies the parallelogram equality:*

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

Remark 17.5 *Exercise 17.4 implies that l^p, $L^p[a, b]$, $p \neq 2$ and $C[a, b]$ are not inner product spaces. Let us show this explicitly for l^p. Take in l^p*

$$x = (1, 1, 0, 0, \dots), \quad y = (1, -1, 0, 0, \dots).$$

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

$$\|x + y\|^2 + \|x - y\|^2 = 2^2 + 2^2 \neq 2\left(2^{\frac{2}{p}} + 2^{\frac{2}{p}}\right) = 2(\|x\|^2 + \|y\|^2)$$

unless $p = 2$.

Lemma 17.6 *Let $x_n \to x$ and $y_n \to y$ in X. Then $\langle x_n, y_n\rangle \to \langle x, y\rangle$.*

Proof:

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

□

Definition 17.7 *An inner product space X that is complete in norm generated by the inner product is said to be a Hilbert space.*

A Hilbert space is a Banach space. A subspace Y of an inner product space X is defined to be a vector subspace of X take with the inner product on X restricted to $Y \times Y$.

Theorem 17.8 *Let Y be a subspace of a Hilbert space H.*

1. *Y is complete if and only if Y is closed in H.*

2. *If Y is finite-dimensional, then Y is complete.*

3. *If H is separable, then Y is also separable.*

Proof: 1) is a direct consequence of Theorem 13.7. 2) follows from the fact that every finite-dimensional space is closed. □

17.3 Orthogonality

Definition 17.9 *An element x of an inner product space X is said to be orthogonal to an element $y \in X$ if $\langle x, y\rangle = 0$. We also say that x and y are orthogonal and write $x \perp y$. Similarly, for subsets $A, B \subset X$, $x \perp A$ if $x \perp a$, $\forall\, a \in A$ and $A \perp B$ if $a \perp b$, $\forall\, a \in A$, $b \in B$.*

Now we are interested in finding a perpendicular from x to a subspace Y. Let M be a non-empty subset of X and let us define the distance from x to M as

$$\delta = \inf_{\tilde{y} \in M} \|x - \tilde{y}\|.$$

We want to know if there exists a unique $y \in M$ such that $\delta = \|x - y\|$.

Example 17.10 *Take $X = \mathbb{R}^2$.*

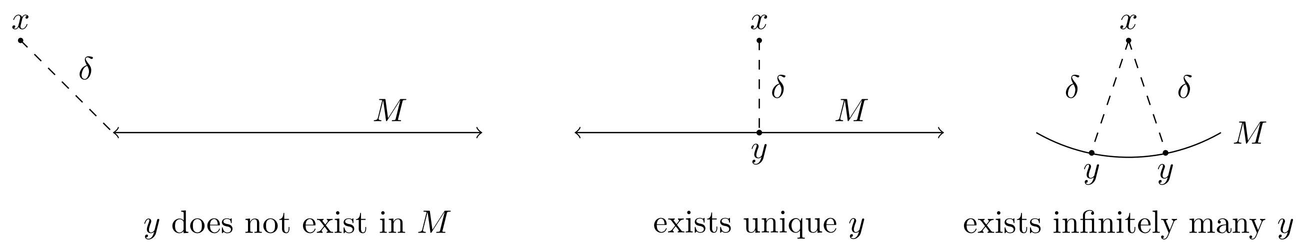

Figure 17.1

A subset M of X is convex if $\forall\, x, y \in M$ we have $\alpha x + (1-\alpha)y \in M$, $\forall\, \alpha \in [0, 1]$.

Theorem 17.11 *Let X be an inner product space and $M \neq \emptyset$ a complete convex subset of X. Then for every given $x \in X$ there exists a unique $y \in M$ such that*

$$\delta = \inf_{\tilde{y} \in M} \|x - \tilde{y}\| = \|x - y\|.$$

Proof Idea: We need to take a sequence $y_n \in M$ such that $\delta_n = \|x - y_n\| \to \delta$, $n \to \infty$ and show that it is a Cauchy sequence in M. Then there exists $y \in M$ such that $y_n \to y$. $\square$

Lemma 17.12 *If in Theorem 17.11 $M = Y$, where Y is a complete subspace of X, and $x \in X$ is fixed, then $z = x - y$ is orthogonal to Y.*

Let H be a Hilbert space and Y a closed subspace of H. We define the orthogonal complement as

$$Y^{\perp} = \{z \in H : z \perp Y\},$$

which is a vector subspace of H.

18 Orthonormal Sets (Lecture Notes)

18.1 Direct Sums

Let X be an inner product space over K. Assume that $Y \subset X$ is a complete subspace of X. Then we know that $z = x - y \perp Y$, where $\|x - y\| = \inf_{\tilde{y} \in Y} \|x - \tilde{y}\|$. We define the following subspace of X:

$$Y^{\perp} = \{z \in X : z \perp Y\}.$$

Theorem 18.1 *Let Y be any complete subspace of X. Then for every $x \in X$ unique $y \in Y$ and $z \in Y^{\perp}$ exist such that $x = y + z$.*

Proof: The existence of y and z follows from Theorem 17.11 and Lemma 17.12. Indeed, take $y \in Y$ such that

$$\inf_{\tilde{y} \in Y} \|x - \tilde{y}\| = \|x - y\|$$

and $z = x - y$. Then $z \in Y^{\perp}$, so $x = y + x - y = y + z$. To prove the uniqueness of y and z, we assume that $x = y + z = y_1 + z_1$, where $y, y_1 \in Y$ and $z, z_1 \in Y^{\perp}$. Then $Y \ni y - y_1 = z_1 - z \in Y^{\perp}$ and

$$\langle y - y_1, z_1 - z \rangle = \langle y - y_1, y - y_1 \rangle = 0$$

since $Y \perp Y^{\perp}$. This implies $y_1 = y$ and hence $z_1 = z$. □

Definition 18.2 *A vector space X is said to be a direct sum of two subspaces Y and Z of X, written $X = Y \oplus Z$, if $\forall\, x \in X$, $\exists!\, y \in Y$, $z \in Z : x = y + z$.*

Remark 18.3 *Let Y be a closed subspace. Then $X = Y \oplus Y^{\perp}$.*

18.2 Orthonormal Sets

Definition 18.4

- *An orthogonal set M in X is a subset of X whose elements are pairwise orthogonal:*

$$\langle x, y \rangle = 0,\ \forall\, x, y \in M,\ x \neq y.$$

- *An orthogonal set M is called orthonormal if*

$$\langle x, y \rangle = \begin{cases} 1, & x = y \\ 0, & x \neq y. \end{cases}$$

Exercise 18.5

1. *Show that for every $x, y \in X$, $x \perp y$ we have $\|x + y\|^2 = \|x\|^2 + \|y\|^2$.*

2. *Prove that an orthonormal set is linearly independent.*

Example 18.6

1. *$M = \{(1,0,0),(0,1,0),(0,0,1)\}$ and $M = \left\{\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0\right), \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0\right), (0,0,1)\right\}$ are orthonormal sets in $X = \mathbb{R}^3$.*

2. *Take $X = l^2$. The set $M = \{e_n,\, n \geqslant 1,\, e_n\}$, where $e_1 = (1,0,0,\dots)$, $e_2 = (0,1,0,\dots)$ and so on, is an orthonormal set.*

3. *Take $X = L^2[0,2\pi]$. The sets $M = \{e_n,\, n \geqslant 0\}$, where*

$$e_0(t) = \frac{1}{\sqrt{2\pi}}, \quad e_n(t) = \frac{\cos nt}{\sqrt{\pi}}$$

and $M = \{e_n,\, n \geqslant 1\}$, where

$$e_n(t) = \frac{\sin nt}{\sqrt{\pi}}$$

are orthonormal sets.

Remark 18.7 *Let $M = \{e_1, \dots, e_n\}$ be a basis in X. Then $\forall\, x \in X$, $\exists!\, \alpha_1, \dots, \alpha_n$ such that*

$$x = \alpha_1 e_1 + \cdots + \alpha_n e_n.$$

If M is orthonormal, that is, $\langle e_k, e_l \rangle = \delta_{kl}$, then

$$\langle x, e_k \rangle = \langle \alpha_1 e_1 + \cdots + \alpha_k e_k + \cdots + \alpha_n e_n, e_k \rangle = \alpha_1 \langle e_1, e_k \rangle + \cdots + \alpha_k \langle e_k, e_k \rangle + \cdots + \alpha_n \langle e_n, e_k \rangle = \alpha_k.$$

Now we want to extend the idea of Remark 18.7 to infinite-dimensional inner product spaces. Let $\{e_1, \dots, e_n\}$ be an orthonormal set in an infinite-dimensional space X. With some $x \in X$, take

$$y := \sum_{k=1}^{n} \langle x, e_k \rangle e_k, \quad z := x - y.$$

Then, applying the Pythagorean theorem

$$\begin{aligned} \langle z, y \rangle = \langle x - y, y \rangle = \langle x, y \rangle - \langle y, y \rangle &= \left\langle x, \sum_{k=1}^{n} \langle x, e_k \rangle e_k \right\rangle - \left\| \sum_{k=1}^{n} \langle x, e_k \rangle e_k \right\|^2 \\ &= \sum_{k=1}^{n} \overline{\langle x, e_k \rangle} \langle x, e_k \rangle - \sum_{k=1}^{n} \| \langle x, e_k \rangle e_k \|^2 = \sum_{k=1}^{n} |\langle x, e_k \rangle|^2 - \sum_{k=1}^{n} |\langle x, e_k \rangle|^2 \|e_k\|^2 = 0. \end{aligned}$$

Again by the Pythagorean theorem $\|x\|^2 = \|y\|^2 + \|z\|^2 \geqslant \|y\|^2 = \sum_{k=1}^{n} |\langle x, e_k \rangle|^2$.

Theorem 18.8 (Bessel Inequality) *Let $\{e_k,\, k \geqslant 1\}$ be an orthonormal sequence in an inner product space X. Then $\forall\, x \in X$*

$$\sum_{k=1}^{\infty} |\langle x, e_k \rangle|^2 \leqslant \|x\|^2.$$

Let $\{x_n,\, n \geqslant 1\}$ be linearly independent. We want to construct a sequence $\{e_n,\, n \geqslant 1\}$ such that

$$\operatorname{span}\{x_1, \dots, x_n\} = \operatorname{span}\{e_1, \dots, e_n\}, \ \forall\, n.$$

We use the Gram-Schmidt procedure:

$$e_1 := \frac{x_1}{\|x_1\|},$$

$$v_2 := x_2 - \langle x_2, e_1\rangle e_1, \quad e_2 := \frac{v_2}{\|v_2\|},$$

and in general

$$v_n := x_n - \sum_{k=1}^{n-1} \langle x_n, e_k\rangle e_k, \quad e_n := \frac{v_n}{\|v_n\|}.$$

18.3 Series Related to Orthonormal Sequences

Given any orthonormal sequence $\{e_k,\, k \geqslant 1\}$ we consider

$$\sum_{k=1}^{\infty} \alpha_k e_k,\ \alpha_k \in K. \tag{18.1}$$

We want to find for which α_k, $k \geqslant 1$ this series converges.

Theorem 18.9 *Let $\{e_k,\, k \geqslant 1\}$ be an orthonormal sequence in a Hilbert space H.*

1. *(18.1) converges in H if and only if*

$$\sum_{k=1}^{\infty} |\alpha_k|^2 < \infty.$$

2. *If (18.1) converges and*

$$x := \sum_{k=1}^{\infty} \alpha_k e_k,$$

   then $\alpha_k = \langle x, e_k\rangle$, $k \geqslant 1$.

3. *For every $x \in H$ the series*

$$\sum_{k=1}^{\infty} \langle x, e_k\rangle e_k$$

   converges, but not necessarily to x.

Proof:

1. Proving that $\sum_{k=1}^{n} \alpha_k e_k$ converges in H if and only if $\sum_{k=1}^{\infty} |\alpha_k|^2$ converges is equivalent to proving that $S_n = \alpha_1 e_1 + \cdots + \alpha_n e_n$ is a Cauchy sequence if and only if $R_n = |\alpha_1|^2 + \cdots + |\alpha_n|^2$ is a Cauchy sequence. We compute for $n < m$

$$\|S_m - S_n\|^2 = \|\alpha_{n+1} e_{n+1} + \cdots + \alpha_m e_m\|^2 = |\alpha_{n+1}^2 + \cdots + |\alpha_m|^2 = R_m - R_n.$$

   Indeed $\{S_n\}_{n\geqslant 1}$ is a Cauchy sequence in H if and only if $\{R_n\}_{n\geqslant 1}$ is a Cauchy sequence in $\mathbb{R}$.

2. Let $x = \sum_{k=1}^{\infty} \alpha_k e_k$. We compute for $k \leqslant n$ that $\langle S_n, e_k\rangle = \alpha_k$. Since $S_n \to x$, by the continuity of the inner product $\alpha_k = \langle S_n, e_k\rangle \to \langle x, e_k\rangle$, $n \to \infty$.

3. Using the Bessel inequality and the proof of 1), we have

$$\sum_{k=1}^{\infty} |\langle x, e_k\rangle|^2 \leqslant \|x\|^2 \Rightarrow \sum_{k=1}^{\infty} |\langle x, e_k\rangle|^2 < \infty \Rightarrow \sum_{k=1}^{\infty} \langle x, e_k\rangle e_k < \infty.$$

$\square$

18.4 Total Orthonormal Sets

Definition 18.10

- *A set $M \subset X$ is called a total orthonormal set if $\overline{\operatorname{span} M} = X$, that is, if $\overline{\operatorname{span} M}$ is dense in X.*
- *A total orthonormal family in X is called an orthonormal basis.*

Theorem 18.11 *In every Hilbert space H there exists a total orthonormal set.*

Theorem 18.12 (Parseval Equality) *Let M be an orthonormal set in a Hilbert space H. Then M is total in H if and only if*

$$\sum_k |\langle x, e_k \rangle|^2 = \|x\|^2, \, \forall \, x \in H.$$

Theorem 18.13 *Let H be a Hilbert space.*

1. *If H is separable, then every orthonormal set in H is countable.*
2. *If H contains a total orthonormal sequence, then H is separable.*

19 Adjoint Operators (Lecture Notes)

19.1 Examples of Orthonormal Bases

1. Legendre Polynomials
   We consider the space $L^2[-1,1]$ which is separable and is the space of all real-valued functions x given on $[-1,1]$ such that $\int_{-\infty}^{\infty} |x(t)|^2\,dt < \infty$. We want to find an orthonormal basis of functions for this space. For that we consider the linearly independent set of polynomials $M = \{x_n,\, n \geqslant 0\}$, where $x_n(t) = t^n$, $t \in [-1,1]$. Then $\overline{\operatorname{span} M} = L^2[-1,1]$, so M is a total set. However it is not orthonormal because

$$\langle x_k, x_l\rangle = \int_{-1}^{1} t^k t^l\,dt = \int_{-1}^{1} t^{k+l} \neq 0$$

   if $k+l$ is even. So we need to use the Gram-Schmidt procedure. In general we find

$$e_n(t) = \sqrt{\frac{2n+1}{2}}P_n(t), \quad P_n(t) = \frac{1}{2^n n!}\frac{d^n}{dt^n}(t^2-1)^n,$$

   where $P_n(t)$ are called the Legendre polynomials. The set $\{e_n,\, n \geqslant 0\}$ is an orthonormal basis in $L^2[-1,1]$:

$$x = \sum_{n=0}^{\infty}\langle x, e_n\rangle e_n,\ \forall\, x \in L^2[-1,1].$$

2. Hermite Polynomials
   We consider $L^2(\mathbb{R})$. Now $t^n \notin L^2(\mathbb{R})$ because $\int_{-\infty}^{\infty} |t^n|^2\,dt = \infty$. Instead we take $M = \{x_n,\, n \geqslant 0\}$, where $x_n(t) = t^n e^{-\frac{t^2}{2}}$, $t \in \mathbb{R}$. After normalization we find

$$e_n(t) = \frac{1}{\sqrt{2^n n!\sqrt{\pi}}}e^{-\frac{t^2}{2}}H_n(t), \quad H_n(t) = (-1)^n e^{t^2}\frac{d^n}{dt^n}e^{-t^2},$$

   where $H_n(t)$ are called the Hermite polynomials. The set $\{e_n,\, n \geqslant 0\}$ is an orthonormal basis in $L^2(\mathbb{R})$.

3. Laguerre Polynomials
   We consider $L^2[0,\infty)$ and $M = \{x_n,\, n \geqslant 0\}$, where $x_n(t) = t^n e^{-\frac{t}{2}}$, $t \geqslant 0$. Then

$$e_n(t) = e^{-\frac{t}{2}}L_n(t), \quad L_n(t) = \frac{e^t}{n!}\frac{d^n}{dt^n}(t^n e^{-t}),$$

   where $L_n(t)$ are called the Laguerre polynomials. The set $\{e_n,\, n \geqslant 0\}$ is an orthonormal basis in $L^2[0,\infty)$.

19.2 Adjoint Operators

Let H be a Hilbert space.

Theorem 19.1 (Riesz Representation Theorem) *Every bounded linear functional f on H can be written in terms of an inner product:*

$$f(x) = \langle x, z\rangle,$$

where z is a uniquely determined element of H, and $\|f\| = \|z\|$.

Definition 19.2 *Let H_1 and H_2 be Hilbert spaces and $T : H_1 \mapsto H_2$ a bounded linear operator. Then the adjoint operator T^* of T is the operator $T^* : H_2 \mapsto H_1$ such that*

$$\langle Tx, y\rangle = \langle x, T^*y\rangle, \forall\, x \in H_1,\, y \in H_2.$$

Theorem 19.3 *The adjoint operator T^* of T exists, is unique, and is bounded, with $\|T^*\| = \|T\|$.*

The existence of T^* follows from Theorem 19.1. Namely, consider for a fixed $y \in H_2$ the map

$$f(x) = \langle Tx, y\rangle,\, x \in H_1.$$

Then $f : H_1 \mapsto K$ is a bounded linear functional:

$$|f(x)| = |\langle Tx, y\rangle| \leqslant \|Tx\|\|y\| \leqslant \|T\|\|x\|\|y\| = C\|x\|.$$

By Theorem 19.1, there exists $z \in H_1$ such that $f(x) = \langle x, z\rangle$. We set $T^*y := z$.

Theorem 19.4 *Let H_1 and H_2 be Hilbert spaces and $T, S : H_1 \mapsto H_2$ bounded linear operators.*

1. $\langle T^*y, x\rangle = \langle y, Tx\rangle$, $x \in H_1$, $y \in H_2$
2. $(S + T)^* = S^* + T^*$
3. $(\alpha T)^* = \overline{\alpha} T^*$, $\alpha \in K$
4. $(T^*)^* = T$
5. $\|T^*T\| = \|TT^*\| = \|T\|^2$
6. $T^*T = 0 \Leftrightarrow T = 0$
7. $(ST)^* = T^*S^*$ *(if $H_1 = H_2$)*

19.3 Self-Adjoint, Unitary, and Normal Operators

We assume that H is a Hilbert space.

Definition 19.5 *A bounded linear operator $T : H \mapsto H$ is said to be*

- *self-adjoint if $T^* = T$*
- *unitary if T is bijective and $T^* = T^{-1}$*
- *normal if $TT^* = T^*T$*

Remark that if T is self-adjoint or unitary then it is normal. The inverse is not true.

Example 19.6 *If we take $T = 2iI$, where I is the identity operator, then $T^* = -2iI$. So $TT^* = T^*T$, but $T^* \neq T^{-1} = -\frac{1}{2}iI$ and $T \neq T^*$.*

Example 19.7 *Consider $\mathbb{C}^n$ with inner product*

$$\langle x, y\rangle = \sum_{k=1}^{n} \xi_k \overline{\eta}_k,\; x = (\xi_k)_{k=1}^n,\; y = (\eta_k)_{k=1}^n.$$

Any bounded linear operator $T : \mathbb{C}^n \mapsto \mathbb{C}^n$ can be given by a matrix M_T, that is, $y = Tx$ can be expressed as

$$\begin{pmatrix} \eta_1 \\ \vdots \\ \eta_n \end{pmatrix} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_n \end{pmatrix}.$$

If M_T is the matrix of T, then M_{T^*}, the matrix of T^*, is the conjugate transpose of M_T.

Theorem 19.8 *Let $T : H \mapsto H$ be a bounded linear operator.*

1. *If T is self-adjoint, then $\langle Tx, x\rangle$ is real for all $x \in H$.*
2. *If H is complex ($K = \mathbb{C}$) and $\langle Tx, x\rangle$ is real, then T is self-adjoint.*

Proof:

1. If T is self-adjoint, then

$$\overline{\langle Tx, x\rangle} = \langle x, Tx\rangle = \langle T^*x, x\rangle = \langle Tx, x\rangle \Rightarrow \langle Tx, x\rangle \in \mathbb{R}.$$

2. If $\langle Tx, x\rangle$ is real, then

$$\langle Tx, x\rangle = \overline{\langle Tx, x\rangle} = \overline{\langle x, T^*x\rangle} = \langle T^*x, x\rangle.$$

Hence

$$0 = \langle Tx, x\rangle - \langle T^*x, x\rangle = \langle Tx - T^*x, x\rangle = \langle (T - T^*)x, x\rangle \Rightarrow T = T^*.$$

□

Theorem 19.9

1. *The product of two bounded self-adjoint operators S and T is self-adjoint if and only if $ST = TS$.*
2. *Let T_n, $n \geqslant 1$ be self-adjoint operators on H such that $T_n \to T$ in $B(H, H)$. Then T is self-adjoint.*

Proof: We will only prove 2). We need to show that $T = T^*$. Consider

$$\|T_n^* - T^*\| = \|(T_n - T)^*\| = \|T_n - T\| \to 0.$$

So $T_n^* \to T^*$ and since $T_n = T_n^*$, then $T_n \to T^*$. This implies that $T = T^*$. □

20 Spectral Theory of Bounded Linear Operators (Lecture Notes)

We assume that all spaces are complex.

20.1 Basic Concepts

Assume $X \neq \emptyset$ is a complex normed space and consider the operators

$$T : \mathcal{D}(T) \mapsto X, \quad T - \lambda I : \mathcal{D}(T) \mapsto X,$$

where $Ix = x$ and $\lambda \in \mathbb{C}$. If it exists, denote

$$R_\lambda := R_\lambda(T) = (T - \lambda I)^{-1}.$$

Note that R_λ is a linear operator.

Definition 20.1

- *A regular value of T is a complex number λ such that*

  (R1) $R_\lambda(T)$ exists,

  (R2) $R_\lambda(T)$ is bounded,

  (R3) $R_\lambda(T)$ is defined on a dense subset of X.

- *The resolvent set $\rho(T)$ is the set of all regular values of T.*

- *The set $\sigma(T) = \mathbb{C} \setminus \rho(T)$ is called the spectrum of T.*

- *$\lambda \in \sigma(T)$ is called a spectral value of T.*

The spectrum $\sigma(T)$ is partitioned into three disjoint sets.

Definition 20.2

- *The point spectrum or discrete spectrum $\sigma_p(T)$ is the set such that $R_\lambda(T)$ does not exist.*

- *The continuous spectrum $\sigma_c(T)$ is the set such that $R_\lambda(T)$ exists and is defined on a dense subset of X, but $R_\lambda(T)$ is unbounded.*

- *The residual spectrum $\sigma_r(T)$ is the set such that $R_\lambda(T)$ exists but the domain of $R_\lambda(T)$ is not dense in X.*

Remark that $\sigma(T) = \sigma_p(T) \cup \sigma_c(T) \cup \sigma_r(T)$. We also note that $R_\lambda(T)$ does not exist if and only if $T - \lambda I$ is not injective, that is

$$\exists\, x \neq 0 : (T - \lambda I)x = Tx - \lambda x = 0.$$

Then $\lambda \in \sigma_p(T) \Leftrightarrow \exists\, x \neq 0 : Tx - \lambda x = 0$ and the vector x is called an eigenvector of T. If X is finite-dimensional then

$$\sigma_c(T) = \sigma_r(T) = \emptyset.$$

Example 20.3 *Take* $X = l^2 = \{x = (\xi_k)_{k\geqslant 1} : \xi_k \in \mathbb{C}, \sum_{k=1}^{\infty} |\xi_k|^2 < \infty\}$ *and define* $T : l^2 \mapsto l^2$ *such that*

$$Tx = (0, \xi_1, \xi_2, \xi_3, \dots),\ x = (\xi_k)_{k\geqslant 1}.$$

T *is called a right-shift operator and* $\mathcal{D}(T) = l^2$*. We have*

$$\|Tx\|^2 = \sum_{k=1}^{\infty} |\xi_k|^2 = \|x\|^2 \Rightarrow \|T\| = 1.$$

Now consider $\lambda = 0$*. We have* $R_0 = T^{-1}$, $\mathcal{D}(T^{-1}) = \{y = (\eta_k)_{k\geqslant 1} : \eta_1 = 0\}$*, and*

$$T^{-1}y = (\eta_2, \eta_3, \dots),\ y \in \mathcal{D}(T^{-1}).$$

In this case R_0 *exists but* $\mathcal{D}(T^{-1})$ *is not dense in* X*. Thus* $\lambda = 0$ *belongs to the residual spectrum of* T*.*

Proposition 20.4 *Let* X *be a complex Banach space and take* $T \in B(X, X)$ *and* $\lambda \in \rho(T)$*. Then* $R_\lambda(T)$ *is defined on the entire set* X *and is bounded.*

20.2 Spectral Properties of Bounded Linear Operators

Theorem 20.5 *Take* $T \in B(X, X)$*, where* X *is a Banach space. If* $\|T\| < 1$*, then* $(I - T)^{-1}$ *exists, belongs to* $B(X, X)$*, and*

$$(I - T)^{-1} = \sum_{k=0}^{\infty} T^k = I + T + T^2 + \dots,$$

where the series converges in $B(X, X)$*.*

Proof: Note that $\|T^k\| \leqslant \|T\|^k$. Since $\|T\| < 1$, we have

$$\sum_{k=0}^{\infty} \|T^k\| \leqslant \sum_{k=0}^{\infty} \|T\|^k < \infty.$$

This implies that the series

$$S := \sum_{k=0}^{\infty} T^k$$

converges. Now compute

$$(I - T)(I + T + T^2 + \dots + T^n) = (I + T + T^2 + \dots + T^n)(I - T) = I - T^{n+1}.$$

Since $\|T^{n+1}\| \leqslant \|T\|^{n+1} \to 0$, we get $(I - T)S = S(I - T) = I$, and thus $S = (I - T)^{-1}$. □

Theorem 20.6 *The resolvent set* $\rho(T)$ *of* $T \in B(X, X)$ *on a complex Banach space* X *is open. Hence the spectrum* $\sigma(T)$ *is closed.*

Theorem 20.7 *The spectrum* $\sigma(T)$ *of* $T \in B(X, X)$ *on a complex Banach space* X *is compact and lies in the disk* $|\lambda| \leqslant \|T\|$*.*

Proof: Take $\lambda \neq 0$ and denote $\theta = \frac{1}{\lambda}$. From Theorem 20.5 we obtain that

$$R_\lambda = (T - \lambda I)^{-1} = -\theta(I - \theta T)^{-1} = -\theta \sum_{k=0}^{\infty} (\theta T)^k = -\frac{1}{\lambda} \sum_{k=0}^{\infty} \left(\frac{1}{\lambda} T\right)^k,$$

where the series converges because

$$\left\| \frac{1}{\lambda} T \right\| = \frac{\|T\|}{|\lambda|} < 1.$$

So by Theorem 20.5 $R_\lambda \in B(X, X)$. Since $\sigma(T)$ is closed by Theorem 20.6 and bounded, we have that $\sigma(T)$ is compact. $\square$

Theorem 20.8 *Let X be a Banach space and $T \in B(X, X)$. Then for every $\lambda_0 \in \rho(T)$ the resolvent $R_\lambda(T)$ has the representation*

$$R_\lambda(T) = \sum_{k=0}^{\infty} (\lambda - \lambda_0)^k R_{\lambda_0}^{k+1},$$

where the series absolutely converges for λ in the open disk

$$|\lambda - \lambda_0| < \frac{1}{\|R_{\lambda_0}\|}$$

in the complex plane.

Definition 20.9 *The spectral radius $r_\sigma(T)$ of $T \in B(X, X)$ is the radius*

$$r_\sigma(T) = \sup_{\lambda \in \sigma(T)} |\lambda|.$$

One can show that $r_\sigma(T) = \lim_{n \to \infty} \sqrt[n]{\|T^n\|}$.

Theorem 20.10 (Resolvent Equation, Commutativity) *Let X be a complex Banach space and take $T \in B(X, X)$ and $\lambda, \mu \in \rho(T)$.*

1. $R_\mu - R_\lambda = (\mu - \lambda) R_\mu R_\lambda$
2. *R_λ commutes with any $S \in B(X, X)$ which commutes with T.*
3. $R_\lambda R_\mu = R_\mu R_\lambda$

Proof:

1. We have

$$\begin{aligned} R_\mu - R_\lambda &= R_\mu I - I R_\lambda = R_\mu\big((T - \lambda I) R_\lambda\big) - \big(R_\mu (T - \mu I)\big) R_\lambda = R_\mu (T - \lambda I - T + \mu I) R_\lambda \\ &= R_\mu (\mu - \lambda) R_\lambda = (\mu - \lambda) R_\mu R_\lambda. \end{aligned}$$

2. The assumption $TS = ST$ implies $(T - \lambda I)S = S(T - \lambda I)$. Thus

$$R_\lambda S = R_\lambda S (T - \lambda I) R_\lambda = R_\lambda (T - \lambda I) S R_\lambda = S R_\lambda.$$

3. R_λ commutes with T by 2). Hence R_λ commutes with R_μ.

$\square$

Theorem 20.11 *Let X be a complex Banach space. Take $T \in B(X, X)$ and the polynomial*

$$p(\lambda) = \alpha_n \lambda^n + \alpha_{n-1}\lambda^{n-1} + \cdots + a_0,\ \alpha_n \neq 0.$$

Then

$$\sigma\big(p(T)\big) = p\big(\sigma(T)\big),$$

where $p(T) = \alpha_n T^n + \alpha_{n-1}T^{n-1} + \cdots + \alpha_0 I$ and $p\big(\sigma(T)\big) = \{p(\lambda) \in \mathbb{C} : \lambda \in \sigma(T)\}$.

Theorem 20.12 *Eigenvectors $\{x_1, \ldots, x_n\}$ corresponding to different eigenvalues $\lambda_1, \ldots, \lambda_n$ of a linear operator T on a vector space X are linearly independent.*

21 Spectral Representation of Bounded Self-Adjoint Operators I
(**Lecture Notes**)

We will assume that H is a complex Hilbert space and $T : H \mapsto H$ is a bounded linear operator. We recall that T is self-adjoint if $T^* = T$, that is,

$$\forall\, x, y \in H,\ \langle Tx, y\rangle = \langle x, Ty\rangle.$$

21.1 Spectral Representation of Self-Adjoint Operators in Finite Dimensions

Here we will assume that H is finite-dimensional. From Mathematics 2, Lecture 12 we know that there exists an orthonormal basis $\{e_1, \dots, e_n\}$ consisting of eigenvectors of T. In particular

$$Te_k = \lambda_k e_k,\ \forall\, k = 1, \dots, n$$

and λ_k are called eigenvalues and are real. Since $x = \sum_{k=1}^n \langle x, e_k\rangle e_k$, we get

$$Tx = T\left(\sum_{k=1}^n \langle x, e_k\rangle e_k\right) = \sum_{k=1}^n \langle x, e_k\rangle Te_k = \sum_{k=1}^n \lambda_k \langle x, e_k\rangle e_k.$$

Let us define operators $P_k x = \langle x, e_k\rangle e_k$ which are projections onto $\operatorname{span}\{e_k\}$. Then

$$Tx = \sum_{k=1}^n \lambda_k P_k x$$

or

$$T = \sum_{k=1}^n \lambda_k P_k. \tag{21.1}$$

But this formula cannot be extended to infinite-dimensional Hilbert spaces. For instance, take

$$H = L^2[0,1], \quad (Tx)(t) = tx(t).$$

Then $T^* = T$ and $\sigma(T) = \sigma_c(T) = [0,1]$. So we need to rewrite (21.1) in a more appropriate fashion. Let us assume for simplicity that $\lambda_1 < \lambda_2 < \cdots < \lambda_n$. We introduce

$$E_\lambda = \sum_{k:\lambda_k \leqslant \lambda} P_k.$$

Remark that

$$\begin{aligned}
&E_\lambda = 0,\ \lambda < \lambda_1\\
&E_\lambda = P_1,\ \lambda_1 \leqslant \lambda < \lambda_2\\
&E_\lambda = P_1 + P_2,\ \lambda_2 \leqslant \lambda < \lambda_3\\
&\quad\vdots\\
&E_\lambda = P_1 + \cdots + P_k,\ \lambda_k \leqslant \lambda < \lambda_{k+1}\\
&E_\lambda = I,\ \lambda \geqslant \lambda_n.
\end{aligned}$$

Remark also that

$$E_\lambda x = \sum_{j=1}^{k} \langle x, e_j \rangle e_j, \ \lambda_k \leqslant \lambda < \lambda_{k+1}$$

is the projection onto $\operatorname{span}\{e_1, \dots, e_k\}$. Moreover, it "increases" and is right continuous. Then, in particular, $E_{\lambda_k} = P_1 + \dots + P_k$ and $P_k = E_{\lambda_k} - E_{\lambda_{k-1}} = E_{\lambda_k} - E_{\lambda_{k-}}$. Consequently

$$T = \sum_{k=1}^{n} \lambda_k (E_{\lambda_k} - E_{\lambda_{k-}}) = \int_{-\infty}^{\infty} \lambda \, dE_\lambda,$$

where the integral is a Riemann-Stieltjes integral. We need to understand the last equality as follows:

$$\langle Tx, y \rangle = \sum_{k=1}^{n} \lambda_k \big(\langle E_{\lambda_k} x, y \rangle - \langle E_{\lambda_{k-}} x, y \rangle \big) = \int_{-\infty}^{\infty} \lambda \, d\langle E_\lambda x, y \rangle.$$

Later we will extend this formula to infinite-dimensional spaces. Namely, we will show that there exists an "increasing" right-continuous family of projection operators E_λ, $\lambda \in \mathbb{R}$ such that $E_{-\infty} = 0$ and $E_\infty = I$, and

$$T = \int_{-\infty}^{\infty} \lambda \, dE_\lambda.$$

21.2 Spectral Properties of Bounded Self-Adjoint Operators

Example 21.1 *Consider $H = L^2[0,1]$ and*

$$(Tx)(t) = tx(t), \ t \in [0,1], \ x \in L^2[0,1].$$

1. *T is self-adjoint. Indeed*

$$\langle Tx, y \rangle = \int_0^1 tx(t)\overline{y(t)} \, dt = \int_0^1 x(t)\overline{ty(t)} \, dt = \langle x, Ty \rangle.$$

2. *We want to find the spectrum and resolvent sets. Consider $T_\lambda := T - \lambda I$. We compute*

$$(T_\lambda x)(t) = (Tx - \lambda x)(t) = tx(t) - \lambda x(t) = (t - \lambda) x(t) = y(t).$$

Then

$$(R_\lambda y)(t) = \frac{1}{t - \lambda} y(t), \ t \in [0,1].$$

(a) *If $\lambda \in \mathbb{C} \setminus [0,1]$, then $\frac{1}{t-\lambda}$ is bounded so*

$$\|R_\lambda y\|^2 = \int_0^1 \frac{1}{|t-\lambda|^2} |y(t)|^2 \, dt \leqslant \sup_{t \in [0,1]} \frac{1}{|t-\lambda|^2} \int_0^1 |y(t)|^2 \, dt \leqslant \sup_{t \in [0,1]} \frac{1}{|t-\lambda|^2} \|y\|^2.$$

Hence R_λ is a bounded linear operator defined on the whole space $L^2[0,1]$, implying $\lambda \in \rho(T)$.

(b) *If $\lambda \in [0,1]$, then $\frac{1}{t-\lambda}$ is not bounded and R_λ is not defined on the whole space $L^2[0,1]$. Say for the function $y(t) = \sqrt{t-\lambda}\,\mathbb{I}_{[\lambda,1]}(t)$, $t \in [0,1]$ we get*

$$R_\lambda y(t) = \frac{\sqrt{t-\lambda}}{t-\lambda}\mathbb{I}_{[\lambda,1]}(t) = \frac{1}{\sqrt{t-\lambda}}\mathbb{I}_{[\lambda,1]}(t)$$

and

$$\|R_\lambda y\|^2 = \int_0^1 \frac{1}{\sqrt{t-\lambda}^2}\mathbb{I}_{[\lambda,1]}(t)\,dt = \int_\lambda^1 \frac{1}{t-\lambda}\,dt = \infty$$

if $\lambda < 1$. So R_λ is only defined on the set

$$\mathcal{D}(R_\lambda) = \left\{ y \in L^2[0,1] : \int_0^1 \frac{|y(t)|^2}{|t-\lambda|}\,dt < \infty \right\}.$$

One can show that $\mathcal{D}(R_\lambda)$ is dense in $L^2[0,1]$ so $\lambda \in \sigma_c(T)$. Additionally $\sigma_c(T) = [0,1]$, $\sigma_p(T) = \sigma_r(T) = \emptyset$ and $\rho(T) = \mathbb{C} \setminus [0,1]$.

Theorem 21.2 *Let H be a complex Hilbert space and $T : H \mapsto H$ a bounded self-adjoint operator.*

1. *All eigenvalues of T (if they exist) are real.*
2. *Eigenvectors corresponding to different eigenvalues of T are orthogonal.*

Theorem 21.3 (Resolvent Set) *Let H be a complex Hilbert space and $T : H \mapsto H$ a bounded self-adjoint operator. Then $\lambda \in \rho(T)$ if and only if there exists $C > 0$ such that*

$$\|Tx - \lambda x\| \geqslant C\|x\|, \forall\, x \in H.$$

Theorem 21.4 (Spectrum) *Let H be a complex Hilbert space and $T : H \mapsto H$ a bounded self-adjoint operator. Then the spectrum $\sigma(T)$ of T is real and belongs to the interval $[m, M]$, where $m = \inf\limits_{\|x\|=1}\langle Tx, x\rangle$ and $M = \sup\limits_{\|x\|=1}\langle Tx, x\rangle$. Moreover, m and M are spectral values of T.*

Theorem 21.5 (Residual Spectrum) *The residual spectrum $\sigma_r(T)$ of a bounded self-adjoint operator $T : H \mapsto H$ on a complex Hilbert space H is empty.*

21.3 Positive Operators

We introduce a partial order "$\leqslant$" on the set of self-adjoint operators on H. If T is a self-adjoint operator, then we know that $\langle Tx, x\rangle$ is real.

Definition 21.6

- *Let $T_1, T_2 : H \mapsto H$ be bounded self-adjoint operators. We write $T_1 \leqslant T_2$ if $\langle T_1 x, x\rangle \leqslant \langle T_2 x, x\rangle$ for all $x \in H$.*
- *A bounded self-adjoint operator T is called positive if $T \geqslant 0$, that is, $\langle Tx, x\rangle \geqslant 0, \forall\, x \in H$.*

We remark that a sum of positive operators is positive.

Theorem 21.7 *Every positive bounded self-adjoint operator $T : H \mapsto H$ on a complex Hilbert space H has a positive square root $T^{\frac{1}{2}}$, that is, $(T^{\frac{1}{2}})^2 = T$, which is unique. This operator commutes with every bounded linear operator on H that commutes with T.*

22 Spectral Representation of Bounded Self-Adjoint Operators II
(Lecture Notes)

22.1 Projection Operators

Let H be a Hilbert space and Y a closed subspace of H. In Lecture 18, we showed that $H = Y \oplus Y^\perp$, that is, for every $x \in H$ there exists a unique $y \in Y$ and $z \in Y^\perp$ such that $x = y + z$. We defined y as the minimizer of the function $Y \ni \tilde{y} \mapsto \|x - \tilde{y}\|$, i.e.

$$\|x - y\| = \inf_{\tilde{y} \in Y} \|x - \tilde{y}\|.$$

We define the operator $P : H \mapsto H$ such that $Px := y$, which is called an orthogonal projection on H. More specifically, P is called the projection of H onto Y.

Exercise 22.1 *Show that P is a bounded linear operator on H with $\|P\| = 1$.*

Remark 22.2 *If P is the projection of H onto Y, then $P(H) = \{Px : x \in H\} = Y$ and $\ker P = Y^\perp$.*

Theorem 22.3 *A bounded linear operator $P : H \mapsto H$ on a Hilbert space H is a projection on H if and only if $P^* = P$ and $P^2 = P$, that is, if it is self-adjoint and idempotent.*

Proof: Assume that P is a projection. Take $x \in H$. Then

$$Px = y + z = Px + 0,$$

where $y \in Y$ and $z \in Y^\perp$. Thus $P(Px) = Px$. Now take $x_1 = y_1 + z_1$ and $x_2 = y_2 + z_2$, where $y_1, y_2 \in Y$ and $z_1, z_2 \in Y^\perp$. Then

$$\langle Px_1, x_2 \rangle = \langle y_1, y_2 + z_2 \rangle = \langle y_1, y_2 \rangle + \langle y_1, z_2 \rangle = \langle y_1, y_2 \rangle$$

and

$$\langle x_1, Px_2 \rangle = \langle y_1 + z_1, y_2 \rangle = \langle y_1, y_2 \rangle + \langle z_1, y_2 \rangle = \langle y_1, y_2 \rangle.$$

This implies that $P^* = P$. Conversely, assume $P^* = P^2 = P$ is given. Set $Y := P(H)$. We need to show that if $x = y + z$, where $y \in Y$ and $z \in Y^\perp$, then $y = Px$. We write

$$x = Px + x - Px$$

and check that $x - Px \in Y^\perp$. Take $u \in P(H) \Leftrightarrow u = Pv$, $v \in H$. Compute

$$\langle u, x - Px \rangle = \langle Pv, x - Px \rangle = \langle Pv, x \rangle - \langle Pv, Px \rangle = \langle Pv, x \rangle - \langle P^2 v, x \rangle = 0.$$

$\square$

Example 22.4 *Consider $H = L^2[0, 1]$. Define for $\lambda \in [0, 1]$*

$$(Px)(t) = \mathbb{I}_{[0,\lambda]}(t) x(t) = \begin{cases} x(t), \ t \leqslant \lambda \\ 0, \ t > \lambda. \end{cases}$$

Let us check that P is a projection. According to Theorem 22.3, we have to show that $P^2 = P^* = P$. It is clear that $P^2 = P$.

Next

$$\langle Px_1, x_2\rangle = \int_0^1 (Px_1)(t)\overline{x_2(t)}\,dt = \int_0^1 \mathbb{I}_{[0,\lambda]}(t)x_1(t)\overline{x_2(t)}\,dt = \int_0^1 x_1(t)\overline{\mathbb{I}_{[0,\lambda]}(t)x_2(t)}\,dt = \langle x_1, Px_2\rangle.$$

This implies that P is a projection on H. We define

$$Y = P(H) = \{x \in L^2[0,1] : x(t) = 0,\ t \in (\lambda, 1]\}.$$

22.2 Properties of Projection Operators

Assume that H is a Hilbert space and P_1, P_2, P are projections on H. Denote $Y_i = P_i(H) = \operatorname{Im} P_i$ and $Y = P(H) = \operatorname{Im} P$.

1. P is positive and $\langle Px, x\rangle = \|Px\|^2$.
2. P_1P_2 is a projection if and only if $P_1P_2 = P_2P_1$. Then P_1P_2 projects H onto $Y_1 \cap Y_2$.
3. $P_1 + P_2$ is a projection on H if and only if $Y_1 \perp Y_2$. In this case $P_1 + P_2$ projects H onto $Y_1 \oplus Y_2$.
4. $P_2 - P_1$ is a projection on H if and only if $Y_1 \subset Y_2$.

Theorem 22.5 (Partial Order) *The following conditions are equivalent.*

1. $P_1P_2 = P_2P_1 = P_1$
2. $Y_1 \subset Y_2$
3. $\ker P_1 \supset \ker P_2$
4. $\|P_1x\| \leqslant \|P_2x\|$
5. $P_1 \leqslant P_2$ *($P_2 - P_1$ is positive)*

22.3 Spectral Family

Let H be a complex Hilbert space.

Definition 22.6

- *A real spectral family is a family $\{E_\lambda,\ \lambda \in \mathbb{R}\}$ of projections E_λ on H such that*
  1. $E_\lambda \leqslant E_\mu,\ \forall\, \lambda < \mu$
  2. $\lim\limits_{\lambda\to-\infty} E_\lambda x = 0,\ \lim\limits_{\lambda\to\infty} E_\lambda x = x,\ \forall\, x \in H$
  3. $E_{\lambda+0}x := \lim\limits_{\mu\to\lambda+0} E_\mu x = E_\lambda x,\ \forall\, x \in H$
- $\{E_\lambda,\ \lambda \in \mathbb{R}\}$ *is called a spectral family on an interval* $[a, b]$ *if* $E_\lambda = 0,\ \lambda < a$ *and* $E_\lambda = I,\ \lambda \geqslant b$.

We define a spectral family for a bounded self-adjoint operator $T : H \mapsto H$. Fix $\lambda \in \mathbb{R}$ and consider $T_\lambda = T - \lambda I$. Define the positive operator $B_\lambda = (T_\lambda^2)^{\frac{1}{2}}$. Remark that B_λ is the unique positive self-adjoint operator such that $B_\lambda^2 = T_\lambda^2$. Define $T_\lambda^+ = \frac{1}{2}(B_\lambda + T_\lambda)$ as the positive part of the operator T.

Example 22.7 *Let $H = L^2[0,1]$ and take $(Tx)(t) = tx(t)$. We want to construct E_λ. We compute*

$$(T_\lambda x)(t) = (Tx)(t) - \lambda x(t) = (t-\lambda)x(t),\ t \in [0,1].$$

Then $(T_\lambda^2 x)(t) = (t-\lambda)^2 x(t)$ and $(B_\lambda x)(t) = \sqrt{(t-\lambda)^2}x(t) = |t-\lambda|x(t)$, $t \in [0,1]$. So the positive part of T is

$$(T_\lambda^+ x)(t) = \frac{1}{2}\big((B_\lambda x)(t) + (T_\lambda x)(t)\big) = \frac{1}{2}\big(|t-\lambda|x(t) + (t-\lambda)x(t)\big) = (t-\lambda)^+ x(t),\ t \in [0,1],$$

where

$$s^+ = \begin{cases} s,\ s \geqslant 0 \\ 0,\ s < 0. \end{cases}$$

So

$$(T_\lambda^+ x)(t) = \begin{cases} x(t),\ t > \lambda \\ 0,\ t \leqslant \lambda. \end{cases}$$

Then $\ker T_\lambda^+ = \{x : T_\lambda^+ x = 0\} = \{x : x(t) = 0,\ t > \lambda\}$. From Example 22.4 we know that the projection E_λ of H onto $\ker T_\lambda^+$ is defined as

$$(E_\lambda x)(t) = \mathbb{I}_{[0,\lambda]}(t)x(t).$$

Theorem 22.8 *The family $\{E_\lambda,\ \lambda \in \mathbb{R}\}$, where E_λ is the projection of H onto T_λ^+, is the spectral family on the interval $[m, M]$, which is the smallest interval containing the spectrum of T (see Theorem 21.4).*

Theorem 22.9 (Spectral Theorem for Bounded Self-Adjoint Linear Operators) *Let $T : H \mapsto H$ be a bounded self-adjoint linear operator on a complex Hilbert space H. Then*

$$T = \int_{-\infty}^{\infty} \lambda \, dE_\lambda = \int_{m-0}^{M} \lambda \, dE_\lambda,$$

where E_λ is the spectral family associated with T. In particular

$$\langle Tx, y\rangle = \int_{-\infty}^{\infty} \lambda \, d\langle E_\lambda x, y\rangle = \int_{m-0}^{M} \lambda \, d\langle E_\lambda x, y\rangle,\ \forall\, x, y \in H.$$

Coming back to $(Tx)(t) = tx(t)$, we compute

$$(Tx)(t) = \int_{-\infty}^{\infty} \lambda \, dE_\lambda x(t) = \int_0^1 \lambda \, d\mathbb{I}_{[0,\lambda]}(t)x(t) = x(t)\int_0^1 \lambda \, d\mathbb{I}_{[0,\lambda]}(t) = x(t) \cdot t \cdot 1 = tx(t).$$

23 Compact Linear Operators (Lecture Notes)

23.1 Definition and Properties of Compact Linear Operators on Normed Spaces

Let X be a normed space. We first recall that $F \subset X$ is compact in X if every open cover of F contains a finite subcover, that is, for every family $\{G_\alpha\}$ of open sets in X such that $F \subset \bigcup_\alpha G_\alpha$ there exists $\{G_{\alpha_1}, \dots, G_{\alpha_n}\} \subset \{G_\alpha\}$ such that $F \subset \bigcup_{k=1}^n G_{\alpha_k}$.

Theorem 23.1 *F is compact in X if and only if every sequence $\{x_n\}_{n\geqslant 1} \subset F$ has a subsequence that is convergent in F.*

Definition 23.2 *A set $F \subset X$ is called relatively compact if $\overline{F}$ is compact.*

Every bounded set in a finite-dimensional normed space is relatively compact.

Exercise 23.3 *Show that F is relatively compact if and only if $\forall \{x_n\}_{n\geqslant 1} \subset F$ there exists a subsequence $\{x_{n_k}\}_{k\geqslant 1}$ such that $x_{n_k} \to x$, where x is not necessarily in F.*

Definition 23.4 *Let X and Y be normed spaces. An operator $T : X \mapsto Y$ is called a compact linear operator if T is linear and if for every bounded subset $M \subset X$ the image $T(M)$ is relatively compact.*

Theorem 23.5 (Compactness Criterion) *Let X and Y be normed spaces and $T : X \mapsto Y$ a linear operator. Then T is compact if and only if it maps every bounded sequence $\{x_n\}_{n\geqslant 1}$ in X onto a sequence $\{Tx_n\}$ in Y that has a convergent subsequence, that is, for all bounded $\{x_n\}_{n\geqslant 1}$ in X there exists a subsequence $\{Tx_{n_k}\}_{k\geqslant 1}$ of $\{Tx_n\}_{n\geqslant 1}$ such that $Tx_{n_k} \to y$ in Y.*

Theorem 23.6 *If $T : X \mapsto Y$ is bounded and $\operatorname{Im} T = T(X)$ is finite-dimensional, then T is compact.*

Example 23.7 *Take $X = Y = l^2$ over the field K. The operator T defined by*

$$Tx = (2\xi_1, \xi_2, \xi_3 + \xi_4, 0, 0, 0, \dots)$$

for $x = (\xi_k)_{k=1}^\infty$ is compact. Indeed $T(X) = \{(\eta_1, \eta_2, \eta_3, 0, 0, 0\dots) : \eta_1, \eta_2, \eta_3 \in K\}$ is a 3-dimensional subspace of l^2. By Theorem 23.6 T is compact.

Theorem 23.8 *Let $\{T_n\}_{n\geqslant 1}$ be a sequence of compact linear operators from a normed space X to a Banach space Y. If $T_n \to T$ in $B(X,Y)$, then T is compact.*

Example 23.9 *We consider $X = Y = l^2$ and*

$$Tx = \left(\xi_1, \frac{\xi_2}{2}, \frac{\xi_3}{3}, \dots\right).$$

Let us prove that T is compact. Take

$$T_n x = \left(\xi_1, \frac{\xi_2}{2}, \frac{\xi_3}{3}, \dots, \frac{\xi_n}{n}, 0, 0, \dots\right).$$

Then T_n is bounded and $\dim\big(T_n(X)\big) = n$. So by Theorem 23.6 it is compact. Let us compute

$$\|(T - T_n)x\|^2 = \left\|\left(0, 0, \dots, 0, \frac{\xi_{n+1}}{n+1}, \frac{\xi_{n+2}}{n+2}, \dots\right)\right\|^2 = \sum_{k=n+1}^\infty \frac{\xi_k^2}{k^2} \leqslant \frac{1}{(n+1)^2} \sum_{k=n+1}^\infty \xi_k^2 \leqslant \frac{1}{(n+1)^2}\|x\|^2.$$

Hence $\|T - T_n\| \leqslant \frac{1}{n+1} \to 0$, $n \to \infty$. By Theorem 23.8 T is compact.

23.2 Spectral Properties of Compact Self-Adjoint Operators

In this section we will assume that H is a separable Hilbert space.

Theorem 23.10 *Let $T : H \mapsto H$ be a bounded linear operator. The following statements are equivalent.*

1. *T is compact.*
2. *T^* is compact.*
3. *If $\langle x_n, y \rangle \to \langle x, y \rangle$, $\forall\, y \in H$, then $Tx_n \to Tx$ in H.*
4. *There exists a sequence T_n of operators of finite rank such that $\|T - T_n\| \to 0$.*

Theorem 23.11 (Hilbert-Schmidt Theorem) *Let T be a self-adjoint compact operator.*

1. *There exists an orthonormal basis consisting of eigenvectors of T.*
2. *All eigenvalues of T are real and for every eigenvalue $\lambda \neq 0$ the corresponding eigenspace is finite dimensional.*
3. *Two eigenvalues of T that correspond to different eigenvalues are orthogonal.*
4. *If T has a countable (not finite) set of eigenvalues $\{\lambda_n\}_{n\geqslant 1}$, then $\lambda_n \to 0$, $n \to \infty$.*

Corollary 23.12 *Let T be a compact self-adjoint linear operator on a complex Hilbert space H. Then there exists an orthonormal basis $\{e_k\}_{k\geqslant 1}$ such that*

$$Tx = \sum_{n=1}^{\infty} \lambda_n \langle x, e_n \rangle e_n, \; x \in H.$$

24 Unbounded Linear Operators (Lecture Notes)

24.1 Examples of Unbounded Linear Operators

Take $H = L^2(-\infty, \infty)$. Consider first the multiplication operator

$$(Tx)(t) = tx(t),\, t \in \mathbb{R}, \quad \mathcal{D}(T) = \left\{ x \in L^2(-\infty,\infty) : \int_{-\infty}^{\infty} t^2 |x(t)|^2 \, dt < \infty \right\}.$$

Remark that $\mathcal{D}(T) \neq L^2(-\infty,\infty)$. Indeed

$$x(t) = \begin{cases} \frac{1}{t},\, t \geqslant 1 \\ 0,\, t < 1 \end{cases} \in L^2(-\infty,\infty)$$

because

$$\|x\|^2 = \int_{-\infty}^{\infty} |x(t)|^2 \, dt = \int_{1}^{\infty} \frac{1}{t^2} \, dt = 1,$$

but

$$\|Tx\|^2 = \int_{-\infty}^{\infty} t^2 |x(t)|^2 \, dt = \int_{1}^{\infty} 1 \, dt = \infty.$$

Let us recall that a linear operator $T : \mathcal{D}(T) \mapsto H$ is bounded if

$$\exists\, C \geqslant 0 : \|Tx\| \leqslant C\|x\|,\, \forall\, x \in \mathcal{D}(T).$$

We take

$$x_n = \begin{cases} 1,\, n \leqslant t < n+1 \\ 0,\, \text{otherwise.} \end{cases}$$

Then

$$\|x_n\|^2 = \int_{-\infty}^{\infty} |x_n(t)|^2 \, dt = \int_{n}^{n+1} dt = 1,$$

but

$$\|Tx_n\|^2 = \int_{-\infty}^{\infty} t^2 |x_n(t)|^2 \, dt = \int_{n}^{n+1} t^2 \, dt \geqslant n^2.$$

So $\|Tx_n\|^2 \geqslant n^2 \|x_n\|$, $\forall\, n \geqslant 1$, hence T is unbounded. The differentiation operator

$$(Tx)(t) = ix'(t), \quad \mathcal{D}(T) \subset L^2(-\infty,\infty)$$

is also unbounded. Later we will explain what $\mathcal{D}(T)$ is. Here we only remark that all continuously differentiable functions with compact support and Hermite polynomials belong to $\mathcal{D}(T)$.

24.2 Symmetric and Self-Adjoint Linear Operators

Let H be a complex Hilbert space. Let $T : \mathcal{D}(T) \mapsto H$ be a densely defined ($\overline{\mathcal{D}(T)} = H$) linear operator. The adjoint operator $T^* : \mathcal{D}(T^*) \mapsto H$ of T is defined as follows. The domain $\mathcal{D}(T^*)$ of T^* consists of all $y \in H$ such that $\exists\, y^* \in H$ satisfying

$$\langle Tx, y\rangle = \langle x, y^*\rangle,\ \forall\, x \in \mathcal{D}(T). \tag{24.1}$$

For each such $y \in \mathcal{D}(T^*)$ define $T^*y := y^*$. Remark that $\mathcal{D}(T^*)$ is not necessarily equal to H. Since $\mathcal{D}(T)$ is dense in H, for every $y \in \mathcal{D}(T^*)$ there exists a unique y^* satisfying (24.1). Before we discuss the properties of adjoint operators, we will first discuss the extension of a linear operator. Let us come back to the operator

$$(T_1 x)(t) = ix'(t).$$

We can define T_1 only for functions from

$$\mathcal{D}(T_1) = C_0^1(\mathbb{R}) = \{f \in C^1(\mathbb{R}) : f = 0\,\text{outside}\,\text{some}\,\text{interval}\}.$$

Now let

$$(T_2 x)(t) = ix'(t), \quad \mathcal{D}(T_2) = \left\{ f \in C(\mathbb{R}) : \int\limits_{-\infty}^{\infty} |f|^2\,dt < \infty,\ \int\limits_{-\infty}^{\infty} |f'|^2\,dt < \infty \right\}.$$

T_1 and T_2 are different operators, but $\mathcal{D}(T_1) \subset \mathcal{D}(T_2)$ and $T_1 = T_2\big|_{\mathcal{D}(T_1)}$.

Definition 24.1 *An operator T_2 is called an extension of another operator T_1 if $\mathcal{D}(T_1) \subset \mathcal{D}(T_2)$ and $T_1 = T_2\big|_{\mathcal{D}(T_1)}$. In this case we write $T_1 \subset T_2$.*

Theorem 24.2 *Let $S : \mathcal{D}(S) \mapsto H$ and $T : \mathcal{D}(T) \mapsto H$ be densely defined linear operators.*

1. *If $S \subset T$ then $T^* \subset S^*$.*
2. *If $\mathcal{D}(T^*)$ is dense in H, then $T \subset (T^*)^*$.*
3. *If T is injective and $\operatorname{Im} T$ is dense in H, then T^* is injective and $(T^*)^{-1} = (T^{-1})^*$.*

Definition 24.3 *Let $T : \mathcal{D}(T) \mapsto H$ be a densely defined linear operator on H. T is called a symmetric linear operator if*

$$\langle Tx, y\rangle = \langle x, Ty\rangle,\ \forall\, x, y \in \mathcal{D}(T).$$

Remark T being symmetric does not imply that $T = T^*$. Indeed, take

$$(Tx)(t) = ix(t), \quad \mathcal{D}(T) = C_0(\mathbb{R}).$$

Then

$$\begin{aligned}\langle Tx, y\rangle &= \int\limits_{-\infty}^{\infty} ix'(t)\overline{y(t)}\,dt = \int\limits_{-\infty}^{\infty} \overline{iy(t)}dx(t) = iy(t)x(t)\Big|_{-\infty}^{\infty} - \int\limits_{-\infty}^{\infty} x(t)\,d(\overline{iy(t)}) = 0 - 0 - \int\limits_{-\infty}^{\infty} x(t)\overline{iy'(t)}\,dt \\ &= \int\limits_{-\infty}^{\infty} x(t)\overline{iy'(t)}\,dt = \langle x, Ty\rangle,\ \forall\, x, y \in \mathcal{D}(T) = C_0(\mathbb{R}).\end{aligned}$$

However, $T^* \neq T$. For instance, $y(t) = e^{-t^2}$, $t \in \mathbb{R}$ does not belong to $\mathcal{D}(T) = C_0(\mathbb{R})$ but $y \in \mathcal{D}(T^*)$ because for $y^*(t) = i(-2t)e^{-t^2}$ one has

$$\langle Tx, y \rangle = \langle x, y^* \rangle, \, \forall \, x \in \mathcal{D}(T).$$

Lemma 24.4 *A densely defined linear operator T is symmetric if and only if $T \subset T^*$.*

Definition 24.5 *Let $T : \mathcal{D}(T) \mapsto H$ be a densely defined linear operator. T is called self-adjoint if $T = T^*$.*

Remark 24.6 *Every self-adjoint operator is symmetric but not every symmetric operator is self-adjoint.*

24.3 Closed Linear Operators

Definition 24.7 *Let $T : \mathcal{D}(T) \mapsto H$ be a linear operator, where $\mathcal{D}(T) \subset H$. T is called a closed linear operator if its graph*

$$\mathrm{Gr}(T) = \{(x, y) : x \in \mathcal{D}(T), \, y = Tx\}$$

is closed in $H \times H$, where the norm on $H \times H$ is defined as

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

Theorem 24.8 *Let $T : \mathcal{D}(T) \mapsto H$ be a linear operator, where $\mathcal{D}(T) \subset H$.*

1. *T is closed if and only if $x_n \to x$, $x_n \in \mathcal{D}(T)$ and $Tx_n \to y$ imply $x \in \mathcal{D}(T)$ and $Tx = y$.*

2. *If T is closed and $\mathcal{D}(T)$ is closed, then T is bounded.*

3. *Let T be bounded. Then T is closed if and only if $\mathcal{D}(T)$ is closed.*

Exercise 24.9 *Show that the multiplication operator is closed.*

Theorem 24.10 *Let T be a densely defined operator on H. Then the adjoint operator T^* is closed.*

Definition 24.11

- *If a linear operator T has an extension T_1 which is a closed linear operator, then T is called closable.*

- *If T is closable, then there exists a minimal closed operator $\overline{T}$ satisfying $T \subset \overline{T}$. The operator $\overline{T}$ is called the closure of T.*

Theorem 24.12 *Let $T : \mathcal{D}(T) \mapsto H$ be a densely defined linear operator. If T is symmetric, its closure $\overline{T}$ exists and is unique.*

25 Spectral Representation of Unbounded Self-Adjoint Operators, Curves in $\mathbb{R}^3$ (Lecture Notes)

25.1 Spectral Representation

Let H be a complex Hilbert space. We recall that a bounded operator $U : H \mapsto H$ is called unitary if $U^* = U^{-1}$.

Theorem 25.1 *Let $U : H \mapsto H$ be a unitary operator. Then there exists a spectral family $\{E_\theta\}_\pi$ on $[-\pi, \pi]$ such that*

$$U = \int_{-\pi}^{\pi} e^{i\theta}\, dE_\theta, \tag{25.1}$$

where the integral is understood in the sense of uniform operator convergence.

Proof Idea: One can show that there exists a bounded self-adjoint linear operator S with $\sigma(S) \subset [-\pi, \pi]$ such that

$$U = e^{iS} = \cos S + i \sin S.$$

Let $\{E_\theta\}$ be a spectral family for S on $[-\pi, \pi]$. Then

$$S = \int_{-\pi}^{\pi} \theta\, dE_\theta.$$

Hence

$$U = e^{iS} = \int_{-\pi}^{\pi} \cos\theta\, dE_\theta + i \int_{\pi}^{\pi} \sin\theta\, dE_\theta = \int_{-\pi}^{\pi} e^{i\theta}\, dE_\theta.$$

□

Let $T : \mathcal{D}(T) \mapsto H$ be a self-adjoint linear operator, where $\mathcal{D}(T)$ is dense in H and T may be unbounded. We take a new operator

$$U = (T - iI)(T + iI)^{-1}$$

called the Cayley transform of T. It is defined on the whole Hilbert space since $-i \notin \sigma(T) \subseteq \mathbb{R}$. One can also check that it is unitary and

$$T = i(I + U)(I - U)^{-1}.$$

Theorem 25.2 (Spectral Representation for Unbounded Self-Adjoint Operators) *Let $T : \mathcal{D}(T) \mapsto H$ be a self-adjoint linear operator and let $\mathcal{D}(T)$ be dense in H. Let U be the Cayley transform of T and $\{\tilde{E}_\theta\}$ a spectral family in the spectral representation (25.1) for $-U$. Then*

$$T = \int_{-\pi}^{\pi} \tan\frac{\theta}{2}\, d\tilde{E}_\theta = \int_{-\infty}^{\infty} \lambda\, dE_\lambda,$$

where $E_\lambda = \tilde{E}_{2\arctan\lambda}$, $\lambda \in \mathbb{R}$.

We remark that $T = i(I + U)(I - U)^{-1} = f(-U)$, where $f(\theta) = i\frac{1-\theta}{1+\theta}$. Let

$$-U = \int_{-\pi}^{\pi} e^{i\theta}\, d\tilde{E}_\theta.$$

Then

$$T = \int_{-\pi}^{\pi} f(e^{i\theta})\,d\tilde{E}_\theta = \int_{-\pi}^{\pi} i\frac{1-e^{i\theta}}{1+e^{i\theta}}\,d\tilde{E}_\theta = \int_{-\pi}^{\pi} i\frac{(1-\cos\theta)-i\sin\theta}{(1+\cos\theta)+i\sin\theta}\,d\tilde{E}_\theta$$
$$= \int_{-\pi}^{\pi} i\frac{-2i\sin\theta}{2+2\cos\theta}\,d\tilde{E}_\theta = \int_{-\pi}^{\pi} \tan\frac{\theta}{2}\,d\tilde{E}_\theta.$$

Example 25.3 (Spectral Representation of the Multiplication Operator) *Let $H = L^2(-\infty,\infty)$ be taken over $\mathbb{C}$ and take*

$$(Tx)(t) = tx(t),\ t\in\mathbb{R}, \quad \mathcal{D}(T) = \left\{ x\in L^2(-\infty,\infty) : \int_{-\infty}^{\infty} t^2|x(t)|^2\,dt < \infty \right\}.$$

Then T is self-adjoint and the spectral family associated with T is

$$(F_\lambda x)(t) = \begin{cases} x(t),\ t<\lambda \\ 0,\ t\geqslant\lambda. \end{cases}$$

25.2 Some Definitions

We consider a map $x : I \mapsto \mathbb{R}^3$, where $x(t) = \big(x_1(t), x_2(t), x_3(t)\big)$, $t\in I = [a,b]$. We assume that x_i are r times continuously differentiable and $x'(t) = \big(x_1'(t), x_2'(t), x_3'(t)\big) \neq 0$, $\forall\, t\in I$. The set of points represented by x we will call a curve. A curve can have different representations. Indeed, let us consider a transformation $t = t(t^*)$ such that

1. $t : [a^*, b^*] \mapsto [a,b]$, $t(a^*) = a$, $t(b^*) = b$ (or $t(a^*) = b$, $t(b^*) = a$)

2. the function is r times continuously differentiable

3. $\dfrac{dt}{dt^*}$ is different from zero on I^*

Then $x\big(t(t^*)\big) =: x(t^*)$ is another parametrization of the curve x.

Example 25.4

1. *$x(t) = (a_1+b_1t, a_2+b_2t, a_3+b_3t)$ describes a line passing through (a_1,a_2,a_3) and parallel to (b_1,b_2,b_3).*

2. *$x(t) = (a\cos t, b\sin t, 0)$ describes an ellipse with axes a and b in the plane spanned by x_1 and x_2.*

3. *$x(t) = (r\cos t, r\sin t, ct)$, $c\neq 0$ describes a circular helix.*

We recall that

$$L = \int_a^b |x'(t)|\,dt$$

is the total length of a curve while

$$s(t) = \int_{t_0}^{t} |x'(r)|\,dr$$

is the length of the part of the curve between t_0 and t.

Since $s'(t) = |x'(t)| > 0$, the function $s : [a, b] \mapsto [a^*, b^*]$ is strictly increasing. This means that the inverse map $t = t(s)$, $s \in [-L_1, L_2]$ exists. The parametrization $x(s) := x\big(t(s)\big)$ is called a natural parametrization. Remark that the point t_0 for $s = 0$ is chosen arbitrarily. For a natural parametrization we use the notation

$$\dot{x} = \frac{dx}{ds}, \quad \ddot{x} = \frac{d^2x}{ds^2}$$

and for an arbitrary parametrization we use the notation

$$x' = \frac{dx}{dt}, \quad x'' = \frac{d^2x}{dt^2}.$$

We remark that

$$\dot{x}(s) = \frac{dx}{ds} = \frac{dx}{dt}\frac{dt}{ds} = x'(t)\frac{1}{|x'\big(t(s)\big)|} \Rightarrow |\dot{x}(s)| = 1.$$

Lemma 25.5 *Let x be naturally parametrized. Then $|\dot{x}(s)| = 1$.*

26 Curves in $\mathbb{R}^3$ (Lecture Notes)

26.1 Tangent, Principal Normal and Binormal Vectors

Let $x : I \mapsto \mathbb{R}^3$, where $x(t) = \big(x_1(t), x_2(t), x_3(t)\big)$, $t \in I = [a, b]$, be a curve C. Let $x = x(s)$ be a natural parametrization of x, that is $x(s) = x\big(t(s)\big)$, where $t = t(s)$ is the inverse function to

$$s = \int_{t_0}^{t} |x'(t)|\, dt = \int_{t_0}^{t} \sqrt{\big(x_1'(u)\big)^2 + \big(x_2'(u)\big)^2 + \big(x_3'(u)\big)^2}\, du.$$

Remark 26.1 *$x = x(t)$ is the natural parametrization ($s(t) = t$) if and only if $|x'(t)| = 1$ for every $t \in I$.*

Definition 26.2

- *The vector*
$$\vec{t}(s) = \lim_{h\to 0} \frac{x(s+h) - x(s)}{h} = \frac{dx(s)}{ds} = \dot{x}(s)$$
is called the unit tangent vector to the curve C at the point $x(s)$. Remark that if $x(t)$ is not a natural parametrization, then
$$\vec{t}(t) = \frac{x'(t)}{|x'(t)|}.$$

- *The plane orthogonal to $\vec{t}(s)$ and passing through $x(s)$ is called the normal plane. It can be written in the form*
$$\dot{x}(s) \cdot z + x(s) = 0,\ z = (z_1, z_2, z_3) \in \mathbb{R}^3.$$

Example 26.3 *We consider the circular helix $x(t) = (r\cos t, r\sin t, ct)$, $t \in I$, $c \neq 0$. We first calculate*

$$x'(t) = (-r\sin t, r\cos t, c) \Rightarrow |x'(t)| = \sqrt{r^2\sin^2 t + r^2\cos^2 t + c^2} = \sqrt{r^2 + c^2},$$

$$s(t) = \int_0^t \sqrt{r^2 + c^2}\, dt = \sqrt{r^2 + c^2}\, t \Rightarrow t(s) = \frac{1}{\sqrt{r^2 + c^2}} s = \frac{s}{w},\ w := \sqrt{r^2 + c^2}.$$

Thus we obtain the natural parametrization and unit tangent vector

$$x(s) = \left(r\cos\frac{s}{w}, r\sin\frac{s}{w}, \frac{c}{w}s\right), \quad \vec{t}(s) = \left(-\frac{r}{w}\sin\frac{s}{w}, \frac{r}{w}\cos\frac{s}{w}, \frac{c}{w}\right).$$

Definition 26.4

- *The rate of change of the unit tangent vector*
$$\kappa(s) = |\dot{\vec{t}}(s)| = |\ddot{x}(s)|$$
is called the curvature of the curve C at the point $x(s)$.

- *The plane passing through $x(s)$ and parallel to $\dot{x}(s)$ and $\ddot{x}(s)$ (if $\ddot{x}(s) \neq 0$) is called the osculating plane.*

- *The vector*
$$\vec{p}(s) = \frac{\dot{\vec{t}}(s)}{|\dot{\vec{t}}(s)|} = \frac{\ddot{x}(s)}{|\ddot{x}(s)|} = \frac{\ddot{x}(s)}{\kappa(s)}$$
is called the principal normal to the curve C at the point $x(s)$.

Example 26.3 *Coming back to the circular helix, we have*

$$\ddot{x}(s) = \left(-\frac{r}{w^2}\cos\frac{s}{w}, -\frac{r}{w^2}\sin\frac{s}{w}, 0\right) \Rightarrow \kappa(s) = |\ddot{x}(s)| = \frac{r}{w^2},$$

$$\vec{p}(s) = \left(-\cos\frac{s}{w}, -\sin\frac{s}{w}, 0\right).$$

Definition 26.5

- *The vector*

$$\vec{b}(s) = \vec{t}(s) \times \vec{p}(s)$$

is called the binormal vector of C at the point $x(s)$.

- *The plane parallel to $\vec{t}(s)$ and $\vec{b}(s)$ and passing through $x(s)$ is called the rectifying plane.*

Example 26.3 *For the circular helix, we have*

$$\vec{p}(s) = \left(-\cos\frac{s}{w}, -\sin\frac{s}{w}, 0\right),$$

$$\vec{b}(s) = \begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ -\frac{r}{w}\sin\frac{s}{w} & \frac{r}{w}\cos\frac{s}{w} & \frac{c}{w} \\ -\cos\frac{s}{w} & -\sin\frac{s}{w} & 0 \end{vmatrix} = \left(\frac{c}{w}\sin\frac{s}{w}, -\frac{c}{w}\cos\frac{s}{w}, \frac{r}{w}\right).$$

Next we introduce torsion. Roughly speaking, the torsion measures the rate of rotation of the curve, that is, the rate of change of the osculating plane. Assume that $\kappa(s) > 0$.

Definition 26.6 *The scalar*

$$\tau(s) = -\vec{p}(s) \cdot \dot{\vec{b}}(s) = \frac{\left(\dot{x}(s), \ddot{x}(s), \dddot{x}(s)\right)}{|\ddot{x}(s)|}$$

is called the torsion of the curve C at the point $x(s)$.

Example 26.3 *For the circular helix, we calculate the torsion:*

$$\dot{\vec{b}}(s) = \left(\frac{c}{w^2}\cos\frac{s}{w}, \frac{c}{w^2}\sin\frac{s}{w}, \frac{1}{w}\right) \Rightarrow \tau(s) = \frac{c}{w^2}.$$

Theorem 26.7 *A curve (of class $r \geqslant 3$) with $\kappa(s) \neq 0$, $\forall\, s$ is a helix if and only if $\tau = \text{const.}$, $\kappa = \text{const.}$.*

We remark that the vectors $\vec{t}, \vec{p}, \vec{b}$ form a basis. Consequently, every vector can be rewritten as a linear combination of these vectors. In particular, we obtain the Frenet formulae:

$$\begin{pmatrix} \dot{\vec{t}} \\ \dot{\vec{p}} \\ \dot{\vec{b}} \end{pmatrix} = \begin{pmatrix} 0 & \kappa & 0 \\ -\kappa & 0 & \tau \\ 0 & -\tau & 0 \end{pmatrix} \begin{pmatrix} \vec{t} \\ \vec{p} \\ \vec{b} \end{pmatrix}.$$

Theorem 26.8 *A curve with $\kappa \neq 0$ belongs to a plane if and only if $\tau(s) = 0$, $\forall\, s$.*

Theorem 26.9 *If the curve C is given by an arbitrary parametrization, then*

$$\vec{t}(t) = \frac{x'(t)}{|x'(t)|},\ \vec{p}(t) = \vec{b}(t) \times \vec{t}(t),\ \vec{b}(t) = \frac{x'(t) \times x''(t)}{|x'(t) \times x''(t)|},\ \kappa(t) = \frac{x'(t) \times x''(t)}{|x'(t)|^3},\ \tau(t) = \frac{\left(x'(t), x''(t), x'''(t)\right)}{|x'(t) \times x''(t)|^2}$$

26.2 Topological Spaces

Let X be a set and τ a class of subsets which satisfies the following properties:

(T1) $\emptyset \in \tau$, $X \in \tau$

(T2) Any arbitrary (finite or infinite) union of sets from τ belongs to τ

(T3) The intersection of a finite number of sets from τ belongs to τ

Definition 26.10 *The pair (X, τ) is called a topological space, where τ satisfies (T1) - (T3). Sets from τ are called open sets.*

Example 26.11

1. *Let X be a metric space and τ a family of all open subsets from X. Then X is a topological space.*
2. *Take $X = [0, 1]$ and $\tau = \{[0, b) : b \in (0, 1)\} \cup \{\emptyset, X\}$. Then X is also a topological space.*

Definition 26.12 *A topological space (X, τ) is called a Hausdorff space if $\forall\, x, y \in X$, $\exists\, A, B \in \tau$ such that $A \cap B = \emptyset$ and $x \in A$, $y \in B$.*

Definition 26.13 *Let (X, τ) and (X', τ') be topological spaces. A function $f : X \mapsto X'$ is continuous if $f^{-1}(A) \in \tau$, $\forall\, A \in \tau'$.*

Remark 26.14 *If X, X' are metric spaces, then $f : X \mapsto X'$ is continuous as a function between metric spaces if and only if it is continuous according to Definition 26.13.*

Definition 26.15 *A map $f : X \mapsto X'$ is called a homeomorphism if f is a bijection and f and f^{-1} are continuous.*

Let us consider a way of constructing a topology. Assume that B is a collection of subsets from X such that B covers X and for all $B_1, B_2 \in B$ and $x \in B_1 \cap B_2$, there exists $B_3 \in B$ such that $B_3 \subset B_1 \cap B_2$. Then the collection of arbitrary (finite or infinite) unions of subsets from B is a topology on X. This topology is called the topology generated by B and B is called the base of this topology.

27 Differentiable Manifolds (Lecture Notes)

27.1 Main Definitions

Assume that M is a connected, Hausdorff topological space. Connected means that no open sets U, V exist such that $M = U \cup V$ and $U \cap V = \emptyset$.

Definition 27.1

- *An m-dimensional coordinate chart on M is a pair (U, φ), where U is an open subset of M (called the domain of the coordinate chart) and φ is a homeomorphism of U onto an open subset of $\mathbb{R}^m$.*
- *If $U = M$, then the coordinate chart is globally defined, otherwise it is locally defined.*

Definition 27.2 *Let (U_1, φ_1) and (U_2, φ_2) be m-dimensional coordinate charts with $U_1 \cap U_2 \neq \emptyset$. Then the function*

$$\varphi_2 \circ \varphi_1^{-1} : \varphi_1(U_1 \cap U_2) \mapsto \varphi_2(U_1 \cap U_2)$$

is called the overlap function.

Definition 27.3

- *An atlas of dimension m on M is a family of m-dimensional coordinate charts $\{(U_i, \varphi_i)\}_{i \in I}$, where I is an index set, such that M is covered by $\{U_i\}_{i \in I}$ and each overlap function $\varphi_j \circ \varphi_i^{-1}$, $i, j \in I$ is infinitely differentiable.*
- *An atlas is said to be complete if it is maximal, that is, it is not contained in any other atlas.*
- *For a complete atlas, the family $(U_i, \varphi_i)_{i \in I}$ is called a differential structure on M of dimension m. The topological space M is called a differentiable manifold.*

Definition 27.4 *A point $p \in U \subset M$ has the coordinates $\big(\varphi^1(p), \dots, \varphi^m(p)\big)$ with respect to the chart (U, φ). The coordinates of p are often written as $\big(x^1(p), \dots, x^m(p)\big)$.*

27.2 Some Example of Differentiable Manifolds

1. The circle $S^1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ is a differentiable manifold. One can define the differential structure on S^1 by introducing the following charts:

$$U_1 := \{(x, y) \in S^1 : x > 0\},\ \varphi_1(x, y) := y$$

$$U_2 := \{(x, y) \in S^1 : x < 0\},\ \varphi_2(x, y) := y$$

$$U_1 := \{(x, y) \in S^1 : y > 0\},\ \varphi_3(x, y) := x$$

$$U_1 := \{(x, y) \in S^1 : y < 0\},\ \varphi_4(x, y) := x.$$

Let us show that the overlap functions are from C^∞. Consider the overlap of U_1 and U_3:

$$\varphi_1(x, y) = y, \quad \varphi_3^{-1}(x) = \big(x, \sqrt{1 - x^2}\big).$$

Hence

$$\varphi_1 \circ \varphi_3^{-1} = \sqrt{1 - x^2},\ x \in (0, 1)$$

is infinitely differentiable on $(0, 1)$.

2. The n-sphere $S^n = \{x \in \mathbb{R}^{n+1} : |x|^2 = 1\}$ is a differentiable manifold. The differential structure can be given by means of a stereographic projection from the north and south poles (φ_1 and φ_2 respectively):

$$\varphi_1(x_1, \ldots, x_{n+1}) := \left(\frac{x_1}{1 - x_{n+1}}, \frac{x_2}{1 - x_{n+1}}, \ldots, \frac{x_n}{1 - x_{n+1}} \right) \in \mathbb{R}^n$$

$$\varphi_2(x_1, \ldots, x_{n+1}) := \left(\frac{x_1}{1 + x_{n+1}}, \frac{x_2}{1 + x_{n+1}}, \ldots, \frac{x_n}{1 + x_{n+1}} \right) \in \mathbb{R}^n.$$

27.3 Differentiable Maps and Tangent Spaces

Definition 27.5

- *A local representative of a function $f : M \mapsto N$ with respect to coordinate charts (U, φ) and (V, ψ) on M and N respectively is the map*

$$\psi \circ f \circ \varphi^{-1} : \varphi(U) \mapsto \mathbb{R}^n.$$

- *A map $f : M \mapsto N$ is a C^2-function if for all covers of M and N, the local representatives are r times continuously differentiable. If f is a C^1 function, then f is called differentiable. If f is a C^∞ function, then f is called smooth.*

Definition 27.6

- *A curve on a manifold M is a smooth map σ from some interval $(-\epsilon, \epsilon)$ of the real line into M.*

- *Two curves σ_1 and σ_2 are tangent at a point p in M if $\sigma_1(0) = \sigma_2(0) = p$ and if in some local coordinate system $(x^1, \ldots, x^m)$ around the point p*

$$\frac{dx^i}{dt}\big(\sigma_1(t)\big)\bigg|_{t=0} = \frac{dx^i}{dt}\big(\sigma_2(t)\big)\bigg|_{t=0}, \; i = 1, \ldots, m.$$

  Remark that if σ_1 and σ_2 are tangent in one coordinate system, then they are tangent in any other coordinate system.

- *A tangent vector at $p \in M$ is an equivalence class of tangent curves in p. Then tangent class will be denoted by $[\sigma]$.*

A tangent vector $v = [\sigma]$ can be used as a directional derivative for functions $f : M \mapsto \mathbb{R}$ by defining

$$v(f) := \frac{df\big(\sigma(t)\big)}{dt}\bigg|_{t=0},$$

where σ is any curve from $[\sigma]$. Remark that v does not depend on the choice of σ from $[\sigma]$. Indeed, take any chart (U, φ) such that $p \in U$. Let σ_1 and σ_2 be two curves such that

$$\frac{dx^i}{dt}\big(\sigma_1(t)\big)\bigg|_{t=0} = \frac{dx^i}{dt}\big(\sigma_2(t)\big)\bigg|_{t=0}, \quad \sigma_1(0) = \sigma_2(0) = p.$$

Then

$$\left.\frac{df\big(\sigma_1(t)\big)}{dt}\right|_{t=0} = \left.\frac{d\big(f\circ\varphi^{-1}(\varphi\circ\sigma_1)\big)}{dt}\right|_{t=0} = \left.\sum_{i=1}^{m}\frac{\partial(f\circ\varphi^{-1})}{\partial x^i}\frac{dx^i(\sigma_1)}{dt}\right|_{t=0}$$
$$= \left.\sum_{i=1}^{m}\frac{\partial(f\circ\varphi^{-1})}{\partial x^i}\frac{dx^i(\sigma_2)}{dt}\right|_{t=0} = \left.\frac{df\big(\sigma_2(t)\big)}{dt}\right|_{t=0}.$$

Definition 27.7

The tangent space T_pM to M at a point $p \in M$ is the set of all tangent vectors at the point p. The tangent bundle TM is defined as

$$TM = \bigcup_{p\in M} T_pM.$$

Example 27.8 *Take $M = S^n = \{x \in \mathbb{R}^{n+1} : |x|^2 = 1\}$. Then*

$$T_pS^n = \{v \in \mathbb{R}^{n+1} : p\cdot v = 0\}, \quad TS^n = \{(p,v) \in \mathbb{R}^{n+1}\times\mathbb{R}^{n+1} : |p|^2 = 1,\, p\cdot v = 0\}$$

27.4 The Vector Space Structure on T_pM

The set T_pM can be made a vector space. Let v_1 and v_2 be two tangent vectors from T_pM. Let σ_1 and σ_2 be two representative curves for v_1 and v_2 respectively. σ_1 and σ_2 cannot be added directly since M is not a vector space, but we can consider the sum

$$t \mapsto \varphi\circ\sigma_1(t) + \varphi\circ\sigma_2(t),$$

which is a curve in $\mathbb{R}^m$. So we can define

$$v_1 + v_2 := \big[\varphi^{-1}\circ(\varphi\circ\sigma_1 + \varphi\circ\sigma_2)\big], \quad rv_1 := \big[\varphi^{-1}\circ(r\cdot\varphi\circ\sigma_1)\big],\, r\in\mathbb{R}. \tag{27.1}$$

This definition is independent of the choice of chart (U,φ) and representatives σ_1 and σ_2 of the tangent vectors v_1 and v_2. Under the operations defined by (27.1), the set T_pM is a vector space. A tangent vector also can be defined as a derivative:

$$v(f) = \left.\frac{df\big(\sigma(t)\big)}{dt}\right|_{t=0},$$

where $[\sigma] = v$.

Definition 27.9

- *A derivation at a point $p \in M$ is a map $v : C^\infty(M) \mapsto \mathbb{R}$ such that*
  1. *$v(f+g) = v(f) + v(g)$, $v(rf) = rv(f)$, $r\in\mathbb{R}$, $f,g \in C^\infty(M)$*
  2. *$v(fg) = f(p)v(g) + g(p)v(f)$, $\forall\, f,g \in C^\infty(M)$*
- *The set of all derivations is denoted by D_pM.*

Theorem 27.10 *The linear map $L : T_pM \mapsto D_pM$ defined by*

$$L(v)(f) := \left.\frac{df\big(\sigma(t)\big)}{dt}\right|_{t=0},\ [\sigma] = v$$

is an isomorphism.