Functional Analysis, Fall 2019

泛函分析笔记

A ruined gambler

May 10, 2020

Contents

1 Normed Linear Spaces **6**

1.1 Fundamentals . . . 6

Equivalence between norms . . . 7

Banach spaces . . . 8

Examples . . . 9

1.2 Bounded linear operators . . . 13

Examples . . . 15

1.3 Finite dimensional space . . . 18

1.4 Quotient space . . . 21

1.5 Other topics . . . 24

Separable spaces, Schauder bases . . . 24

Direct sum . . . 25

1.6 Hahn-Banach theorems . . . 26

Extension theorems . . . 26

Geometric Hahn-Banach theorems . . . 30

2 Hilbert Space **36**

2.1 Fundamentals . . . 36

2.2 Orthogonality . . . 42

Best approximation . . . 42

Orthogonal decomposition . . . 44

2.3 Orthonormal Bases . . . 46

2.4 Isometric Isomorphism 50

3 TVS & LCS **53**

3.1 Elementary properties 54

Invariance of the local base 54

Separation properties 56

Types of topological vector spaces 57

3.2 Locally convex spaces 59

3.3 Linear mappings 63

3.4 Finite-dimensional spaces 65

3.5 Quotient spaces 67

3.6 Hahn-Banach theorems 70

Separation theorems 70

Geometric interpretations 73

Extension theorem 74

3.7 Weak topologies on LCS 75

Duality 76

Weak closure 77

Annihilators 79

3.8 Appendix 2: Metrization of TVS 81

4 Opertators and Functionals **83**

4.1 Convergence of operaters 83

4.2 Applications of Baire's Category 86

Principle of uniform boundedness 86

The open mapping theorem 89

The closed graph theorem 90

4.3 Linear functionals 92

Hahn-Banach Theorem in Normed linear space 93

4.4 Dual Spaces 95

Represent of dual space 96

The Dual of Subspace and Quotient Space 98
4.5 Bidual space and Reflexivity 100
4.6 Weak topologies 104
Convergence 105
Compact, sequentially compact sets in weak* topology 107
Compact, sequentially compact sets in weak topology : reflexivity revisited 109
4.7 Uniformly convex space 114
4.8 Projections, Direct Sum Decomposition 119
4.9 Adjoints 122
The closed range theorem 125
4.10 Compact Operators 128

5 Operators on Hilbert Space 129
5.1 The Adjoint of an Operator 130
5.2 Three operators on Hilbert space 134
Self-adjoint operator 134
Normal operator 136
Unitary 137
5.3 Projections and Idempotents; Invariant and Reducing Subspaces 138

A 143
A.1 Metric space 143
Compact Subset of a Metric Space 143
Baire's Category Theorem 145
A.2 不可数级数的和 147

Baisc Linear algebra

Let X and Y be linear space over the same field $\mathbb{F}$. A **linear operator** from X into Y is a mapping $T : X \to Y$ such that

$$T(\alpha x_1 + \beta x_2) = \alpha T x_1 + \beta T x_2, \text{ for all } x_1, x_2 \in X \text{ and all } \alpha, \beta \in \mathbb{F}.$$

We shall denote by $\mathcal{L}(X, Y)$ the set of all linear operators from X into Y, and we shall write $\mathcal{L}(X)$ for $\mathcal{L}(X, X)$.

The **range** of a linear operator $T : X \to Y$ is the set

$$\operatorname{ran}(T) \coloneqq \{y \in Y \mid y = Tx \text{ for some } x \in X\} = T(X)$$

and the **kernel** or the **null space** of T is the set

$$\ker(T) \coloneqq \{x \in X \mid Tx = 0\} = T^{-1}(0)$$

$\ker(T)$ is a linear subspace of X and $\operatorname{ran}(T)$ is a linear subspace of Y. We say T is invertible if and only if T is both injective and bijective.

Proposition 0.1. *Let X and Y be linear space over $\mathbb{F}$. Suppose that $T \in \mathcal{L}(X, Y)$ is an invertible. Then*

(a) T^{-1} *is also invertible and* $T^{-1} \in \mathcal{L}(Y, X)$, $(T^{-1})^{-1} = T$.

(b) $T^{-1}T = I_X$ *and* $T^{-1}T = I_Y$.

Let X and Y be linear sapce over $\mathbb{F}$. For all $T, S \in \mathcal{L}(X, Y)$ and $\alpha \in \mathbb{F}$, define the operations of addition and scalar mutiplication as follows:

$$(T + S)(x) = Tx + Sx, \text{ for each } x \in X$$
$$(\alpha T)(x) = \alpha T x, \text{ for each } x \in X.$$

Then $\mathcal{L}(X, Y)$ is a linear space over $\mathbb{F}$.

Lemma 0.2. *X is a vector space over $\mathbb{F}$. $\ell, \ell_1, \cdots, \ell_n$ are linear functionals on X. Then $\ell \in \operatorname{span}\{\ell_1, \cdots, \ell_n\}$ if and only if*

$$\ker(\ell) \subset \bigcap_{k=1}^{n} \ker(\ell_k)\,. \tag{1}$$

Proof. We only prove the sufficiency. Define $T : X \to \mathbb{F}^n$ by

$$Tx = (\ell_1 x, \cdots, \ell_n x)\ .$$

If $Tx = Tx'$, then $\ker(\ell) \subset \cap_k \ker(\ell_k)$ implies that $\ell x = \ell x'$. Thus we can define a linear functional f on $\operatorname{ran}(T)$ by letting

$$f(Tx) = \ell x \quad \text{for each } Tx \in \operatorname{ran}(T)\,.$$

We can extend f to a linear functional on $\mathbb{F}^n$. This means that there exist scalars $\alpha_i, i = 1, 2 \cdots, n$ such that

$$f\left(u_1, \ldots, u_n\right) = \alpha_1 u_1 + \cdots + \alpha_n u_n\,.$$

Thus

$$\ell x = f(Tx) = f\left(\ell_1 x, \ldots, \ell_n x\right) = \sum_{i=1}^{n} \alpha_i \ell_i x\,. \qquad \square$$

Chapter 1

Normed Linear Spaces

1.1 Fundamentals

Definition 1.1. A **seminorm** on a linear space X is a nonnegative real-valued function $p : X \to [0, \infty)$ which satisfies the following properties.

(a) Absolute homogeneity: $p(\lambda x) = |\lambda| p(x)$ for all $\lambda \in \mathbb{F}$ and $x \in X$.

(b) Trigonometric inequality: $p(x + y) \leqslant p(x) + p(y)$ for all $x, y \in X$.

It follows from (a) that $p(0) = 0$. A **norm** is a seminorm p satisfying:

(c) Positive definiteness: $p(x) = 0$ if and only if $x = 0$.

Usually a norm is denoted by $\|\cdot\|$, A *normed linear space* is a pair $(X, \|\cdot\|)$, where X is a linear space and $\|\cdot\|$ a norm on X. If $(X, \|\cdot\|)$ is a normed linear space, then

$$d(x, y) \coloneqq \|x - y\| \text{ for all } x, y \in X$$

defines a metric on X. Such a metric d is said to be induced or generated by the norm $\|\cdot\|$. Thus, every normed linear space is a metric space, and unless otherwise specified, we shall henceforth regard any normed linear space as a metric space with respect to the metric induced by its norm.

Remark 1.1. It's natural to ask when a metric is induced by a norm or when a linear space X equipped with a metric d become a normed linear space? It's not hard to prove the following proposition, which shows that a metric d on linear space X is induced by a norm if and only if d is *translation-invariant* and *absolutely homogeneous*.

Exercise 1.1. If d is a metric on a linear space X satisfying for all $x, y, z \in X$ and $\lambda \in \mathbb{F}$,

(a) $d(x, y) = d(x + z, y + z)$ (translation invariance)

(b) $d(\lambda x, \lambda y) = |\lambda| d(x, y)$ (absolute homogeneity)

then $\|x\| \coloneqq d(x, 0)$ defines a norm on X and d is induced by this norm.

In the following proposition, we collect some elementary but fundamental facts baout normed linear spaces.

Exercise 1.2. $(X, \|\cdot\|)$ is a normed linear space over $\mathbb{F}$. Let $(x_n)_{n\geq 1}, (y_n)_{n\geq 1}$ be sequences in X and $(\lambda_n)_{n\geq 1}$ be a sequence in $\mathbb{F}$.

(a) Norm is a continuous funcrion, i.e., $x_n \to x$ (w.r.t. metric d) implies $\|x_n\| \to \|x\|$.

(b) Vector addition and scalar multiplicationis are continuous, i.e., $x_n \to x$ and $y_n \to y$ implies $x_n + y_n \to x + y$. is continuous, and $\lambda_n \to \lambda$ and $x_n \to x$ implies $\lambda_n x_n \to \lambda x$.

Equivalence between norms

Definition 1.2. If $\| \cdot \|_1$ and $\| \cdot \|_2$ are two norms on X, they are said to be **equivalent** if they define the same topology on X.

We will give a much direct condition to judge equivalence.

Theorem 1.1. *Let* $\|\cdot\|_1$ *and* $\|\cdot\|_2$ *be two norms on* X*, then they are equivalent iff there are positive constants* C_1 *and* C_2 *such that*

$$C_1\|x\|_1 \leqslant \|x\|_2 \leqslant C_2\|x\|_1 \text{ for all } x \in X\,.$$

Proof. Note that the two topology coincides iff, for any sequence (x_n) and point x in X,

$$\|x_n - x\|_1 \to 0 \Leftrightarrow \|x_n - x\|_2 \to 0\,.$$ □

Banach spaces

Definition 1.3. A normed linear space is complete with respect to the metric induced by the norm is called a **Banach space**.

There are two types of properties of a Banach space: those that are topological and those that are metric. The metric properties depend on the precise norm (such as uniformly convex space); the topological ones depend only on the equivalence class of norms.

Definition 1.4. Let $(x_n)_{n=1}^{\infty}$ be a sequence in $(X, \|\cdot\|)$. If the sequence $(s_n)_{n=1}^{\infty}$ of partial sums, where $s_n = \sum_{k=1}^{n} x_k$ for each $n \in \mathbb{N}$, converges to s, then we say the series $\sum_{k=1}^{\infty} x_k$ *converges* and its *sum* is s. In this case we write $\sum_{k=1}^{\infty} x_k = s$. The series $\sum_{k=1}^{\infty} x_k$ is said to be *absolutely convergent* if $\sum_{k=1}^{\infty} \|x_k\| < \infty$.

Theorem 1.2. $(X, \|\cdot\|)$ *is a Banach space if and only if every absolutely convergent series in* X *is convergent.*

Proof. Let X be a Banach space and suppose that $\sum_{j=1}^{\infty} \|x_j\| < \infty$. For any $\epsilon > 0$ and $n \in \mathbb{N}$, let $s_n = \sum_{j=1}^{n} x_j$. Let $K = K_\epsilon$ be a positive integer such that $\sum_{j=K+1}^{\infty} \|x_j\| < \epsilon$. Then, for all $m > n > K$, we have

$$\|s_m - s_n\| = \left\|\sum_{n+1}^{m} x_j\right\| \leq \sum_{n+1}^{\infty} \|x_j\| \leq \sum_{K+1}^{\infty} \|x_j\| < \epsilon\,.$$

Hence the sequence (s_n) of partial sums forms a Cauchy sequence in X. since X is complete, the sequence (s_n) converges to some element $s \in X$. That is, the series $\sum_{j=1}^{\infty} x_j$ converges.

Conversely, assume that $(X, \|\cdot\| \cdot \|$ is a normed linear space in which every absolutely convergent series converges. Let (x_n) be a Cauchy sequence in X. Then there is an sequence (x_{n_k}) such that for each $k \in \mathbb{N}$, $\left\|x_{n_{k+1}} - x_{n_k}\right\| < 2^{-k}$. Then $\sum_{k=1}^{\infty} \left\|x_{n_{k+1}} - x_{n_k}\right\| < \infty$. By our assumption, the series $\sum_{k=1}^{\infty}(x_{n_{k+1}} - x_{n_k})$ is convergent to some $s \in X$. It follows that as $j \to \infty$

$$s_j = \sum_{k=1}^{j} (x_{n_k+1} - x_{n_k}) = x_{n_j+1} - x_{n_1} \to s\,.$$

Thus, the subsequence (x_{n_k}) of (x_n) converges in X. But if a Cauchy sequence has a convergent subsequence, then then the sequence itself also converges to the same limit as the subsequence. Hence X is complete. $\square$

Examples

We give some examples of normed linear spaces to end the section. We denote by $\mathbb{N}$ all the positive integers, and by $\mathbb{N}_0$ all the non-negative integers.

Example 1.1. Let $n \in \mathbb{N}$, for each $x = (x_1, x_2, \cdots, x_n) \in \mathbb{F}^n$, define

$$\|x\|_p \coloneqq \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}, \text{ for } 1 \leqslant p < \infty$$
$$\|x\|_\infty \coloneqq \max_{1 \leqslant i \leqslant n} |x_i|$$

Then $(\mathbb{F}^n, \|\cdot\|_p)$ and $(\mathbb{F}^n, \|\cdot\|_\infty)$ are Banach spaces.

Example 1.2. Let $M_n(\mathbb{F})$ be the linear space of all $n \times n$ matrices over $\mathbb{F}$.

For each $A = (a_{ij}) \in M_n(\mathbb{F})$, define

$$\|A\|_{tr} \coloneqq \left(\sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}|^2 \right)^{\frac{1}{2}}$$

$$\|A\|_{\infty} \coloneqq \max_{1 \leqslant i,j \leqslant n} |a_{ij}|$$

Then $(M_n(\mathbb{F}), \|\cdot\|_{tr})$ and $(M_n(\mathbb{F}), \|\cdot\|_{\infty})$ are normed linear spaces.

Example 1.3. For each $x = (x_i)_{i \in \mathbb{N}} \in \ell^p$, where $1 \leqslant p < \infty$, define

$$\|x\|_p \coloneqq \left(\sum_{i \in \mathbb{N}} |x_i|^p \right)^{\frac{1}{p}}$$

Then $(\ell^p, \|\cdot\|_p)$ is a normed linear space.

Example 1.4. Let ℓ_0 be the set of all sequences $x = (x_i)_{i \in \mathbb{N}}$ of real or complex numbers such that $x_i = 0$ for all but finitely many incluses i. let c be the set of all convergent sequences $x = (x_i)_{i \in \mathbb{N}}$ of real or complex numbers. Let c_0 is the set of all convergent sequences $x = (x_i)_{i \in \mathbb{N}}$ of real or complex numbers which converge to 0. Suppose that $X = \ell_\infty, \ell_0, c$ or c_0. For each For each $x = (x_i)_{i \in \mathbb{N}} \in X$, define

$$\|x\|_\infty \coloneqq \max_{i \in \mathbb{N}} |x_i|$$

Then ℓ_∞, c, c_0 is Banach space, but ℓ_0 is an incomplete subspace of ℓ_∞.

Example 1.5. Let X be any Hausdorff space [1] and let $C_b(X)$ be all bounded continuous functions $f : X \to \mathbb{F}$. Define

$$\|f\|_\infty \coloneqq \sup_{x \in X} |f(x)| \, .$$

Then $C_b(X)$ is a Banach space. Particularly, when X is compact, let $C(X)$ be the set of all continuous $\mathbb{F}$-valued functions, then $C(X) = C_b(X)$.

[1] All topological spaces in this notes are assumed to be Hausdorff unless the contrazy is specified

Example 1.6. Let X is LCH (locally compact Hausdorff) space and $C_0(X)$ be all continuous functions $f : X \to \mathbb{F}$ such that that vanish at infinity: For any $\epsilon > 0, \{|f| \geqslant \epsilon\}$ is compact in X. Then for each f, Define

$$\|f\|_\infty \coloneqq \sup_{x\in X} |f(x)|\,.$$

then $C_0(X)$ is a closed subspace of $C_b(X)$ and hence is a Banach space. If X is compact, $C_0(X) = C_b(X) = C(X)$.

To prove this, observe that $C_0(X)$ is a linear subspace in $C_b(X)$. It will only be shown that $C_0(X)$ is closed in $C_b(X)$. Let $\{f_n\} \subset C_0(X)$ and suppose $f_n \to f$ in $C_b(X)$. Given $\epsilon > 0$, there is an integer n such that

$$\|f_n - f\| < \frac{\epsilon}{2}$$

Thus,

$$\{|f| \geqslant \epsilon\} \subset \left\{|f_n| \geqslant \frac{\epsilon}{2}\right\}$$

so that $f \in C_0(X)$.

Remark 1.2. Let X be LCH space and let $X_\infty = X \cup \{\infty\}$ be the one-point compactification of X. Then one can show that $\{f \in C\left(X_\infty\right) : f(\infty) = 0\}$, with the norm it inherits as a subspace of $C\left(X_\infty\right)$, is *isometrically isomorphic* to $C_0(X)$. Paricularly,

- $C_0(\mathbb{R})$ = all of the $\mathbb{F}$-valued continuous functions f such that $|f(x)| \to 0$ as $|x| \to \infty$.
- Let $\mathbb{N}$ equipped with discrete topology, then $C_0(\mathbb{N}) = c_0$.

Exercise 1.3. Let X be LCH and define $C_c(X)$ to be the continuous functions $f : X \to \mathbb{F}$ such that $\mathrm{supp}(f)$ (the closure of $\{f \neq 0\}$) is compact. Show that $C_c(X)$ is dense in $C_0(X)$.

The next example is usually proved in courses on integration and no proof is given here.

Example 1.7. For $1 \leqslant p \leqslant \infty$, $L^p(X, \mathcal{F}, \mu)$ is Banach space.

Example 1.8. Let I be a set and $1 \leqslant p < \infty$. Define $\ell^p(I)$ to be the set of all functions $f : I \to \mathbb{F}$ such that $\sum_{i \in I} |f(i)|^p < \infty$, and define

$$\|f\|_p = \left(\sum_{i \in I} |f(i)|^p \right)^{1/p}.$$

Then $\ell^p(I)$ is a Banach space. If $I = \mathbb{N}$, then $\ell^p(\mathbb{N})$ is often denoted as ℓ^p. In fact, let 2^I is all the subset of I and $\#$ be counting measure on $(I, 2^I)$. We have $\ell^p(I) = L^p(I, 2^I, \#)$.

Example 1.9. Let $n \geqslant 1$ and let $C^{(n)}[0,1]$ be the collection of functions $f : [0,1] \to \mathbb{F}$ such that f has n continuous derivatives. Define

$$\|f\| = \sup_{0 \leq k \leq n} \sup_{x \in [0,1]} |f^{(k)}(x)|$$

Then $C^{(n)}[0,1]$ is a Banach space.

Example 1.10 (Sobolev space)**.** Let $1 \leqslant p < \infty$ and $n \in \mathbb{N}$ and let $W_p^n[0,1]$ be all the functions $f : [0,1] \to \mathbb{F}$ such that f has $n-1$ continuous derivatives, $f^{(n-1)}$ is absolutely continuous, and $f^{(n)} \in L^p[0,1]$. For f in $W_p^n[0,1]$, define

$$\|f\| = \sum_{k=0}^{n} \left[\int_0^1 \left| f^{(k)}(x) \right|^p dx \right]^{1/p}$$

Then $W_p^n[0,1]$ is a Banach space.

1.2 Bounded linear operators

In this section, we always assume X and Y are two vector sapce over the same field $\mathbb{F}$.

Definition 1.5. $A \in \mathcal{L}(X, Y)$ is said to be **bounded** if there exists a constant $M > 0$ such that

$$\|Ax\| \leqslant M\|x\|,^{2} \text{ for all } x \in X.$$

We shall denote by $\mathcal{B}(X, Y)$ the set of all bounded linear operators from X into Y. We shall write $\mathcal{B}(X)$ for $\mathcal{B}(X, X)$.

Obviously, A is bounded if and only if A maps a bounded set into a bounded set. Moreover, boundedness is equivalent to continuity.

Exercise 1.4. Let $A \in \mathcal{L}(X, Y)$. Then the following statements are equivalent: (i) A is bounded. (ii) A is continuous on X. (iii) A is continuous at some point x_0 in X.

Definition 1.6. Let $A \in \mathcal{B}(X, Y)$. The **operator norm** (or simply norm) of A, denoted by $\|A\|$, is defined as

$$\|A\| \coloneqq \inf\{M : \|Tx\| \leqslant M\|x\|,\ \forall\, x \in X\}\,.$$

Remark 1.3. As we can see, the definition of $\|A\|$ can be change as the following equations:

$$\|A\| = \sup_{x \neq 0} \frac{\|Tx\|}{\|x\|} = \sup\{\|Tx\| : \|x\| = 1\} = \sup\{\|Tx\| : \|x\| \leqslant 1\}\,.$$

It's easy to check that the function $\|\cdot\|$ defined above is a norm on $\mathcal{B}(X, Y)$, so $\mathcal{B}(X, Y)$ become a normed linear space. So, it's natural to ask: When it becomce a Banach space ?

[2] It should be emphasised that the norm on the left side is in Y and that on the right side is in X .

Theorem 1.3. *$\mathcal{B}(X,Y)$ is a Banach space if and only if Y is a Banach space.*

Proof. Sufficiency. Let $\{T_n\}$ be a Cauchy sequence in $\mathcal{B}(X,Y)$. Then for any $x \in X$, $\{T_n x\}$ is Cauchy sequence in Y. So we can define $T : X \to Y$ by letting

$$Tx \coloneqq \lim_{n\to\infty} T_n x \quad \text{for } x \in X\,.$$

Evidently, $T \in \mathcal{B}(X,Y)$. To show that $T_n \to T$ in $\mathcal{B}(X,Y)$, note that for all $x \le 1$,

$$\|T_n x - Tx\| = \lim_{m\to\infty} \|T_n x - T_m x\| \le \liminf_{m\to\infty} \|T_n - T_m\|\,.$$

Since $\{T_n\}$ is a Cauchy sequence, we can see that $T_n \to T$.

Necessity. Let $\{y_n\}$ be a Cauchy sequence in Y. Let ℓ be a continuous non-zero linear functional on X. Define $T_n : X \to Y$ by letting

$$T_n x \coloneqq \ell(x) y_n \quad \text{for } x \in X\,.$$

Claerly $T_n \in \mathcal{B}(X,Y)$, and $\|T_n - T_m\| \le \|\ell\|\,\|y_n - y_m\|$, so $\{T_n\}$ is a Cauchy sequence, and there exists $T \in \mathcal{B}(X,Y)$ so that $T_n \to T$ in $\mathcal{B}(X,Y)$. Choose $x_0 \in X$ so that $\ell(x_0) = 1$, then

$$T_n x_0 = y_n \to T x_0\,.$$

So Y is complete. □

Remark 1.4. Let X be a normed linear space and Y be a Banach space, Denote $\tilde{X}$ be the completion of X. Define $\rho : \mathcal{B}(\tilde{X},Y) \to \mathcal{B}(X,Y)$ by $\rho(A) = A|_X$, then ρ is an isometric isomorphism.

Let $T : X \to Y$ and $S : Y \to Z$. We define the composition of T and S as the map $ST : X \to Z$ defined by

$$(ST)(x) = (S \circ T)(x) = S(T(x)) \text{ for all } x \in X.$$

Exercise 1.5. Let $T \in \mathcal{B}(X,Y), S \in \mathcal{B}(Y,Z)$. Then $ST \in \mathcal{B}(X,Z)$ and $\|ST\| \leqslant \|S\|\|T\|$.

Examples

We give some examples of normed linear spaces to end the section.

Example 1.11. Let $X = \mathbb{F}^n$ with the norm $\|\cdot\|_\infty$ and $A = (\alpha_{ij}) \in M_n(\mathbb{F})$. For $x = (x_1, x_2 \cdots, x_n)' \in \mathbb{F}^n$, define $T : X \to X$ by

$$Ax = A(x_1, x_2 \cdots, x_n)' = \left(\sum_{j=1}^{n} \alpha_{1j} x_j, \sum_{j=1}^{n} \alpha_{2j} x_j \cdots, \sum_{j=1}^{n} \alpha_{nj} x_j\right)'.$$

Then $A \in \mathcal{B}(X)$ and $\|A\| = \sup\limits_{1 \leqslant i \leqslant n} \sum\limits_{j=1}^{n} |\alpha_{ij}|$.

Example 1.12. Define an operator $L : \ell^2 \to \ell^2$ by

$$Lx = L(x_1, x_2, x_3, \cdots) = (x_2, x_3, \cdots)$$

Then $L \in \mathcal{B}(\ell^2)$ and $\|L\| = 1$. The operator L is called the *shift operator*.

Example 1.13. Let $X = \mathcal{P}[0, 1]$, the set of polynomials on the interval $[0, 1]$ with the uniform norm $\|\cdot\|_\infty$. For each $x \in X$, define $A : X \to X$ by

$$Tx = x'(t), \text{ for all } t \in [0, 1]$$

Then A is a linear operator but NOT bounded.

Example 1.14 (Multiplication operator)**.** Let $(X, \mathcal{F}, \mu)$ be a σ-finite measure space and $\phi \in L^\infty(\mu)$. Let $1 \leqslant p \leqslant \infty$. Define $M_\phi : L^p(\mu) \to L^p(\mu)$, by

$$M_\phi f = \phi f \text{ for all } f \in L^p(\mu)$$

Then $M_\phi \in \mathcal{B}(L^p(\mu))$ and $\|M_\phi\| = \|\phi\|_\infty$. The operator M_ϕ is called a *multiplication operator*. The function ϕ is it's *symbol*. To see this, it's easy to see that if $f \in L^p(\mu)$, then

$$\int |\phi f|^p d\mu \leqslant \|\phi\|_\infty^p \int |f|^p d\mu.$$

Thus, $M_\phi \in \mathcal{B}(L^p(\mu))$ and $\|M_\phi\| \leqslant \|\phi\|_\infty$. On the other hand (assume that $\|\phi\|_\infty > 0$), for any $\epsilon > 0$, the σ-finiteness of the measure space implies that there is a set B in $\mathcal{F}$, such that $0 < \mu(B) < \infty$ and

$$|\phi(x)| \geqslant \|\phi\|_\infty - \epsilon \quad \text{ for } x \in B\,.$$

Let $f = \mu(B)^{-1/p}\chi_B$, then $f \in L^p(\mu)$ and $\|f\|_p = 1$. So

$$\|M_\phi\|^p \geqslant \|\phi f\|_p^p = (\mu(B))^{-1}\int_B |\phi|^p d\mu \geqslant (\|\phi\|_\infty - \epsilon)^p\,.$$

Letting $\epsilon \to 0$, we get that $\|M_\phi\| \geqslant \|\phi\|_\infty$. Thus $\|M_\phi\| = \|\phi\|_\infty$.

Remark 1.5. We should note that if the measure space $(X, \mathcal{F}, \mu)$ is not σ-finite, then the conclusion is not necessarily valid. Indeed, let $\mathcal{F}$ be the Borel subsets of $[0, 1]$ and define μ on $\mathcal{F}$ by

$$\mu(A) = \begin{cases} \lambda(A), & 0 \notin A. \\ \infty, & 0 \in A. \end{cases}$$

where λ is the Lebesgue measure. This measure μ has an infinite *atom* at 0 and therefore, is not σ-finite. Let $\phi = \chi_{\{0\}}$. Then $\phi \in L^\infty(\mu)$ and $\|\phi\|_\infty = 1$. We claim that $M_\phi = 0$, so $\|M_\phi\| = 0 < 1 = \|\phi\|_\infty$.

To see this, note that for any $f \in L^p(\mu)$, $f(0) = 0$, then $M_\phi f = f(0)\chi_{\{0\}} = 0$.

Example 1.15 (Integral operator)**.** Let $(X, \mathcal{F}, \mu)$ be a measure space and suppose $k : X \times X \to \mathbb{F}$ is an $\mathcal{F} \times \mathcal{F}$-measurable function for which there are constants C_1 and C_2 such that

$$\int |k(x, y)|\mu(dy) \leqslant C_1\,, \ \mu\text{-a.e. } x\,.$$
$$\int |k(x, y)|\mu(dx) \leqslant C_2\,, \ \mu\text{-a.e. } y\,.$$

Define $K : L^p(\mu) \to L^p(\mu)$ $(1 < p < \infty)$ by

$$(Kf)(x) = \int f(y)k(x, y)\mu(dy)\,,$$

then K is a bounded linear operator and

$$\|K\| \leqslant C_1^{1/q} C_2^{1/p}, \text{ where } \frac{1}{p} + \frac{1}{q} = 1\,.$$

The operator described above is called an *integral operator* and the function k is called its *kernel*. (There exists other conditions on the kernel imply that K is bounded.) To see this, it must be shown that $Kf \in L^p(\mu)$, but it will follow from the argument that demonstrates the boundedness of K. For $f \in L^p(\mu)$,

$$\begin{aligned}
|Kf(x)| &\leqslant \int |k(x,y)||f(y)|\mu(dy) = \int |k(x,y)|^{1/q}|k(x,y)|^{1/p}|f(y)|\mu(dy) \\
&\leqslant \left[\int |k(x,y)|\mu(dy)\right]^{1/q} \left[\int |k(x,y)||f(y)|^p\mu(dy)\right]^{1/p} \\
&\leqslant C_1^{1/q} \left[\int |k(x,y)||f(y)|^p\mu(dy)\right]^{1/p}
\end{aligned}$$

Hence

$$\begin{aligned}
\int |Kf(x)|^p\mu(dx) &\leqslant C_1^{p/q} \iint |k(x,y)||f(y)|^p\mu(dy)\mu(dx) \\
&= C_1^{p/q} \int |f(y)|^p\mu(dy) \int |k(x,y)|\mu(dx) \\
&\leqslant C_1^{p/q} C_2 \|f\|^p
\end{aligned}$$

Thus $\|K\| \leqslant C_1^{1/q} C_2^{1/p}$.

A particular example of an integral operator is the Volterra operator defined below.

Example 1.16. Let $k : [0,1] \times [0,1] \to \mathbb{R}$ be the characteristic function of $\{(x,y) : y < x\}$. The corresponding operator $V : L^2(0,1) \to L^2(0,1)$ defined by $Vf(x) = \int_0^1 k(x,y)f(y)dy$ is called the *Volterna operator*. Note that

$$Vf(x) = \int_0^x f(y)dy$$

1.3 Finite dimensional space

Suppose that X is a finite-dimensional linear space over $\mathbb{F}$. Let $\{e_1, \cdots, e_n\}$ be a basis of X. Then we can define a "natural" norm on X. For $x = \sum_{j=1}^{n} \alpha_j e_j$, let

$$|x| \coloneqq \sum_{j=1}^{n} |\alpha_j| .$$

It's easy to check that $|\cdot|$ is a norm on X. A suprisingly result told us, this norm is the unique norm on X in the sense of equivalence.

Theorem 1.4. *All the norms on X are equivalent.*

Proof. Let $\|\cdot\|$ be a norm on X. We shall show that $|\cdot|$ is equivalent to $\|\cdot\|$. Clearly, $S = \left\{(\alpha_1, \cdots, \alpha_n) \in \mathbb{F}^n : \sum_{j=1}^{n} |\alpha_j| = 1\right\}$ is compact subset of $\mathbb{F}^n$. We define $f : S \to \mathbb{R}$ by

$$f(\alpha_1, \alpha_2, \cdots, \alpha_n) \coloneqq \left\| \sum_{j=1}^{n} \alpha_j e_j \right\| .$$

Clearly, f is continuous on S. Since S is compact, f attains its minimum and maximum on S, which is positive. Thus there exists a sonstant C_1, C_2 so that for all x

$$C_1|x| \leq \|x\| \leq C_2|x| .$$

Hence we complete the proof. □

By the theorem above, we can find that

Corollary 1.5. *Every finite-dimensional normed linear space $(X, \|\cdot\|)$ is Banach space.*

Corollary 1.6. *In a finite-dimensional normed linear space $(X, \|\cdot\|)$, a subset $K \subset X$ is compact if and only if it is closed and bounded.*

Corollary 1.7. *Let X and Y be normed linear spaces and* $\dim(X) < \infty$*. Then* $\mathcal{B}(X,Y) = \mathcal{L}(X,Y)$*.*

We give a topological characterization of the algebraic concept of finite dimensionality. The following lemma is needed, and we will need it when discussing compact operator.

Lemma 1.8 (Riesz's Lemma)**.** *Let M be a poper closed subspace of* $(X, \|\cdot\|)$*. Then for each $\epsilon \in (0,1)$, there is an element $y \in X$, depending on ϵ, such that*

$$\|y\| = 1 \quad \textit{and} \quad d(y, M) > 1 - \epsilon\,.$$

Proof. Choose $x \in M^c$ and denote $d = d(x, M)$. $d > 0$ since M is closed. Then for any given $\epsilon > 0$, there is a $m \in M$ such that

$$d \leq \|x - m\| < (1+\epsilon)d\,.$$

Then for any $u \in M$,

$$\left\|\frac{x-m}{\|x-m\|} - u\right\| \geqslant \frac{d}{\|x-m\|} > \frac{d}{(1+\epsilon)d} > 1 - \epsilon\,.$$

Let $y = \frac{x-m}{\|x-m\|}$, we have completed the proof. □

We now give a topological characterization of the algebraic concept of finite dimensionality.

Theorem 1.9. $(X, \|\cdot\|)$ *is finite-dimensional iff its closed unit ball* $B_X = \{x : \|x\| \leq 1\}$ *is compact.*

Proof. Assume that the closed unit ball $B_X = \{x \in X : \|x\| \leq 1\}$ is compact. Then B_X is totally bounded. Hence there is a finite $\frac{1}{2}$-net

$$\{x_1, x_2, \ldots, x_n\} \subset B_X\,.$$

Let $M = \operatorname{span}\{x_1, x_2, \ldots, x_n\}$. Then M is a finite-dimensional linear subspace of X and hence closed.

If M is a proper subspace of X, then, by Riesz's lemma, there is an element $x_0 \in B_X$ such that $d(x_0, M) > \frac{1}{2}$. In particular, $\|x_0 - x_k\| > \frac{1}{2}$ for all $k = 1, 2, \ldots, n$. However this contradicts the fact that $\{x_1, x_2, \ldots, x_n\}$ is a $\frac{1}{2}$-net in B_X. Hence $M = X$ and, consequently, X is finite-dimensional. $\square$

1.4 Quotient space

Assume that M is a subspace of X. We learned the quotinent space X/M in the course of linear algebra. We denoe by $\widetilde{x}$ the element in X/M, i.e., for $x \in X$,

$$\widetilde{x} := x + M = \{x + m : m \in M\}\,.$$

Next, we shall make X/M become a normed vector space.

Theorem 1.10. *Let M be a closed linear subspace of a normed linear space X over $\mathbb{F}$ .Then the quotient space X/M is a normed linear space with respect to the quotient norm defined by*

$$\|\widetilde{x}\| := \inf_{y \in \widetilde{x}} \|y\| = d(x, M)\ ,\ \ \textit{where } \widetilde{x} \in X/M. \tag{1.1}$$

Proof. Observe that

$$\|\widetilde{x}\| = \inf_{y \in \widetilde{x}} \|y\| = \inf_{m \in M} \|x + m\| = \inf_{m \in M} \|x - m\| = d(x, M)$$

Then it's easy to check the three conditions which makes $\|\cdot\|$ be a norm. In fact, since M is closed, we have

$$\|\widetilde{x}\| = 0 \Leftrightarrow d(x, M) = 0 \Leftrightarrow x \in M \Leftrightarrow \widetilde{x} = \widetilde{0}\,.$$

Since the metric d is absolutely homogeneous,

$$\|\lambda\widetilde{x}\| = d(\lambda x, M) = |\lambda| d(x, M) = |\lambda| \|\widetilde{x}\|\,.$$

By the trigonometric inequality and translation invaruance of d,

$$\|\widetilde{x} + \widetilde{y}\| = d(x + y, M) \le d(x, M) + d(y, M) = \|\widetilde{x}\| + \|\widetilde{y}\|.$$

So the desired result follows. $\square$

Remark 1.6. From the proof above, we should note that if M is a subspace but not closed, then (1.1) is only a seminorm.

Let M be a closed subspace of the normed linear space X. The mapping $Q = Q_M : X \to X/M$ defined by

$$Q_M(x) \coloneqq \widetilde{x} \quad \text{for all } x \in X\,,$$

is called the **quotient map** (or **natural map**) of X onto X/M.

Theorem 1.11. *M be a closed subspace of normed linear space X, Q is the quotinent map. Then*

(a) Q is continuous linear operator and $\|Q\| = 1$.

(b) Q is an open surjective map.

(c) The topology induced by quotient norm on X/M coinsides with the quotient topology on X/M. In words, a subset U of X/M is open if and only if $Q^{-1}(U)$ is open.

Proof. (a). Clearly Q is linear. By the definition of quotinent norm, $\|Q\| \le 1$. By Lemma 1.8, $\|Q\| = 1$.

(b). Clearly Q is surjective. Fix a open subset G, and for any $\widetilde{x} \in Q(G)$, without loss of generality, we can asuume $x \in G$. Then there exists $\delta > 0$ so that $B_X(x, \delta) \subset G$. We shall show that $B_{X/M}(\widetilde{x}, \delta) \subset Q(G)$, and thus $Q(G)$ is open. To this end, for any $\widetilde{y} \in B_{X/M}(\widetilde{x}, \delta)$, there exists $m \in M$ so that $\|y - m - x\| < \delta$. So $y - m \in B_X(x, \delta)$ and $\widetilde{y} = Q(y - m) \in Q(G)$.

(c) If U is open, $Q^{-1}(U)$ is open since Q is continuous. If $Q^{-1}(U)$ is open, then $U = Q(Q^{-1}(U))$ is open since Q is a open mapping. $\square$

Remark 1.7. In fact, for topology space X and Y, if $f : X \to Y$ is a continuous open mapping, then the topology on Y is the quotinent topology (w.r.t. f).

Corollary 1.12. *Let X be a normed space. Let M e a closed subspace of X. Let N be a finite dimensional subspace of X. Then $M + N$ is a closed subspace of X.*

Proof. Let Q_M be the quotient map. Observing that

$$M + N = Q_M^{-1}(Q_M(N))\,,$$

$Q_M(N)$ is a finite-dimensional subspace of X/M, and Q_M is continuous, the desired result follows. $\square$

Theorem 1.13. *Let X be a Banach space. Let M be a subspace. Then the quotient space X/M, equipped with the quotient norm, is a Banach space.*

Proof. Let $(\widetilde{x_n})_{n\geq 1}$ be a sequence in X/M such that $\sum_{j=1}^{\infty} \|\widetilde{x}_j\| < \infty$. For each $j \in \mathbb{N}$, choose an element $m_j \in M$ such that

$$\|x_j - m_j\| \leq \|\widetilde{x_j}\| + 1/2^j.$$

It now follows that $\sum_{j=1}^{\infty} \|x_j - m_j\| < \infty$. Since X is a Banach space, the series $\sum_{j=1}^{\infty} (x_j - m_j)$ converges to some element $z \in X$. Since the quotient mapping is contious, the series $\sum_{j=1}^{\infty} \widetilde{x_j}$ converges to $\widetilde{z}$. Hence, every absolutely convergent series in X/M is convergent, and so X/M is complete. $\square$

1.5 Other topics

Separable spaces, Schauder bases

A subset S of a normed linear space $(X, \|\cdot\|)$ is said to be *dense* in X if $\overline{S} = X$. If $(X, \|\cdot\|)$ contains a countable dense subset S, we say $(X, \|\cdot\|)$ is **separable**.

Lemma 1.14. *X is separable if and only if it contains a countable set S such that $\overline{\mathrm{span}}(S) = X$.*

Proof. It suffices to show that if there is a countable subset S of X so that $\overline{\mathrm{span}}(S) = X$, then X is seperable. To this end, let $\mathbb{K}$ be a countable dense subset of $\mathbb{F}$ and define

$$A = \left\{ \sum_{i=1}^{n} \lambda_i x_i : n \geq 1, \lambda_i \in \mathbb{K}, x_i \in S \right\}.$$

Clearly A is countable dense subset of $\mathrm{span}(S)$. Hence A is a countable dense subset of X. $\square$

Example 1.17. $\mathbb{R}, \mathbb{C}$ are separable, ℓ^p is separable, ℓ_∞ is not separable. To see the last one, consider

$$\{(a_n) : a_n \in \{0, 1\} \text{ for all } n \in \mathbb{N}\}.$$

Definition 1.7. A sequence $(e_n)_{n\in\mathbb{N}}$ in a separable Banach space is called a **Schauder basis** if, for any $x \in X$, there is a *unique* sequence $(\alpha_n(x))$ of scalars such that

$$x = \sum_{n=1}^{\infty} \alpha_n(x) e_n.$$

It is clear from this definition that if $(e_n)_{n\in\mathbb{N}}$ is a Schauder basis, then $\overline{\mathrm{span}}\{e_n : n \in \mathbb{N}\} = X$ Uniqueness of the expansion clearly implies that the set $\{e_n : n \in \mathbb{N}\}$ is linear independent.

Example 1.18.

- The sequence $(e_n)_{n\in\mathbb{N}}$ where $e_n = (\delta_{nm})_{m\in\mathbb{N}}$ is a Schauder basis for ℓ^p.
- The sequence $(e_n)_{n\in\mathbb{N}}$ where $e_n = (\delta_{nm})_{m\in\mathbb{N}}$ is a Schauder basis for c_0.
- The sequence $\{e\} \cup (e_n)$ where $e = (1, 1, 1, \cdots)$ is a Schauder basis for c.
- ℓ_∞ has no Schauder basis.

Remark 1.8. In 1937, Per Enflo constructed a separeble Banach space with no Schauder basis.

Direct sum

Now for the product or direct sum of normed spaces. Here there is a difficulty because, unlike Hilbert space, there is no canonical way to proceed. Suppose $\{X_i : i \in I\}$ is a collection of normed spaces. Then $\prod_{i\in I} X_i$ is a vector space if the linear operations are defined coordinatewise. The idea is to put a norm on a linear subspace of this product.

$$\oplus_p X_i \coloneqq \left\{ x \in \prod_i X_i : \|x\| \coloneqq \left[\sum_i \|x_i\|^p \right]^{1/p} < \infty \right\}.$$

and

$$\oplus_\infty X_i \coloneqq \left\{ x \in \prod_i X_i : \|x\| \coloneqq \sup_i \|x_i\| < \infty \right\}$$

Exercise 1.6. Let $\{X_i : i \in I\}$ be a collection of normed spaces and $p \in [0, \infty]$. Denote $X = \oplus_p X_i$, then

(a) X is a normed space and the projection $P_i : X \to X_i$ is a continuous linear map with $\|P_i x\| \leqslant \|x\|$ for each x in X.

(b) X is a Banach space if and only if each X_i is a Banach space.

(c) Each projection P_i is an open map of X onto X_i.

1.6 Hahn-Banach theorems

The Hahn-Banach Theorem is one of the most important results in mathematics. It is used so often it is rightly considered as a cornerstone of functional analysis. It is one of those theorems that when it or one of its immediate consequences is used, it is used without quotation or reference and we should to realize that it is being invoked.

Extension theorems

Let X be a vector space over $\mathbb{F}$. A *positive homogeneous, subadditive* functional is a function $p : X \to \mathbb{R}$ satisfying the following properties.

(a) Positive homogeneity: $p(\alpha x) = \alpha p(x)$ for x in X and $\alpha \geq 0$.

(b) Subadditivity: $p(x + y) \leqslant p(x) + p(y)$ for all x, y in X.

Trivally, every seminorm is a positive homogeneous, subadditive functional, but not conversely. It should be emphasized that a positive homogeneous, subadditive functional is allowed to assume negative values and that (b) in the definition only holds for $\alpha \geqslant 0$.

First of all, we deal with the real vector spaces.

Theorem 1.15. *Let X be a vector space over $\mathbb{R}$. Let p be a positive homogeneous, subadditive functional on X. Let Y be a subspace in X, and ℓ be a linear functional on Y. If ℓ satisfies*

$$\ell(y) \leqslant p(y), \textit{ for all } y \in Y,$$

then ℓ can be extended to X as a linear functional satisfying

$$\ell(x) \leqslant p(x), \textit{ for all } x \in X.$$

Remark 1.9. Note that the substance of the theorem isn't that the extension exists but that an extension can be found that remains *dominated* by p.

Just to find an extension, let $\{e_i\}$ be a Hamel basis for Y and let $\{\varepsilon_j\}$ be vectors in X such that $\{e_i\} \cup \{\varepsilon_j\}$ is a Hamel basis for X. Now define $L : X \to \mathbb{R}$ by

$$L\left(\sum_i \alpha_i e_i + \sum_j \beta_j \varepsilon_j\right) = \sum_i \alpha_i \ell(e_i) = \ell\left(\sum_i \alpha_i e_i\right).$$

This extends ℓ. If $\{\gamma_j\}$ is any collection of real numbers, then

$$L\left(\sum_i \alpha_i e_i + \sum_j \beta_j \varepsilon_j\right) = \ell\left(\sum_i \alpha_i e_i\right) + \sum_j \beta_j \gamma_j$$

is also an extension of ℓ. Moreover, any extension of ℓ has this form. The difficulty is that we must find one still dominated by p.

Proof. Without loss of generality, we suppose that Y is not all of X.

Step 1. There is some z in X that is not in Y. Denote by $Z = Y \oplus \mathrm{span}\{z\}$. Our aim is to extend ℓ as a linear functional L on Z such that L is domainted by p. Let's see what L must look like, if L exists. Put $\alpha_0 = L(z)$.

(a) Given $t > 0$ and $y \in Y$, we have $L(y + tz) = t\alpha_0 + \ell(y) \leqslant p(y + tz)$. Hence

$$\alpha_0 \leqslant \frac{p(y+tz)}{t} - \frac{\ell(y)}{t} = p\left(\frac{y}{t} + z\right) - \ell\left(\frac{y}{t}\right)$$

for every y in Y. Since $y/t \in Y$, this gives that

$$\alpha_0 \leqslant p(y+z) - \ell(y) \tag{1.2}$$

for all y in Y. On the other hand, if α_0 satisfies (1.2), then by reversing the preceding argument, it follows that $t\alpha_0 + \ell(y) \leqslant p(y+tz)$ whenever $t \geqslant 0$.

(b) Given $t > 0$ and $y' \in Y$, we have $L(y' - tz) = \ell(y') - t\alpha_0 \leqslant p(y' - tz)$. For the sme reason, this is equivalent to

$$\alpha_0 \geqslant \ell\left(y'\right) - p\left(y' - z\right). \tag{1.3}$$

Combining (1.2) and (1.3), we see that that α_0 can be chosen satisfying them both simultaneously. Such an α_0 exists iff for all pairs $y, y' \in Y$,

$$\ell\left(y'\right) - p\left(y' - z\right) \leqslant p(y+z) - \ell(y) \tag{1.4}$$

Using the linearity of ℓ and subadditivity of p we have

$$\ell\left(y + y'\right) \leqslant p(y + y') \leqslant p(y + z) + p\left(y' - z\right) \ .$$

So pick α_0 satisfying $\sup_{y' \in Y} \ell\left(y'\right) - p\left(-z + y'\right) \leqslant \alpha_0 \leqslant \inf_{y \in Y} p\left(z + y\right) - \ell\left(y\right)$, and define

$$L(y + tz) = \ell(y) + t\alpha_0\,, \text{ for } t \in \mathbb{R}\,.$$

we get a extension of ℓ on Z domainted by p.

Step 2. Consider all extensions of ℓ to linear spaces Z containing Y and domainted by p. We order these extensions by defining

$$(Z, \ell) \leqslant \left(Z', \ell'\right)$$

to mean that Z' contains Z, and that ℓ' agrees with ℓ on Z.

Let $\{Z_v, \ell_v\}$ be a totally ordered collection of extensions of ℓ. Then we can define $\widehat{\ell}$ on the union $\widehat{Z} = \cup_v Z_v$ as being ℓ_v on Z_v. Clearly, $\widehat{\ell}$ on $\widehat{Z}$ domainted by p, and $(Z_v, \ell_v) \leqslant (Z, \widehat{\ell})$ for all v. This shows that every totally ordered collection of extensions of ℓ has an upper bound. So the hypothesis of Zorn's lemma is satisfied, and we conclude that there exists a maximal extension. But according to the foregoing, a maximal extension must be to the whole space X. $\square$

To extend the result to complex vector spaces, we need the following lemma.

Lemma 1.16. *Let X be a vector space over $\mathbb{C}$.*

(a) *If $\ell : X \to \mathbb{C}$ is $\mathbb{C}$-linear, let $\ell_1 = \operatorname{Re}\ell$, then ℓ_1 is a $\mathbb{R}$-linear functional, and we have*

$$\ell(x) = \ell_1(x) - i\ell_1(ix)\,, \text{ for all } x \in X\,.$$

(b) *If $\ell_1 : X \to \mathbb{R}$ is an $\mathbb{R}$-linear functional, then $\ell(x) = \ell_1(x) - i\ell_1(ix)$ is a $\mathbb{C}$-linear functional. Moreover,*

$$\ell = \operatorname{Re}\tilde{\ell}\,.$$

(c) *Let p be a seminorm on X and ℓ and ℓ_1 are as in (a) or (b), then*

$$|\ell(x)| \leqslant p(x) \text{ for all } x \Leftrightarrow \ell_1(x) \leqslant p(x) \text{ for all } x\,.$$

Proof. The proofs of (a) and (b) are left as an exercise. To prove (c), suppose $|\ell(x)| \leqslant p(x)$. Then

$$\ell_1(x) = \operatorname{Re}\tilde{\ell}(x) \leqslant |\tilde{\ell}(x)| \leqslant p(x)\,.$$

Now assume that $\ell_1(x) \leqslant p(x)$. for all x. Let $\ell(x) = e^{i\theta}|\ell(x)|$. Hence

$$|\ell(x)| = \ell\left(e^{-i\theta}x\right) = \operatorname{Re}\ell\left(e^{-i\theta}x\right) = \ell_1\left(e^{-i\theta}x\right) \leqslant p\left(e^{-i\theta}x\right) = p(x)\,. \quad \square$$

Theorem 1.17 (Hahn-Banach Theorem I)**.** *Let X be a vector space over $\mathbb{F}$, let Y be a subspace, let p be a seminorm, and ℓ is a linear functional on Y. If ℓ is domainted by p,*

$$|\ell(y)| \leqslant p(y), \text{ for all } y \in Y\,,$$

then ℓ can be extended to X as a linear functional, still dominated by p:

$$|\ell(x)| \leqslant p(x), \text{ for all } x \in X\,.$$

Proof. There are two case for $\mathbb{F}$. When $\mathbb{F}$ is $\mathbb{R}$, then $\ell(y) \leqslant |\ell(y)| \leqslant p(y)$ for y in Y. By Theorem 1.15 , ℓ can be extended to X such that $\ell(x) \leqslant p(x)$ for all $x \in X$. Hence $-\ell(x) = \ell(-x) \leqslant p(-x) = p(x)$, That is, $|\ell(x)| \leqslant p(x)$.

When $\mathbb{F} = \mathbb{C}$. Let $\ell_1 = \operatorname{Re} \ell$. By Lemma 1.16, $\ell_1 \leqslant p$. From the proof above, ℓ_1 can be extended on X as $\mathbb{R}$-linear functional such that $|\ell_1| \leqslant p$. Let

$$\ell(x) = \ell_1(x) - i\ell_1(ix)$$

for all x in X, then ℓ is a extension. By Lemma 1.16, $|\ell| \leqslant p$. $\square$

Geometric Hahn-Banach theorems

In spite (or perhaps because) of its nonconstructive proof, the Hahn-Banach theorem has plenty of very concrete applications. One of the most important is to separation theorems concerning convex sets; these are sometimes called *geometric Hahn-Banach theorems*.

Minkowski functionals

Suppose X is a vector space over $\mathbb{F}$. Let K be a subset of X.

(a) K is called *convex* if, for any $t \in [0, 1]$, $tK + (1-t)K \subset K$.

(b) K is called *absorbing at* $x \in K$, if for any $y \in X$, there exists an $\epsilon > 0$, depending on y, such that

$$x + ty \in K \qquad \text{for all real } t, |t| < \epsilon\,.$$

If K contains 0 and absorbs at 0, we say K is *absorbing* for short.

(c) K is called *balanced*, if for any $\lambda \in \mathbb{F}$ and $|\lambda| \leqslant 1$, we have $\lambda K \subset K$.

Example 1.19. Let p be real-valued functional on X.

(a) If p is a positive homogeneous, subaddititive functional on X, then $\{x : p(x) < 1\}$ is a absorbing convex subset of X.

(b) If p is a seminorm on X, then $\{x : p(x) < 1\}$ is a balanced absorbing convex subset of X.

The following observation is quite useful. If p is a seminorm, letting $K = \{x : p(x) < 1\}$, there holds

$$p(x) = \inf \left\{ t > 0 : \frac{x}{t} \in K \right\} .$$

If we let K be any balanced absorbing convex set, then we can see that the avobe p satisfies that $p(x) < \infty$ for all x. Hence, we introduce the definition:

Definition 1.8. Let K be a balanced absorbing convex set, the *Minkowski functional of* K is defined by

$$p_K(x) = \inf \left\{ t > 0 : \frac{x}{t} \in K \right\} , \text{ for } x \in X . \tag{1.5}$$

Lemma 1.18. *Let K be balanced absorbing convex, then p_K is a seminorm.*

Proof. We only show that for all x, y in X,

$$p_K(x + y) \leqslant p_K(x) + p_K(y) .$$

For any $a, b > 0$, such that $\frac{x}{a}, \frac{y}{b} \in K$, note that

$$\frac{x+y}{a+b} = \frac{a}{a+b} \frac{x}{a} + \frac{b}{a+b} \frac{y}{b} \in K .$$

So $p_K(x + y) \leqslant a + b$. Letting $a \downarrow p_K(x)$ and $b \downarrow p_K(y)$, we get the desired result. $\square$

Lemma 1.19. *For any balanced absorbing convex set K,*

(a) $\{x : p_K(x) < 1\} \subset K \subset \{x : p_K(x) \leq 1\}$.

(b) $p_K(x) < 1$ *if and only if K is absorbing at x.*

Proof. We only show (b). When K is absorbing at x, clearly $p_K(x) < 1$. Suppose now that $p_K(x) < 1$. Given $y \in X$, note that

$$p_K(x + ty) \leq p_K(x) + |t| p_K(y) .$$

So there exists some $\epsilon > 0$, depending on y, so that when $|t| < \epsilon$,

$$p_K(x+ty) \leqslant p_K(x) + |t| p_K(y) \leqslant p_K(x) + \epsilon\, p_K(y) < 1\,.$$

By (a), $x + ty \in K$. $\square$

Note that in (1.5), if we let K be any absorbing convex set, give up the condition that K is balanced, $p_K(x) < \infty$ for all $x \in X$ still holds. Thuswe extends the Minkovski functional to all the absorbing convex set: Let K be absorbing and convex. The *Minkowski functional of* K is defined by

$$p_K(x) = \inf\left\{ t > 0 : \frac{x}{t} \in K \right\}\,.$$

As the preceeding theorems, we have

Lemma 1.20. *K is an absorbing convex set, p_K is the Minkowski functional of K. Then*

(a) *p_K is a positive homogeneous, subadditive functional on X.*

(b) $\{x : p_K(x) < 1\} \subset K \subset \{x : p_K(x) \leq 1\}$.

(c) *$p_K(x) < 1$ if and only if K is absorbing at x.*

Hyperplane and linear functionals

We turn now to the notion of a hyperplane. Let X be a vector space over $\mathbb{F}$, a subspace M is called a *hyperplane* in X if it has codimension 1, in words,

$$\dim(X/M) = 1\,.$$

An *affine hyperplane* in X is a hyperplane shifted from the origin by a vector, i.e., $x_0 + M$ is the affine hyperplane for some $x_0 \in X$.

It's easy to find that is a closed connection between hyperplanes and linear functionals.

- Let ℓ be a *nonzero* linear functional, then $\ker\ell$ is a hyperplane. In fact, there is an isomorhism between $X/\ker\ell$ and $\mathbb{F}$ naturally induced by ℓ.

- Let M be a hyperplane, and $Q : X \to X/M$ be the quotinent map and let $T : X/M \to \mathbb{F}$ be an isomorphism. Then $\ell \coloneqq T \circ Q$ is a linear functional with kernel $\ker\ell = M$.

This is summarized as follows :

Proposition 1.21. *let X be a vector space, $M \subset X$, then M is a hyperplane if and only if there is a nonzero linear functional ℓ such that $M = \ker\ell$. Consequently, M is an affine hyperplane if and only if there is non-zero linear functional ℓ and some scalar c such that $M = \{\ell = c\}$.*

Hyperplane seperation theorem

There is a great advantage inherent in a *geometric discussion* of *real* vector space, X. Since, if ℓ is a nonzero linear functional, then any affine hyperplane $\{\ell = c\}$ "disconnects" the space: all points of X belong to one, and only one, of the following three sets:

$$\{x : \ell(x) < c\}, \quad \{x : \ell(x) = c\}, \quad \{x : \ell(x) > c\}\,.$$

The sets where $\{\ell < c\}$, or$\{\ell > c\}$ are called *open halfspaces*. The sets where $\{\ell \geq c\}$, or $\{\ell \leq c\}$ are called *closed halfspaces*.

However, When X is a complex vector space, and ℓ is nonzero linear functional, then $X\backslash\{l = c\}$ is "connected". However, we can regard any complex vector space as a real vector space, in this case, we say ℓ is "linear" means ℓ is $\mathbb{R}$ -linear, not $\mathbb{C}$-linear. Then the resluts in real vector spaces can be applied in vector spaces over both $\mathbb{R}$ and $\mathbb{C}$ after the slight modification.

We say two subsets A and B of *real* vectors space X are said to be *strictly separated* if they are contained in disjoint open half-spaces; they are *separated* if they are contained in two closed half-spaces whose intersection is a affine hyperplane.

Theorem 1.22 (Hyperplane Separation Theorem)**.** *Let X be a real vector space. Let $K \subset X$ be convex and absorbing at each of it's points. Then, any point $y \notin K$ can be separated from K by a hyperplane $\{x : \ell(x) = c\}$. In words, there is a linear functional ℓ, depending on y, such that*

$$\ell(x) < c = \ell(y) \quad \textit{for all } x \textit{ in } K\,.$$

Proof. Without loss of generality, we assume that $0 \in K$. So K is absorbing. Denote by p_K the Minkovski functional of K. We define ℓ on $\mathrm{span}\{y\}$ by

$$\ell(ay) = a\,, \text{ for all } a \in \mathbb{R}\,.$$

We claim that for all such ay,

$$\ell(ay) \le p_K(ay)\,.$$

This is obvious for $a \le 0$, for then $\ell(ay) = a \le 0$ while $p_K \ge 0$. If $a > 0$, since $y \notin K$, $p_K(y) \ge 1$. So, $p_K(ay) \ge a = \ell(ay)$ for $a > 0$.

Having shown that ℓ, as defined on the above one-dimensional subspace, is dominated by p_K, we conclude from the Hahn-Banach theorem that ℓ can be so extended to all of X. We deduce from this and Lemma 1.20, since K is absorbing at each of it's points, it follows from that $p_K(x) < 1$ for every x in K, thus

$$\ell(x) \le p_K(x) < 1 = l(y)\,. \qquad \square$$

Corollary 1.23. *Let K denote a absorbing convex set. For any y not in K there is a nonzero linear functional ℓ that satisfies*

$$\ell(x) \le \ell(y) \textit{ for all } x \textit{ in } K.$$

Theorem 1.24 (Extended Hyperplane Separation)**.** *X is real vector space. A and B disjoint convex subsets of X. A is absorbing at some $a_0 \in A$. Then A and B can be separated by a hyperplane $\{x : \ell(x) = c\}$. That is, there is a nonzero linear functional ℓ, and a number c, such that*

$$\ell(a) \leqslant c \leqslant \ell(b) \quad \textit{for all}\, a \in A,\, b \in B\,.$$

Proof. Let $G = A - B = \{a - b : a \in A, b \in B\}$; it is easy to verify that G is convex and absorbsing at each of it's points (do it!). Moreover, $0 \notin G$, because $A \cap B = \emptyset$. By hyperplane separation theorem, there is a linear functional ℓ on X such that

$$\ell(a - b) < 0 = \ell(0)\,, \text{for any}\, a \in A, b \in B$$

Thus

$$\sup\{\ell(a) : a \in A\} \leqslant c \leqslant \inf\{\ell(b) : b \in B\}\,. \qquad \square$$

Chapter 2

Hilbert Space

A Hilbert space is the abstraction of the finite-dimensional Euclidean spaces of geometry. Its properties are very regular and contain few surprises, though the presence of an infinity of dimensions guarantees a certain amount of surprise. Historically, it was the properties of Hilbert spaces that guided mathematicians when they began to generalize.

2.1 Fundamentals

Definition 2.1. Let X be a vector space over $\mathbb{F}$. An **inner product** (or **scalar product**) on X is a scalar-valued function $\langle \cdot\,, \cdot \rangle : X \times X \to \mathbb{F}$ such that for all $x, y, z \in X$ and forall $\alpha, \beta \in \mathbb{F}$, we have

(a) $\langle x, x\rangle \geq 0$, and $\langle x, x\rangle = 0$ iff $x = 0$. (Positive definiteness)

(b) $\langle x, y\rangle = \overline{\langle y, x\rangle}$. (Hermitian property)

(c) $\langle \alpha x + \beta y, z\rangle = \alpha\langle x, z\rangle + \beta\langle y, z\rangle$. (Linearity)

If we change (a) to (a′): $\langle x, x\rangle \geq 0$, $\langle \cdot\,, \cdot\rangle$ is called a **semi-inner production**. X is called *inner production space*, or *semi-inner production space*, respectively.

Theorem 2.1 (Cauchy-Bunyakowsky-Schwarz inequality)**.** *Let* $(X, \langle \cdot, \cdot \rangle)$ *be an semi-inner product space over* $\mathbb{F}$*. Then, for all* $x, y \in X$*.*

$$|\langle x, y \rangle|^2 \leq \langle x, x \rangle \langle y, y \rangle \,,$$

Moreover, equality occurs if and only if there are scalars α, β*, both not* 0*, such that* $\langle \alpha x + \beta y, \alpha x + \beta y \rangle = 0$ *.*

Proof. We only show this inequality in the case of $\mathbb{F} = \mathbb{C}$. Note that, for any $\lambda \in \mathbb{C}$ and $x, y \neq 0$, we have

$$\langle x + \lambda y, x + \lambda y \rangle = \langle y, y \rangle |\lambda|^2 + 2Re\{\langle x, y \rangle \overline{\lambda}\} + \langle x, x \rangle \geq 0\,.$$

One can show that $f(z) = |\alpha z|^2 - 2Re\{\beta z\}$, where $\alpha, \beta \in \mathbb{C}, \alpha \neq 0$, achieve its minimum $-\frac{|\beta|^2}{|\alpha|^2}$ if and only if $z = \frac{\overline{\beta}}{|\alpha|^2}$. Then let $\lambda = -\frac{\langle x,y \rangle}{\langle y,y \rangle}$, we get

$$-\frac{|\langle x, y \rangle|^2}{\langle y, y \rangle} + \langle x, x \rangle \geq 0\,.$$

Moreover, the equality orrurs if and only if α, β, both not 0, such that $\langle \alpha x + \beta y, \alpha x + \beta y \rangle = 0$. □

Let $(X, \langle \cdot\,, \cdot \rangle)$ be an inner product (semi-inner product) space over $\mathbb{F}$. For each $x \in X$, define

$$\|x\| := \langle x, x \rangle^{\frac{1}{2}} \tag{2.1}$$

Then $\| \cdot \|$ is a norm, called the norm induced by the inner product. Using this notation, the CBS Inequality now becomes

$$|\langle x, y \rangle| \leq \|x\| \|y\|\,. \tag{2.2}$$

From this, we can see that the inner product is continuous: If $x_n \to x$, $y_n \to y$ w.r.t. the norm, then $\langle x_n, y_n \rangle$ tends to $\langle x, y \rangle$.

A natural question arises: Is every normed linear space an inner product space? If the answer is NO, how then does one recognise among all normed linear spaces those that are inner product spaces in disguise, i.e., those whose norms are induced by an inner product?

Proposition 2.2 (Polarization identity)**.** *Let* $(X, \langle \cdot, \cdot \rangle)$ *be an inner product space over* $\mathbb{F}$*. Given any* $x, y \in X$*,*

- *if* $\mathbb{F} = \mathbb{R}$*,* $\langle x, y \rangle = \dfrac{\|x+y\|^2 - \|x-y\|^2}{4}$.
- *if* $\mathbb{F} = \mathbb{C}$*,* $\langle x, y \rangle = \dfrac{\|x+y\|^2 - \|x-y\|^2}{4} + i\left(\dfrac{\|x+iy\|^2 - \|x-iy\|^2}{4}\right)$.

Proof. We only show the polarization identity in the case of $\mathbb{F} = \mathbb{C}$. Since

$$\|x+y\|^2 = \|x\|^2 + 2\mathrm{Re}\langle x, y \rangle + \|y\|^2 \,,$$
$$\|x-y\|^2 = \|x\|^2 - 2\mathrm{Re}\langle x, y \rangle + \|y\|^2 \,,$$

We get

$$\mathrm{Re}\langle x, y \rangle = \frac{\|x+y\|^2 - \|x-y\|^2}{4} \,.$$

Note that

$$\mathrm{Im}\langle x, y \rangle = \mathrm{Re}\langle x, iy \rangle \,,$$

the desired result follows. $\square$

Proposition 2.3 (Parallelogram identity)**.** *Let* $(X, \langle \cdot, \cdot \rangle)$ *be an inner product space over a field* $\mathbb{F}$*. then for all* $x, y \in X$*,*

$$\|x-y\|^2 + \|x+y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right) \,. \tag{2.3}$$

Moreover, a normed linear space X *over* $\mathbb{F}$ *is an inner product space if and only if the parallelogram identity holds for all* $x, y \in X$*.*

Proof. If $(X, \langle \cdot, \cdot \rangle)$ is an inner product, it's trival that the parallelogram identity holds.

Step 1. If X is normed space over $\mathbb{R}$ and the parallelogram identity holds, define

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

Then we can check $\langle \cdot, \cdot \rangle$ is an inner product. Clearly $\langle \cdot, \cdot \rangle$ is positive definite and symmetric. We firstly show that

$$\langle x+z, y \rangle = \langle x, y \rangle + \langle z, y \rangle \,.$$

This follows from parallelogram identity. By induction, for all $n \in \mathbb{N}$, we have $\langle nx, y\rangle = n\langle x, y\rangle$. Then it follows that for $r \in \mathbb{Q}$, $\langle rx, y\rangle = r\langle x, y\rangle$. Since $\langle\cdot,\cdot\rangle$ is continuous, for any $\lambda \in \mathbb{R}$, we have

$$\langle \lambda x, y\rangle = \lambda\langle x, y\rangle \,.$$

Step 2. If X is normed space over $\mathbb{C}$ and the parallelogram identity holds, define

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

Clearly $\langle\cdot,\cdot\rangle$ is positive definite and Hermitian. We have showed that for given y, $\langle\cdot, y\rangle$ is $\mathbb{R}$-linear in step 1. Observe that

$$\langle ix, y\rangle = i\langle x, y\rangle \,,$$

we have that $\mathrm{Re}\langle\cdot, y\rangle$ is $\mathbb{C}$-linear for given y. So the desired result follows. $\square$

Remark 2.1. From Proposition 2.3 , we know if every two-dimensional linear subspace of normed linear space X is an inner product space, then X is an inner product space.

The mathematical concept of a Hilbert space, named after David Hilbert, generalizes the notion of Euclidean space. Hilbert spaces, as the following definition states, are inner product spaces which in addition are required to be complete,a property that stipulates the existence of enough limits in the space to allow the techniques of calculus to be used.

Definition 2.2. $(H, \langle\cdot,\cdot\rangle)$ is inner product space. If H is complete, with respect to the norm induced by $\langle\cdot,\cdot\rangle$, then we say that H is a **Hilbert space**.

Remark 2.2. Given a linear space with a inner product, it can be completed with respect to the norm derived from the inner product. It follows from the C-B-S inequality that the inner product is a continuous function of its factors; therefore it can be extended to the completed space. Thus the completion is a Hilbert space.

Given a linear space with a inner product, it can be completed with respect to the norm derived from the scalar product. It follows from the Schwarz inequality that the

Example 2.1. Fix a positive integer n. Let $X = \mathbb{F}^n$. For $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$ in X, define

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i \overline{y_i} \, .$$

The space $\mathbb{R}^n$ (resp. $\mathbb{C}^n$) with this inner product is called the Euclidean space (resp. unitary space) and is a (trivial) Hilbert sapce.

Example 2.2. Let $M_n(\mathbb{C})$, the linear space of all $n \times n$ complex matrices. For any $A \in M_n(\mathbb{C})$ let $\operatorname{tr}(A) = \sum_{i=1}^{n} (A)_{ii}$ be the trace of A. For $A, B \in M_n(\mathbb{C})$, define

$$\langle A, B \rangle = \operatorname{tr}(B^* A)$$

where B^* denotes conjugate transpose of matrix B. Then $(M_n(\mathbb{C}), \langle \cdot, \cdot \rangle)$ is a Hilbert space.

Example 2.3. Let $X = \ell_0$, the linear space of finitely non-zero sequences of real or complex numbers. For $x = (x_1, x_2, \ldots)$ and $y = (y_1, y_2, \ldots)$ in X, define

$$\langle x, y \rangle = \sum_{i=1}^{\infty} x_i \overline{y_i} \, .$$

since this is essentially a finite sum, $\langle \cdot \, , \cdot \rangle$ is well-defined. ℓ^0 is an incomplete inner product space.

Example 2.4. Let $X = \ell^2$, the space of all sequences $x = (x_1, x_2, \ldots)$ of real or complex numbers with $\sum_{i=1}^{\infty} |x_i|^2 < \infty$. For $x = (x_1, x_2, \ldots)$ and $y = (y_1, y_2, \ldots)$ in X, define

$$\langle x, y \rangle = \sum_{i=1}^{\infty} x_i \overline{y_i} \, .$$

ℓ^2 is a Hilbert space. Moreover, ℓ^2 is the completion of ℓ_0 in previous example.

Example 2.5. Let I be any noempty set and let $\ell^2(I)$ denote the set of all functions $x : I \to \mathbb{F}$ such that $\sum_{i \in I} |x(i)|^2 < \infty$. For x and y in $\ell^2(I)$ define

$$\langle x, y \rangle = \sum_{i \in I} x(i)\overline{y(i)}.$$

For any noempty set I, $\ell^2(I)$ is a Hilbert space. One can find that $\ell^2(\mathbb{N})$ is exactly ℓ^2.

Example 2.6. Let $(\Omega, \mathcal{F}, \mu)$ be a measure space. Denote by $L^2(\Omega, \mathcal{F}, \mu)$ all $\mathbb{F}$-valued square integrable functions If f and $g \in L^2$, then Hölder's inequality implies $f\overline{g} \in L^1$. Definite

$$\langle f, g \rangle = \int f\overline{g}\, \mathrm{d}\mu\,,$$

then this defines an inner product on L^2. Then L^2 becomes a Hilbert space.

Example 2.7. Let $X = C[a, b]$, the space of all continuous $\mathbb{F}$-valued functions on $[a, b]$. For $x, y \in X$, define

$$\langle x, y \rangle = \int_a^b x(t)\overline{y(t)}dt$$

$C[a, b]$ is an incomplete inner product space. Evidently, $L^2[a, b]$ is the Completion of the $C[a, b]$.

2.2 Orthogonality

Definition 2.3. Two elements x and y in an inner product space $(X, \langle\cdot,\cdot\rangle)$ are said to be **orthogonal**, denoted by $x \perp y$, if

$$\langle x, y\rangle = 0 \; .$$

A subset $\Omega \subset X$ is called *orthogonal* if it consists of non-zero pairwise orthogonal elements.

Pythagorean theorem still holds in this case. In other words, if

$$\{x_1, x_2, \ldots, x_n\}$$

is an orthogonal set , then

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

Definition 2.4. Let M be a subset of X. If $\langle x, m\rangle = 0$ for all $m \in M$, then we say x is *orthogonal to* M and write $x \perp M$. We shall denote by

$$M^\perp = \{x \in X : \langle x, m\rangle = 0 \, , \forall\, m \in M\}$$

the set of all elements orthogonal to M. The set $M^\perp$ is called the **orthogonal complement** of M.

It's easy to find out that for any subset M of X, $M^\perp$ is a closed linear subspace of X, $M \subset M^{\perp\perp} := \left(M^\perp\right)^\perp$, and $M^\perp = (\operatorname{span} M)^\perp = (\overline{\operatorname{span}} M)^\perp$.

Best approximation

Theorem 2.4 (Existence of the unique best approximation)**.** *Let* $(X, \langle\cdot,\cdot\rangle)$ *be a inner product space. Let* K *be a nonempty complete convex subset of* X*. Then for each* $x \in X$ *has a unique best approximation in* K*, i.e. there is a unique point* $k_0 \in K$ *satisfying*

$$\|x - k_0\| = d(x, K) := \inf_{k \in K} \|x - k\| \, .$$

Proof. Without loss of generality, assume $x = 0 \notin K$. Then there exists a sequence $\{k_n\}_{n\geq 1}$ in K such that

$$\|k_n\| \downarrow d(0, K)\,.$$

By parallelogram identity

$$\left\|\frac{k_n - k_m}{2}\right\|^2 + \left\|\frac{k_n + k_m}{2}\right\|^2 = \frac{\|k_n\|^2 + \|k_m\|^2}{2}\,.$$

Thus $\{k_n\}$ is a Cauchy sequence. Since K is complete, there exists $k_0 \in K$ such taht $k_n \to k_0$, then $\|k_0\| = d(0, K)$.

Using parallelogram identity again, we will get the uniqueness of k_0. $\square$

Remark 2.3. The proof above is using the *uniformly convexness* of the norm induced by the inner product. In fact, in any uniformly convex normed linear space, Theorem 2.4 holds.

Theorem 2.5 (Characterization of the unique best approximation)**.** *Let K be a nonempty complete convex subset of a inner product space $(X, \langle\cdot\,,\cdot\rangle)$. Assume $x \in X\backslash K$ and $k_0 \in K$, then k_0 is the best approximation to x from K if and only if*

$$\operatorname{Re}\langle x - k_0, k - k_0\rangle \leq 0,\ \textit{for all } k \in K\,. \tag{2.4}$$

Proof. Take any $k \in K$ and fix it.

$$\begin{aligned}\|x - k\|^2 &= \|x - k_0 - (k - k_0)\|^2\\ &= \|x - k_0\|^2 + \|k - k_0\|^2 - 2\operatorname{Re}\langle x - k_0, k - k_0\rangle\,.\end{aligned}$$

So if (2.4) holds , we get $\|x-k\|^2 \geqslant \|x-k_0\|^2$. so k_0 is a best approximation. The uniqueness is guaranteed by Theorem 2.4.

If $\|x - k\|^2 \geqslant \|x - k_0\|^2$ for all $k \in K$, we have

$$2\operatorname{Re}\langle x - k_0, k - k_0\rangle \leqslant \|k - k_0\|^2\,.$$

For any $\lambda \in (0, 1)$ and $k' \in K$, let $k = \lambda k' + (1 - \lambda)k_0$, then we have

$$2\operatorname{Re}\left\langle x - k_0, k' - k_0\right\rangle \leqslant \lambda\|k' - k_0\|^2\,.$$

So we let λ tends to zero , we get (2.4). $\square$

Orthogonal decomposition

Theorem 2.6. *H is Hilbert space, M is a closed subspace. Given $x \in H$, then $m \in M$ is the unique best approximation to x from M if and only if*

$$x - m \perp M.$$

Proof. By Theorem 2.5, $m \in M$ is the unique best approximation to x from M if and only if

$$\operatorname{Re}\left\langle x - m, m' \right\rangle \leq 0, \text{ for all } m' \in M\,.$$

Let $m'' = -m'$, we can see that

$$\operatorname{Re}\left\langle x - m, m' \right\rangle = 0, \text{ for all } m' \in M\,.$$

If H is a real Hilbert space, then we get $x - m \perp M$ directly. If If H is a complex Hilbert space, note that

$$\operatorname{Im}\left\langle x - m, m' \right\rangle = \operatorname{Re}\left\langle x - m, im' \right\rangle = 0, \text{ for all } m' \in M\,,$$

the desired result follows. □

Theorem 2.6 says that if M is a closed linear subspace of a Hilbert space H, then $P_M(x)$ is the best approximation to x from M if and only if $x - P_M(x) \perp M$. That is, the unique best approximation is obtained by "dropping the perpendicular from x onto M". Therefore, the map

$$P_M : H \to M; x \mapsto P_M(x) \tag{2.5}$$

is also called the **projection** of H onto M. And we get the following important theorem :

Theorem 2.7 (Orthogonal decomposition)**.** *H is Hilbert space, M is a closed subspace. Then*

$$H = M \oplus M^{\perp}\,. \tag{2.6}$$

That is, each $x \in H$ can be uniquely decomposed in the form $x = y + z$ with $y \in M$ and $z \in M^{\perp}$.

Corollary 2.8. *Let S be a nonempty subset of a Hilbert space H. Then*

(a) $(S^{\perp})^{\perp} = \mathrm{cspan}(S)$

(b) $S^{\perp} = \{0\}$ *if and only if* $\mathrm{cspan}(S) = H$.

2.3 Orthonormal Bases

In this section, we always assume that H is an inner product space over the feild $\mathbb{F}$.

Definition 2.5. Let $\mathcal{E} \subset H$. We say that $\mathcal{E} = \{e_i\}_{i\in I}$ is **orthonormal** if $\mathcal{E}$ is orthogonal and $\|e_i\| = 1$ for all $i \in I$.

For any $x \in H$, the numbers $\langle x, e_i\rangle$ are called the *Fourier coefficients.* and the formal series $\sum_{i\in I}\langle x, e_i\rangle e_i$ are called the *Fourier series of x with respect to $\mathcal{E}$.*

Lemma 2.9 (Gram-Schmidt Orthonormalisation Procedure)**.** *Let $\{x_k\}_{k\geq 1}$ be a linearly independent set in H. There exists an orthonormal set $\{e_k\}_{k\geq 1}$ in H such that for all $n \in \mathbb{N}$,*

$$\operatorname{span}\{x_1, x_2, \ldots, x_n\} = \operatorname{span}\{e_1, e_2, \ldots, e_n\} .$$

Proof. Set $e_1 = \frac{x_1}{\|x_1\|}$. Then $\operatorname{span}\{x_1\} = \operatorname{span}\{e_1\}$. Then we define e_k by induction. Now assume $\{e_1, \cdots, e_k\}$ is orthonomal, and

$$\operatorname{span}\{e_1, e_2, \ldots, e_k\} = \operatorname{span}\{x_1, x_2, \ldots, x_k\} .$$

Let $\widehat{x}_{k+1}$ be the projection of x_{k+1} on $\operatorname{span}\{e_1, e_2, \ldots, e_k\}$, and

$$y_{k+1} = x_{k+1} - \widehat{x}_{k+1}\ ,\ e_{k+1} = \frac{y_{k+1}}{\|y_{k+1}\|} .$$

Obviously, $\{e_1, \cdots, e_{k+1}\}$ is orthonomal, and

$$\operatorname{span}\{e_1, e_2, \ldots, e_{k+1}\} = \operatorname{span}\{x_1, x_2, \ldots, x_{k+1}\} . \qquad \square$$

Theorem 2.10 (Riesz-Fischer Theorem)**.** *$\mathcal{E} = \{e_i\}_{i\in I}$ is orthonomal. $\{\lambda_i\}_{i\in I}$ are scalars.*

(a) If $\sum_{i\in I}\lambda_i e_i$ converges, then $\sum_{i\in I}|\lambda_i|^2$ converges, and

$$\left\|\sum_{i\in I}\lambda_i e_i\right\|^2 = \sum_{i\in I}|\lambda_i|^2 .$$

(b) *If H is a Hilbert space, the convergence of* $\sum_{i\in I} |\lambda_i|^2$ *and* $\sum_{i\in I} \lambda_i e_i$ *are equivalent.*

Proof. (a). Suppose $\sum_{i\in I} \lambda_i e_i$ converges to x. Then for any $\epsilon > 0$, there exists a finite subset S of I, depending on ϵ, so that for any finite subset T containing S,

$$\left\| \sum_{i\in T} \lambda_i e_i \right\|^2 = \sum_{i\in T} |\lambda_i|^2 \leq \|x\|^2 + \epsilon\,.$$

Therefore, $\sum_{i\in I} |\lambda_i|^2$ converges, and

$$\sum_{i\in I} |\lambda_i|^2 = \|x\|^2\,.$$

(b). Firstly, we suppose that I is countable. Without loss of generality, let $I = \mathbb{N}$. Then if $\sum_{i=1}^{\infty} |\lambda_i|^2$ converges, since H is complete, $\sum_{i=1}^{\infty} \lambda_i e_i$ absolutely converges. Thus $\sum_{i\in I} \lambda_i e_i$ converges (as a net). If I is uncountable, since

$$\sum_{i\in I} |\lambda_i|^2 < \infty\,,$$

There exists a countable subset J of I so that $i \in J$ if and only if $\lambda_i \neq 0$. So $\sum_{i\in I} \lambda_i e_i$ converges, by the first step. □

Lemma 2.11. *Let $\{e_1, \cdots, e_n\}$ be orthonarmal. Let $M = \operatorname{span}\{e_1, \cdots, e_n\}$. Clearly, M is complete because it is finite dimensional. Let $x \in H$.*

(a) $\hat{x} = \sum_{i=1}^{n} \langle x, e_i\rangle e_i$ *is the projection of x on M, and*

$$\|\hat{x}\|^2 = \sum_{i=1}^{n} |\langle x, e_i\rangle|^2 = \|x\|^2 - \|x - \hat{x}\|^2 \leq \|x\|^2.$$

(b) *For an $\lambda_i \in \mathbb{F}, i = 1, 2, \cdots, n$, we have*

$$\|x - \sum_{i=1}^{n} \langle x, e_i\rangle e_i\| \leq \|x - \sum_{i=1}^{n} \lambda_i e_i\|\,.$$

Theorem 2.12 (Bessel Inequality)**.** *Let $\mathcal{E} = \{e_i\}_{i\in I}$ be orthonarmal. Then, for any $x \in H$, the series $\sum_{i\in I} |\langle x, e_i\rangle|^2$ converges, and*

$$\sum_{i\in I} |\langle x, e_i\rangle|^2 \leq \|x\|^2 .$$

Proof. Using Lemma 2.11 and Proposition A.10, the theorem follows. $\square$

Theorem 2.13 (Fourier series)**.** *Let $\mathcal{E} = \{e_i\}_{i\in I}$ be orthonarmal. Let $x \in H$. Then the propositions following are equivalent.*

(a) The Fourier series of x with respect to $\mathcal{E}$ converges to x , in other words, $x = \sum_{i\in I}\langle x, e_i\rangle e_i$.

(b) Parseval Equality holds: $\|x\|^2 = \sum_{i\in I} |\langle x, e_i\rangle|^2$.

(c) $x \in \overline{\operatorname{span}}(\mathcal{E})$.

Proof. Using the property of net convergence, we can easily prove that (a), (b) are equivalent. Obviously, (a) implies (c). So we only prove (c) implies (a). To this end, take any $x \in \overline{\operatorname{span}}(\mathcal{E})$. Then for each $\epsilon > 0$, there exists $e_{i_1}, \cdots, e_{i_n} \in \mathcal{E}$ and $\lambda_1, \cdots, \lambda_n \in \mathbb{F}$ such that

$$\|x - \sum_{k=1}^{n} \lambda_i k e_{i_k}\| \leqslant \epsilon.$$

Using Lemma 2.11 we have

$$\|x - \sum_{k=1}^{n}\langle x, e_{i_k}\rangle e_{i_k}\| = \|x\|^2 - \sum_{k=1}^{n} |\langle x, e_{i_k}\rangle|^2 \leqslant \epsilon$$

So

$$\|x\|^2 - \sum_{i\in I} |\langle x, e_i\rangle|^2 \leqslant \epsilon$$

We let $\epsilon \downarrow 0$, then we get $\|x\|^2 = \sum_{i\in I} |\langle x, e_i\rangle|^2$. $\square$

Remark 2.4. There exists the projection of x on $\overline{\mathrm{span}}(\mathcal{E})$ if and only if $\sum_{i\in I}\langle x, e_i\rangle e_i$ coverges, when the projection $\hat{x}$ exists, we have

$$\hat{x} = \sum_{i\in I}\langle x, e_i\rangle e_i$$

Definition 2.6. We say an orthonomal set $\mathcal{E} = \{e_i\}_{i\in I}$ is **complete**, if every $x \in H$, $x = \sum_{i\in I}\langle x, e_i\rangle e_i$ holds.

We can see form Theorem 2.13 that an orthonomal set $\mathcal{E} = \{e_i\}_{i\in I}$ is complete is equivalent to the follows propositions. (a) $\overline{\mathrm{span}}(\mathcal{E}) = H$. (b) for any $x \in H$, the Parseval equality holds. (c), for all $x, y \in H$,

$$\langle x, y\rangle = \sum_{i\in I}\langle x, e_i\rangle\overline{\langle y, e_i\rangle}\,.$$

Definition 2.7. An orthonomal set $\mathcal{E} = \{e_i\}_{i\in I}$ is called **total**, if $\mathcal{E}^{\perp} = \{0\}$.

Obviously, a completely orthnormal set must be total, and if H is a Hilbert space, a totally orthnormal set must be complete, which is also called a *Hilbert basis* of H. But in the general case, the converse doesn't hold well. See 夏道行, 《实变函数论与泛函分析下册》, Page 127.

2.4 Isometric Isomorphism

Every mathematical theory has its concept of "isomorphism". In topology there is homeomorphism and homotopy equivalence, algebra calls them isomorphisms. The basic idea is to define a map which preserves the basic structure of the spaces in the category. In this section, We introduce the isometric isomorphism between Hilbert spaces.

If H and H' are inner product spaces, an *isometric isomorphism* between H and H' is a linear bijective isometry. In this case H and H' are said to be *isometricly isomorphic*.

It is easy to see that if $U : H \to H'$ is an isometric isomorphism, then so is $U^{-1} : H' \to H$. Similar such arguments show that the concept of "isometricly isomorphic" is an equivalence relation on inner product spaces.

Proposition 2.14. *If $U : H \to H'$ is bijective. Then U is an isometric isomorphism if anf only if for any $x, y \in H$,*

$$\langle Ux, Uy \rangle = \langle x, y \rangle .$$

Proof. This is a direct result of the polarization identity. □

In linear algebra, we have learned that every two finite dimensional linear space are isomorphic if and only if they have the same dimension. Therefore, it's nature to do the same thing : give a proper definition about "definition" about Hilbert space. The Hilbert basis occurs to us mind easily.

Definition 2.8. The *dimension* of Hilbert space H is the cardinality of a Hilbert basis, and is denoted by $\dim H$.

Is the definition well-defined ?

Lemma 2.15. *Any two Hilbert bases of Hilbert space H have the same cardinality.*

Proof. Let $\{\varepsilon_i\}_{i\in I}$ and $\{e_j\}_{j\in J}$ be two Hilbert basis for H. For any $j \in J$, e_j has a Fourier expansion

$$e_j = \sum_{i\in I} \langle \varepsilon_i, e_j \rangle \varepsilon_i$$

Let $I_j = \{i \in I : \langle \varepsilon_i, e_j\rangle \neq 0\}$, then I_j is countable, and we have $\cup_{j\in J} I_j = I$. Therefore

$$|I| = \left| \bigcup_{j\in J} I_j \right| \leqslant \aleph_0 |J| = |J| \, .$$

For the same reason, $|J| \leqslant |I|$, so $|I| = |J|$. □

Theorem 2.16. *H is an infinite dimensional Hilbert space. Then* $\dim H$ *is* $\aleph_0$ *if and only if H is separable.*

Proof. If $\dim H$ is $\aleph_0$, by Lemma 1.14 we have H is seperable. If H is separable, then there is a countable dense subset $\{x_n\}_{n\geq 1}$. Without loss of generality, we assume it is linearly independent. Using Gram-Schmidt orthonormalisation procedure we get a countable Hilbert basis. □

Now, we can give the isomorphism theorem in Hilbert space which likes to finite linear space .

Theorem 2.17. *Two Hilbert spaces are isomorphic if and only if they have the same dimension. Particularly, all separable infinite dimensional Hilbert spaces are isomorphic to ℓ^2.*

Proof. Let $\{e_i\}_{i\in I}$ be an orthonormal basis for H. Define $T : H \to \ell^2(I)$ by

$$Tx = (\langle x, e_i \rangle)_{i\in I} \text{ for each } x \in H.$$

It follows from Bessel's Inequality that the right hand side is in $\ell^2(I)$. We must show that T is a surjective linear isometry. Clearly, T is linear. By Theorem 2.10, T is a surjection. By Theorem 2.13, T isometry. □

Example 2.8. If for each $k \in \mathbb{Z}$,

$$e_k(t) := \frac{1}{\sqrt{2\pi}}\, \mathrm{e}^{ikt}, \quad t \in [0, 2\pi]\,,$$

then $\{e_k : k \in \mathbb{Z}\}$ is a basis for $L^2([0, 2\pi], \mathbb{C})$. In fact, by using trigonometric polynomials to uniformly approach continuous function, we can prove

$$\overline{\mathrm{span}}\{e_k\} = L^2([0, 2\pi], \mathbb{C})\,.$$

Then for any $f \in L^2([0, 2\pi], \mathbb{C})$, let

$$\widehat{f}(k) := \langle f, e_k\rangle = (2\pi)^{-1/2} \int_0^{2\pi} f(t) e^{-ikt} dt$$

is called the *kth Fourier coefficient of f*, k in $\mathbb{Z}$, and we have

$$f = \sum_{k=-\infty}^{\infty} \widehat{f}(k) e_k\,,$$

where this infinite series converges to f in the metric defined by the norm of $L^2([0, 2\pi], \mathbb{C})$. This is the classic Fourier series.

For f in $L^2([0, 2\pi], \mathbb{C})$, the function $\widehat{f} : \mathbb{Z} \to \mathbb{C}$ is called the Fourier transform of f; the map $U : L^2([0, 2\pi], \mathbb{C}) \to l^2(\mathbb{Z})$ defined by $Uf = \widehat{f}$ is the Fourier transform. As we can see,

The Fourier transform is a linear isometry from $L^2([0, 2\pi], \mathbb{C})$ onto $\ell^2(\mathbb{Z})$.

Chapter 3

Topology Vector Space and Locally Convex Space

There still are many important spaces carrying natural topologies that cannot be induced by norms. Here are some examples:

(a) $C(\Omega)$, the space of all continuous complex functions on some (unbounded) open set Ω in a euclidean space $\mathbb{R}^n$.

(b) $H(\Omega)$, the space of all holomorphic functions in some open set Ω in the complex plane.

(c) C_K^∞, the space of all infinitely differentiable complex functions on $\mathbb{R}^n$ that vanish outside some fixed compact set K with nonempty interior.

So we need a generalization of the concept of a Banach space to describe these spaces, that is topological vector space. As a special case for topological vector space, the locally convex spaces are encountered repeatedly when discussing weak topologies on a Banach space, sets of operators on Hilbert space, or the theory of distributions. We will only skim the surface of this theory, but it will treat locally convex spaces in sufficient detail as to enable us to understand the use of these spaces in the three areas of analysis just mentioned.

3.1 Elementary properties

A vector space equipped with a Hausdorff topology such that the linear structure and the topological structure are *vitally connected*, is called a topological vector space. Here is a more precise way of stating the definition:

Definition 3.1. A **topological vector space** (**TVS**) is a vector space X together with a topology τ such that

(a) (X, τ) is Hausdorff space.

(b) the vector space operations are continuous with respect to τ.

and such topology τ is called a **vector topology** on X.

Remark 3.1. In many texts, (a) is omitted from the definition of a topological vector space. Since (a) is satisfied in almost every application, and since most theorems of interest require (a) in their hypotheses, it seems best to in clude it in the axioms. Later, we willsee that under condition (b), (X, τ) is regular sapce. So (a) can be reduced that X is T_1 space.

Invariance of the local base

Let X be a topological vector space. Associate to each $a \in X$ and to each scalar $\lambda \neq 0$ the *translation operator* T_a and the *multiplication operator* M_λ, by the formulas

$$T_a(x) = a + x, \quad M_\lambda(x) = \lambda x \quad (x \in X)$$

The following simple proposition is very important:

Proposition 3.1. *T_a and M_λ are homeomorphisms of X onto X*

Proof. The vector space axioms alone imply that T_a and M_λ are one-to-one, that they map X onto X, and that their inverses are T_{-a} and $M_{1/\lambda}$, respectively. The assumed continuity of the vector space operations implies that all these mappings are continuous. Hence each of them is a homeomorphism. $\square$

One consequence of this proposition is that every vector topology τ is *translation-invariant* (or simply invariant, for brevity): A set $U \in X$ is open if and only if each of its translates $a + U$ is open. Thus τ is completely determined by any local base.

In the vector space context, the term **local base** will always mean a local base at 0. A local base of a topological vector space X is thus a collection $\mathcal{B}$ of open neighborhoods of 0 such that every neighborhood of 0 contains a member of $\mathcal{B}$. The open sets of X are then precisely those that are unions of translates of members of $\mathcal{B}$.

We check the definition that addition is continuous , which means that the mapping

$$(x, y) \mapsto x + y$$

of the Cartesian product $X \times X$ into X is continuous: if $x_i \in X$ for $i = 1, 2$, and if U is a neighborhood of 0 there should exist open neighborhoods V_i of 0 such that $(x_1 + V_1) + (x_2 + V_2) \subset x_1 + x_2 + V$. Thus let $V = V_1 \cap V_2$, we have

$$V + V \subset U\,.$$

Similarly, the assumption that scalar multiplication ts continuous means that the mapping

$$(\lambda, x) \mapsto \lambda x$$

of $\mathbb{F} \times X$ into X is continuous. Thus for any U is a neighborhood of 0, then for some $\epsilon > 0$ and some open neighborhood V of 0, we have $\lambda V \subset V$ whenever $|\lambda| \leqslant \epsilon$. If we let

$$W = \bigcup_{|\lambda| \leqslant \epsilon} \lambda V\,,$$

then W is a balanced open neighborhood of 0 contained in U. Thus we have

Proposition 3.2. *Every topology vector space has a balanced local base.*

Separation properties

Theorem 3.3. *X is a topological vector space, K is compact, C is closed, and $K \cap C = \varnothing$. Then there is a neighborhood V of 0 such that*

$$(K+V) \cap (C+V) = \varnothing .$$

Proof. For any $x \in K$, there is a balanced open neighborhood V_x of 0 such that $(x + V_x + V_x) \cap C = \varnothing$. Since K is compact, there are finitely many points $x_1, \dots, x_n$ in K such that

$$K \subset (x_1 + V_{x_1}) \cup \cdots \cup (x_n + V_{x_n})$$

Put $V = V_{x_1} \cap \cdots \cap V_{x_n}$. Then

$$K + V \subset \bigcup_{i=1}^{n} (x_i + V_{x_i} + V) \subset \bigcup_{i=1}^{n} (x_i + V_{x_i} + V_{x_i})$$

and no term in this last union intersects $C+V$. This completes the proof. □

Remark 3.2. In the proof we have not uses the assumption that X is Hausdorff space. Let K be any single point, then we konw X is regular. Thus X is T_1 imples X is Hausdorff, see the Remark of Definition 3.1.

Some other simple properties is followed. We omit the proof

Proposition 3.4. *Let X be a TVS, $A, B \subset X$.*

(a) $\overline{A} = \cap(A+V)$, *where V runs through all neighborhoods of* 0.

(b) $\overline{A} + \overline{B} \subset \overline{A+B}$

(c) *If Y is a subspace of X, so is $\overline{Y}$.*

We can also discuss whether the convexity, balance still holds under the topology operations.

Theorem 3.5. *Let X be a TVS, K is subset of X.*

(a) *If K is convex, then so are $\overline{K}$ and K°.*

(b) *If B is balanced, so is $\overline{B}$. In addition, $0 \in B^\circ$ then B° is balanced.*

Types of topological vector spaces

Definition 3.2. X is TVS, a subset B is called **bounded** if for any neighborhood U of 0, there exist some $\epsilon > 0$, depending on U, such that $\epsilon B \subset U$.

Definition 3.3. X is a topological vector space, with topology τ. We say

(a) X is **locally convex** if there is a local base $\mathcal{B}$ whose members are *convex.*

(b) X is **locally bounded** if 0 has a *bounded* neighborhood.

(c) X is **locally compact** if 0 has a neighborhood whose closure is compact.

(d) X is **metrizable** if τ is compatible with some metric d.

(e) X is an **F-space** if τ is induced by a complete *translation invariant* metric.

(f) X is a **Fréchet space** if X is a locally convex F-space.

(g) X is **normable** if a norm exists on X such that the metric induced by the norm is compatible with τ.

The terminology of (e) and (f) is not universally agreed upon: In some texts, local convexity is omitted from the definition of a Frechet space, whereas others use F-space to describe what we have called Frechet space.

Relations

Here is a list of some relations between these properties of a topological vector space X.

(a) If X is locally bounded, then X has a countable local base.

(b) X is metrizable *iff* X has a countable local base.

(c) X is normable *iff* X is locally convex and locally bounded.

(d) X has finite dimension *iff* X is locally compact.

(e) If a locally bounded space X has the Heine-Borel property, i.e., every closed bounded subset of X is compact, then X has finite dimension.

The spaces $H(\Omega)$ and C_B^∞ mentioned before are infinite-dimensional Fréchet spaces with the Heine-Borel property, they are therefore not locally bounded, hence not normable; they also show that the converse of (a) is false.

On the other hand, there exist locally bounded F-spaces that are not locally convex : L^p space when $p \in (0, 1)$.

3.2 Locally convex spaces

Firstly, we give an example of locally convex space. Let X be a vector space, we konw that for any seminorm p on X, given $\epsilon > 0$,

$$V(p, \epsilon) \coloneqq \{x : p(x) < \epsilon\}$$

is a balanced absorbing convex set. If use these sets to induce a vector topology on X, it must be locally convex.

Example 3.1. X be a vector space, $\mathcal{P}$ is a family of seminorms on X. Let τ be the smallest topology on X staisfying p is continuous for all $p \in \mathcal{P}$. It's easy to find a subbase of τ. For any $x_0 \in X$, let

$$V(x_0, \Phi, \epsilon) \coloneqq \{x : p(x - x_0) < \epsilon \text{ for all } p \in \Phi\},$$

where $\Phi \subset \mathcal{P}$ is finite, $\epsilon > 0$. Then it's easy to check that

$$\{V(x_0, \Phi, \epsilon) : \Phi \subset \mathcal{P} \text{ is finite, } \epsilon > 0\}$$

is a balanced convex neighborhood base at x_0.

$V(0, \Phi, \epsilon)$ sometimes is wirted as $V(\Phi, \epsilon)$ for short. It's easy to find that,

$$V(x_0, \phi, \epsilon) = x_0 + V(\phi, \epsilon).$$

Now we try to make (X, τ) be a TVS, then since 0 has a (balanced) convex local base, it is a LCS.

(a) To make (X, τ) be Hausdorff, we should assume that $\mathcal{P}$ is **seperating**:

$$\bigcap_{p \in \mathcal{P}} \{x : p(x) = 0\} = \{0\}. \tag{3.1}$$

In fact, suppose that $x \neq y$. Then there is a p in $\mathcal{P}$ such that $p(x-y) \neq 0$, pick an $\epsilon > 0$ such that $p(x-y) > 2\epsilon$. Then $x+V(p, \epsilon)$ and $y+V(p, \epsilon)$ are disjoint neighborhoods of x and y, respectively. And this condition is necessary too.

(b) For any $V(\Phi,\epsilon)$, $V(\Phi,\epsilon/2)+V(\Phi,\epsilon/2)\subset V(\Phi,\epsilon)$, so addition is continous. Pick $\alpha_0\in\mathbb{F}$ and $x_0\in X$, for any $\alpha_0x_0+V(\Phi,\epsilon)$, let $|\alpha-\alpha_0|\leqslant\delta_1$ and $x\in x_0+V(\Phi,\delta_2)$, then for all $p\in\Phi$,

$$p(\alpha x-\alpha_0x_0)\leqslant|\alpha-\alpha_0|p(x_0)+\alpha p(x-x_0)<\delta_1p(x_0)+(|\alpha_0|+|\delta_1|)\delta_2<\epsilon\,.$$

Thus scalars mutipulication is continuous.

Thus we proved (X,τ) is a LCS, and

$$\{V(\Phi,\epsilon):\ \Phi\subset\mathcal{P}\text{ is finite, }\epsilon>0\}$$

is a local base. Since τ is induced by the separating family $\mathcal{P}$ of seminorms, there are two interesting consequences.

(i) $B\subset X$ is bounded if and only if $\{p(x):x\in B\}$ is bounded for any $p\in\mathcal{P}$.

(ii) A net $\{x_i\}$ in X is convergent to some x_0 if and only if $p(x_i)\to p(x_0)$ for all $p\in\mathcal{P}$.

If $\mathcal{P}$ is a family of seminorms of X that makes X into a LCS, it is often convenient to enlarge $\mathcal{P}$ by assuming that $\mathcal{P}$ is closed under the formation of finite sums and supremums of "bounded families". Sometimes it is convenient to assume that $\mathcal{P}$ consists of all continuous seminorms. In either case the resulting topology on X remains unchanged.

Seminorms and local convexity

Proposition 3.6. *Let X be a TVS and let p be a seminorm on X. The following statements are equivalent.*

(a) *p is continuous.*

(b) $0\in\mathrm{int}V(p,1)$.

(c) *p is continuous at 0 .*

Proof. (a) implies (b) is obviously.

To show (b) implies (c), since $0 \in \text{int}\{x : p(x) < 1\}$, there exists some neighborhood U of 0 contained in $V(p, 1)$. Note that $p(0) = 0$, for any $\epsilon > 0$, each $x \in \epsilon U \subset \epsilon V(p, 1) = V(p, \epsilon)$, so p is continuous at 0.

(c) implies (a) is obviously. □

Corollary 3.7. *Let X be a TVS and let p be a seminorm. Then p is continuous iff there is a continuous seminorm q such that $p \leqslant q$.*

Theorem 3.8. *Every locally convex space has a balanced convex local base.*

Proof. Suppose U is a convex neighborhood of 0. Note that U is convex and absorbing, then $rU \subset U$ for any $0 \leqslant r \leqslant 1$. A balanced convex open neighborhood of 0 contained in U must contained in

$$V = \bigcap_{|\alpha|=1} \alpha U\,.$$

Firstly, being an intersection of convex sets, V is convex. Secondly, there is a balanced neighboehood W of 0 contained in U, then $\alpha W = W \subset \alpha U$ for $|\alpha| = 1$, thus $W \subset V$, Thus, the interior V° of V is a convex neighborhood of 0.

Besides, we can see that V is balanced: choose r and β so that $0 \leqslant r \leqslant 1, |\beta| = 1$.

$$r\beta V = \bigcap_{|\alpha|=1} r\beta\alpha U = \bigcap_{|\alpha|=1} r\alpha U \subset V$$

Then V° is balanced since $0 \in V^\circ$. Thus, V° is balanced convex open neighborhood of 0 contained in U, so LCS has a balanced convex local base. □

Theorem 3.9. *Suppose $\mathcal{B}$ is a convex balanced local base in a LCS (X, τ). Associate to every $U \in \mathcal{B}$ its Minkowski functional p_U. Then $\mathcal{P} = \{p_U : U \in \mathcal{B}\}$ is a separating family of continuous seminorms on X, which induced τ.*

Proof. First, we show $\mathcal{P}$ is separating. If $p_U(x) = 0$ for all $U \in \mathbb{B}$, $x \in U$ for all U. Since τ is Hausdorff, $x = 0$.

Secondly, let τ_1 is the topology induced by p, we show $\tau_1 = \tau$. Since U is open, U is absorbing at each of it's point, then $V(p_U, 1) = U$ by Lemma 1.20, then p_U is continuous for all $U \in \mathcal{B}$. Thus $\tau_1 \subset \tau$. Note that τ is determined by it's local base $\mathcal{B}$, and for any $U \in \mathcal{B}$, $U = V(p_U, 1) \in \tau_1$, so $\tau \subset \tau_1$. $\square$

Metrizable and normable LCS*

Which LCS's are metrizable? That is, which have a topology which is defined by a metric? Which LCS's have a topology that is defined by a norm? Both are interesting questions and both answers could be useful.

The proof of the following two theorems can be found in J. Conway, *A Course in Functional Analysis*, Chapter IV, Section 2.

Theorem 3.10. *Let (X, τ) be a LCS , $\mathcal{P} = \{p_n : n \in \mathbb{N}\}$ is countable seperating family of seminorms which induces τ. For x and y in X, define*

$$d(x, y) = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{p_n(x - y)}{1 + p_n(x - y)} \,.$$

Then d is a translation invariant metric, and the topology defined by d coinsides with τ.

Corollary 3.11. *LCS is metrizable iff its topology is determined by a countable seperating family of seminorms.*

Theorem 3.12. *LCS is normable if and only if it has a nonempty bounded open set.*

3.3 Linear mappings

Here are some properties of linear mappings $\Lambda : X \to Y$ whose proofs are so easy that we omit them; it is assumed that $A \subset X$ and $B \subset Y$:

(a) If A is a subspace (or a convex set, or a balanced set) the same is true of $\Lambda(A)$.

(b) If B is a subspace (or a convex set, or a balanced set) the same is true of of $\Lambda^{-1}(B)$.

The following proposition is obviously .

Proposition 3.13. *Let X and Y be topological vector spaces. If $\Lambda : X \to Y$ is linear and continuous at 0, then Λ is continuous.*

Now we turn to discuss continuous linear functionals on TVS, namely (X, τ), and we denote by $(X, \tau)^*$ or X^* all the continuous linear functional on (X, τ).

Theorem 3.14. *X is a TVS. Let f be a nonzero linear functional on X. Then each of the following three properties implies the other three:*

(a) f is continuous.

(b) The null space $\ker f$ is closed.

(c) f is locally bounded, i.e., there exists $M > 0$ and a neighborhood V of 0 such that $|f(x)| \leqslant M$ for all $x \in V$.

Proof. (a) implies (b) is obviously.

To show (b) implies (c), since $\ker f$ is proper closed subspace, pick any $x \notin \ker f$. Then there exists a balanced neighborhood V of 0, suth that

$$(x + V) \cap \ker f = \varnothing\,.$$

Since $f(V)$ is balanced subset of $\mathbb{F}$, and $-f(x) \notin f(V)$, $f(V)$ must be bounded.

(c) implies (a) : if (c) holds, then $|f(x)| < M$ for all x in V and for some $M < \infty$. Then For any $\epsilon > 0$, let

$$W = \frac{\epsilon}{M} V \,,$$

then $|f(x)| < \epsilon$ for every x in W, hence f is continuous at the origin, and then f is continuous. $\square$

Theorem 3.15. *X is LCS, $\mathcal{P}$ is a separating family of seminorms that defines the topology on X, then f is continuous if and only if there are $p_1, \dots, p_n$ in $\mathcal{P}$ and positive scalars $\alpha_1, \dots, \alpha_n$ such that*

$$|f(x)| \leqslant \sum_{k=1}^{n} \alpha_k p_k(x) \,, \quad \text{for all } x \in X.$$

Proof. Sufficiency is easy. We show necessity. If f is countinuous, there exists $\Phi = \{p_1, \cdots, p_n\} \subset \mathcal{P}$ and $\epsilon > 0$ such that

$$\{x : |f(x)| < 1\} \subset V(\phi, \epsilon) \,.$$

Let

$$q(x) := \sum_{k=1}^{n} \frac{1}{\epsilon} p_k(x) \,, \quad \text{for all } x \in X.$$

Obviously, q is a (continuous) seminorm, and

$$V(q, 1) \subset \{x : |f(x)| < 1\} \,.$$

Thus for any $x \in X$, and $\delta > 0$,

$$q\left(\frac{x}{q(x) + \delta}\right) < 1 \Rightarrow \frac{|f(x)|}{q(x) + \delta} < 1 \Rightarrow |f(x)| \leqslant q(x) \,. \qquad \square$$

3.4 Finite-dimensional spaces

We have shown that on finite-dimensional vector sapce X, all the norm topologys are the same one. It's natural to ask is it to for vector topologys?

Lemma 3.16. *X is TVS, and $f : \mathbb{F}^n \to X$ is linear, then f is continuous.*

Proof. Let $\{e_1, \ldots, e_n\}$ be the standard basis of $\mathbb{F}^n$. Put $u_k = f(e_k)$, for $k = 1, \ldots, n$ Then

$$f(z) = z_1 u_1 + \cdots + z_n u_n \text{ for every } z = (z_1, \ldots, z_n) \in \mathbb{F}^n.$$

Every z_k is a continuous function of z. The continuity of f is therefore an immediate consequence of the fact that addition and scalar multiplication are continuous in X. □

Theorem 3.17. *X is an n-dimensional TVS. Then every isomorphism of $\mathbb{F}^n$ onto X is a linear homeomorphism.*

Proof. Let S be the sphere which bounds the open unit ball B of $\mathbb{F}^n$, i.e.,

$$S = \{\lambda \in \mathbb{F}^n : \Sigma |\lambda_i|^2 = 1, \}$$

Suppose $f : \mathbb{F}^n \to X$ is an isomorphism, That is, f is a linear bijection. Since f is continuous, $f(S)$ is compact. Note that $0 \notin f(S)$, then there is a balanced neighborhood V of 0 in X which does not intersect $f(S)$, i.e.,

$$V \cap f(S) = \varnothing .$$

The set

$$E = f^{-1}(V)$$

is therefore disjoint from S. Since f is linear, V is balanced, hence E is balanced. Note that $0 \in E$, thus $E \subset B$. This implies that the linear map f^{-1} takes V into B. This implies that f^{-1} is locally bounded, by Theorem 3.14 (c) we have f^{-1} is continuous. Thus f is a homeomorphism. □

Corollary 3.18. *X is a finite-dimensional vector sapce and τ_1, τ_2 are two vector topologys on X, then $\tau_1 = \tau_2$.*

Proof. Let f be a linear isomorphism between X and $\mathbb{F}^n$, then f is linear homeomorphism betwwen $\mathbb{F}^n$, (X, τ_1) and (X, τ_2), so $\tau_1 = \tau_2$. □

Sufficient and necessary condition*

Theorem 3.19. *TVS has finite dimension iff it is locally compact.*

Proof. The origin of X has a neighborhood V whose closure is compact. V is bounded, and the sets $2^{-n}V(n = 1, 2, 3, \ldots)$ form a local base for X. The compactness of $\overline{V}$ shows that there exist $x_1, \ldots, x_m$ in X such that

$$\overline{V} \subset \left(x_1 + \frac{1}{2}V\right) \cup \cdots \cup \left(x_m + \frac{1}{2}V\right)$$

Let Y be the vector space spanned by $x_1, \ldots, x_m$. Then $\dim Y \leq m$. Thus Y is a closed subspace of X. Since $V \subset Y + \frac{1}{2}V$ and since $\lambda Y = Y$ for every scalar $\lambda \neq 0$, it follows that

$$\frac{1}{2}V \subset Y + \frac{1}{4}V$$

If we continue in this way, we see that

$$V \subset \bigcap_{n=1}^{\infty} \left(Y + 2^{-n}V\right)$$

Since $\{2^{-n}V\}$ is a local base, we have $V \subset \overline{Y}$. But $\overline{Y} = Y$. Thus $V \subset Y$, which implies that $kV \subset Y$ for $k \in \mathbb{N}_+$. But $X = \cup_{k=1}^{\infty} kV$, thus $Y = X$.. □

Corollary 3.20. *If X is a locally bounded topological vector space with the Heine-Borel property: every closed and bounded subset of X is compact. Then X has finite dimension.*

3.5 Quotient spaces

Let N be a subspace of a vector space X. For every $x \in X$, let $Q(x)$ (sometimes write $\tilde{x}$ or $[x]$) be the coset of N that contains x, thus

$$Q(x) = x + N.$$

These cosets are the elements of a vector space X/N, called the quotient space of X modulo N, in which addition and scalar multiplication are defined by

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

Since N is a vector space, the operations are well defined.

The origin of X/N is $Q(0) = N$. Q is a linear mapping of X onto X/N with N as its null space; Q is often called the **quotient map** or the **natural map** of X onto X/N.

Suppose now that τ is a vector topology on X and that M is a subspace of X. Let τ_M be the *quotient topology* on X/M. That is

$$\tau_M = \{U \in X/M : Q^{-1}(U) \in \tau\}. \tag{3.2}$$

To guarantee that τ_M is T_1 space, $Q^{-1}(Q(x)) = x + M$ must be closed in τ, which is equivalent to that M is *closed*. So we always assume M is closed subspace of X.

Theorem 3.21. *Let M be a closed subspace of a topological vector space X. Let τ be the topology of X and τ_M is quotient topology on X/M.*

(a) The quotient map $Q : X \to X/M$ is linear continuous open mapping.

(b) τ_M is a vector topology on X/M.

(c) If $\mathcal{B}$ is a local base for τ, then $Q(\mathcal{B})$ is a local base for $(X/M, \tau_M)$.

Proof. To show (a), note that the continuity of Q follows directly from the definition of τ_M. Mext, suppose $V \in \tau$. Since

$$Q^{-1}(Q(V)) = M + V$$

and $M + V \in \tau$, it follows that $Q(V) \in \tau_M$. Thus Q is an open mapping.

To show (b), if now W is a neighborhood of 0 in X/M, there is a neighborhood V of 0 in X such that

$$V + V \subset Q^{-1}(W)$$

Hence $Q(V) + Q(V) \subset W$. since Q is open, $Q(V)$ is a neighborhood of 0 in X/M. Addition is therefore continuous in X/M. The continuity of scalar multiplication in X/M is proved in the same manner. This establishes (b).

It is clear that (a) implies (c). □

Corollary 3.22. *Suppose M and F are subspaces of a topological vector space X, M is closed, and F has finite dimension. Then $M + F$ is closed.*

Proof. Let Q be the quotient map of X onto X/M, and give X/M its quotient topology. Then $Q(F)$ is a finite-dimensional subspace of X/M, since X/M is TVS, $Q(F)$ is subspace of X/M, so is closed in X/M. Since $M + F = Q^{-1}(Q(F))$ and Q is continuous, we conclude that $M + F$ is closed. □

Corollary 3.23. *Each of the following properties of X is inherited by X/N: local convexity, local bounded ness, metrizability, normability.*

Proof. Since all of these properties are determined by the local base, and Q do not change these properties. □

Other definitions

If X is a LCS and $\mathcal{P}$ is the separating family of seminorms on X, induced the topology.For any seminorm p on X, define $\tilde{p}$ on X/M by

$$\tilde{p}(\tilde{x}) = \inf\{p(x+m) : m \in M\}\,.$$

then $\tilde{p}$ is a seminorm on X/M. The family $\tilde{\mathcal{P}} := \{\tilde{p} : p \in \mathcal{P}\}$ is a seperating family of seminorms on X/M, and it follows from Theorem 3.21 (c) that $\tilde{\mathcal{P}}$ induces the quotient topology on X/M.

Suppose next that d is an invariant metric on X, compatible with τ. Define ρ by

$$\tilde{\rho}(\tilde{x}, \tilde{y}) = d(x-y, M) = \inf\{d(x-y, m) : m \in M\}\,,$$

is interpreted as the distance from $x - y$ to M. We omit the verifications that are now needed to show that $\tilde{\rho}$ is well defined and that it is an invariant metric on X/N. Since

$$Q(\{x : d(x,0) < r\}) = \{\tilde{x} : \tilde{\rho}(\tilde{x}, 0) < r\}\,,$$

it follows from Theorem 3.21 (c) that $\tilde{\rho}$ is compatible with τ_M.

If X is normed, this definition of ρ specializes to yield what is usually called the quotient norm of X/M :

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

3.6 Hahn-Banach theorems

Separation theorems

Theorem 3.24. *X is TVS, A and B are two disjoint, nonempty, convex subsets. If A is open, there exist $\ell \in X^*$ and $c \in R$ such that*

$$\operatorname{Re} \ell x < c \leqslant \operatorname{Re} \ell y \tag{3.3}$$

for every $x \in A$ and for every $y \in B$.

Proof. It is enough to prove this for real scalars.

Fix $a_0 \in A, b_0 \in B$,and put $x_0 = b_0 - a_0$, put $C = A - B + x_0$. Then C is a convex open neighborhood of 0 in X. Denote by p the Minkowski functional of C. Since $x_0 \notin C$, $p(x_0) \geqslant 1$.

Just like in Theorem 1.22, we define ℓ on $\operatorname{span}\{x_0\}$ by

$$\ell(ax_0) = a \,, \text{ for all } a \in \mathbb{R} \,.$$

Then for all such ax_0

$$\ell(ax_0) \leq p_K(ax_0) \,.$$

So by H-B extension theorem, ℓ can be extened as a linear functional on X domainted by p. Note that $p(x) \leqslant 1$ for all $x \in C$, then

(i) ℓ is continuous, since ℓ is locally bounded (Theorem 3.14 (c) .) indeed,

$$|\ell(x)| < 1 \,, \quad \text{for all } x \in C \cap (-C).$$

(ii) $\ell(a) < \ell(b)$ for all $a \in A, b \in B$, since

$$\ell a - \ell b + 1 = \ell \left(a - b + x_0\right) \leq p \left(a - b + x_0\right) < 1$$

It follows that $\ell(A)$ and $\ell(B)$ are disjoint convex subsets of R, with $\ell(A)$ to the left of $\ell(B)$.

The key is that: *every nonconstant linear functional on X is an open mapping.* Since $\ell(A) = \tilde{\ell}(Q(A))$, where $Q : X \to X/\mathrm{ker} f$ and $\tilde{\ell}(\tilde{x}) = \ell(x)$, Q is open mapping, $\tilde{\ell}$ is linear homeomorphism between $X/\mathrm{ker}\ell$ and $\mathbb{R}$, by Theorem 3.17. (See Figure 3.1) Thus $\ell(A)$ is an open convex set, c be the

$$\begin{tikzcd} X \arrow[r, "Q"] \arrow[dr, "l"'] & \frac{X}{\ker l} \arrow[d, "\tilde{l}"] \\ & \mathbb{R} \end{tikzcd}$$

Figure 3.1: ℓ is open mapping

right end point of $\ell(A)$ to get the conclusion of (3.3). $\square$

Theorem 3.25. *X is LCS, A, B are two disjoint, nonempty, convex subsets. If A is compact, and B is closed, then there exist $\ell \in X^*$ and $c \in \mathbb{R}$ such that*

$$\mathrm{Re}\,\ell x < c < \mathrm{Re}\,\ell y \tag{3.4}$$

for every $x \in A$ and for every $y \in B$.

Proof. It is enough to prove this for real scalars.

By Theorem 3.3, and X is locally convex, there is a convex open neighborhood V of 0 such that

$$(A + V) \cap B = \varnothing.$$

With $A + V$ in place of A, Theorem 3.24 shows that there exists $\ell \in X^*$ such that $\ell(A+V)$ and $\ell(B)$ are disjoint convex subsets of $\mathbb{R}$, with $\ell(A+V)$ open (See Figure 3.1) and to the left of $\ell(B)$. Note that $\ell(A)$ is compact, we obtain the conclusion. $\square$

Corollary 3.26. *X is a LCS, then X^* separates points on X.*

Remark 3.3. In Theorem 3.25,

(a) the hypothesis that X is locally convex does not appear in the preceding results, since the existence of an open convex subset of X is assumed. In this theorem such a set must be manufactured. Without the hypothesis of local convexity it may be that the only open convex sets are the whole space itself and the empty set, see Example 3.2

(b) the fact that one of the two closed convex sets in the preceding theorem is assumed to be compact is necessary. In fact, if $X = \mathbb{R}^2$, $A = \{(x, y) \in \mathbb{R}^2 : y \leqslant 0\}$, and $B = \left\{(x, y) \in \mathbb{R}^2 : y \geqslant x^{-1} > 0\right\}$, then A and B are disjoint closed convex subsets of $\mathbb{R}^2$ that cannot be strictly separated.

Example 3.2. For $0 < p < 1$, let $L^p(0, 1)$ be the collection of equivalence classes of measurable functions $\ell : (0, 1) \to \mathbb{R}$ such that

$$((\ell))_p = \int_0^1 |\ell(x)|^p dx < \infty$$

It will be shown that $d(f, g) = ((f - g))_p$ is a metric on $L^p(0, 1)$ and that with this metric $L^p(0, 1)$ is a Fréchet space. It will also be shown, however, that $L^p(0, 1)$ has only one nonempty open convex set, namely itself. So $L^p(0, 1)$ $0 < p < 1$, is most emphatically not locally convex.

Theorem 3.27. *X is a LCS, M is a closed linear subspace of X, and $x_0 \notin M$, then there is a nonzero continuous linear functional $\ell \in X^*$ such that*

(a) $\ell(x_0) = 1$.

(b) $\ell(y) = 0$ *for all* y *in* M.

Proof. By Theorem 3.25, there is nonzero $\ell \in X^*$ such that $\ell(x_0)$ and $\ell(M)$ are disjoint. Thus $\ell(M)$ is a *proper subspace* of the scalar field. This forces $\ell(M) = \{0\}$ and $\ell(x_0) \neq 0$. The desired functional is obtained by dividing ℓ by $\ell(x_0)$. $\square$

There is an another useful corollary of the separation theorem.

Theorem 3.28. *X is LCS, B is a balanced, closed, convex subset, For any $x_0 \notin B$, there exists $\ell \in X^*$ such that*

$$|\ell x| \leqslant 1 < \ell x_0\,, \quad \textit{for all } x \in B\,.$$

Proof. By Theorem 3.25, there is nonzero $\ell_1 \in X^*$ such that

$$\operatorname{Re} \ell_1 x < c < \operatorname{Re} \ell_1 x_0\,, \quad \text{for all } x \in B\,.$$

Since B is balanced, so is $\ell_1(B)$. Hence

$$|\ell_1 x| < c < |\ell_1 x_0|\,, \quad \text{for all } x \in B\,.$$

Let $\ell_1 x_0 = |\ell_1 x_0| e^{i\theta}$, then $\ell = c^{-1} e^{-i\theta} \ell_1$ has the desired properties. $\square$

Geometric interpretations

The geometric consequences of the Hahn-Banach Theorem are achieved by interpreting that theorem in light of the correspondence between linear functionals and hyperplanes and between sublinear functionals and open convex neighborhoods of the origin.

As bofore, there is a great advantage inherent in a geometric discussion of real TVS's. Namely, if $\ell : X \to \mathbb{R}$ is a nonzero continuous $\mathbb{R}$-linear functional, then the closed hyperplane $\ker\ell$ disconnects the space. Indeed, $X \backslash \ker\ell$ has two connected components: $\{x : \ell(x) > 0\}$ and $\{x : \ell(x) < 0\}$. But If X is a complex TVS and $\ell : X \to \mathbb{C}$ is a nonzero continuous linear function, then $X \backslash \ker\ell$ is connected.

Theorem 3.24 and Theorem 3.25 can be rewrited as

(a) X is a real TVS and A and B are disjoint convex sets with A open, then A and B are separated.

(b) X is a real LCS and A and B are two disjoint closed convex subsets. If A is compact, then A and B are strictly separated.

Proposition 3.29. *If X is a real LCS, A is a subset of X.*

(a) $\overline{\mathrm{co}}(A)$ *is the intersection of the closed half-spaces containing A.*

(b) $\overline{\mathrm{span}}(A)$ *is the intersection of the closed hyperplanes containing A.*

Extension theorem

Theorem 3.30. *Let X be a LCS. ℓ is a continuous linear functional on a subspace Y, then there ℓ can be extended on X as a continuous linear functional.*

Proof. Assume, without loss of generality, that ℓ is not identically 0. Put

$$Y_{\mathrm{o}} = \{y \in Y : \ell(y) = 0\}$$

and pick $y_{\mathrm{o}} \in Y$ such that $\ell(y_{\mathrm{o}}) = 1$. since ℓ is continuous on Y

$$y_{\mathrm{o}} \notin Y_{\mathrm{o}} = \mathrm{cl}_Y(Y_{\mathrm{o}}) = \mathrm{cl}_X(Y_{\mathrm{o}}) \cap Y\,,$$

so $y_{\mathrm{o}} \notin \mathrm{cl}_X(Y_{\mathrm{o}})$. Then there exists a $\Lambda \in X^*$ such that

(a) $\Lambda y_{\mathrm{o}} = 1$

(b) $\Lambda = 0$ on Y_{o}.

Then for any $y \in Y$, note that $y - \ell(y)y_{\mathrm{o}} \in Y_{\mathrm{o}}$, since $\ell(y_{\mathrm{o}}) = 1$. Hence

$$\Lambda y - \ell(y) = \Lambda y - \ell(y)\Lambda y_{\mathrm{o}} = \Lambda\left(y - \ell(y)y_{\mathrm{o}}\right) = 0\,.$$

Thus $\Lambda = \ell$ on Y. $\square$

3.7 Weak topologies on LCS

As before, for a LCS X, let X^* denote the space of continuous linear functionals on X. Obviously, X^* has a natural vector-space structure.

It is convenient and, more importantly, helpful to introduce the notation,because of a certain symmetry,

$$\langle x, x^* \rangle$$

to stand for $x^*(x)$, for x in X and x^* in X^*.

Definition 3.4. (X, τ) is a LCS, the **weak topology** on X, denoted by τ_w or $\sigma(X, X^*)$, is the topology defined by the separating family of seminorms $\{p_{x^*} : x^* \in X^*\}$, where

$$p_{x^*}(x) = |\langle x, x^* \rangle| \text{ , for all } x \in X \text{ .} \tag{3.5}$$

On the other hand, The **weak*** **topology** on X^*, denoted by τ_{w^*} or $\sigma(X^*, X)$, is the topology defined by the separating family of seminorms seminorms $\{p_x : x \in X\}$, where

$$p_x(x^*) = |\langle x, x^* \rangle| \text{ , for all } x^* \in X^* \text{ .} \tag{3.6}$$

Remark 3.4. The weak topology on X is the weakest one with respect to which X^* is continuous. Regarding each $x \in X$ as a linear functional on X^*, denoted as Jx and, given by

$$Jx : x^* \mapsto \langle x, x^* \rangle \text{ , for all } x^* \in X^* \text{ ,} \tag{3.7}$$

the weak* topology on X^* is the weakest one with respect to which $J(X)$ is continuous.

Obviously, (X, τ_w) and (X^*, τ_{w^*}) both are LCS, by Example 3.1, and

(a) (X, τ_w) has a local base, namely,

$$\left\{ V(\Phi, \epsilon) := \{x : |\langle x, x^* \rangle| < \epsilon, x^* \in \Phi\} : \Phi \subset X^* \text{ is finite, } \epsilon > 0 \right\}, \tag{3.8}$$

and (X^*, τ_{w^*}) has a local base, namely,

$$\left\{ V(\Phi, \epsilon) := \{x^* : |\langle x, x^* \rangle| < \epsilon, x \in \Phi\} : \Phi \subset X \text{ is finite, } \epsilon > 0 \right\}. \tag{3.9}$$

(b) A net $\{x_i\}$ in X is convergent to some $x \in X$ in τ_w if and only if

$$\langle x_i, x^* \rangle \to \langle x, x^* \rangle \ , \text{ for all } x^* \in X^* \,. \tag{3.10}$$

A net $\{x_i^*\}$ in X^* is convergent to some $x^* \in X$ in τ_{w^*} if and only if

$$\langle x, x_i^* \rangle \to \langle x, x^* \rangle \ , \text{ for all } x \in X \tag{3.11}$$

(c) $M \subset X$ is bounded if and only if

$$\sup_{x \in M} |\langle x, x^* \rangle| < \infty \,, \text{ for all } x^* \in X^* \,. \tag{3.12}$$

$N \subset X^*$ is bounded if and only if

$$\sup_{x^* \in N} |\langle x, x^* \rangle| < \infty \,, \text{ for all } x \in X \,. \tag{3.13}$$

Duality

Theorem 3.31. *(X, τ) is LCS, then $(X, \tau_w)^* = X^*$,*

Proof. $X^* \subset (X, \tau_w)^*$, since the weak topology on X is the weakest one with respect to which X^* is continuous.

On the other hand, $\tau_w \subset \tau$, thus $(X, \tau_w)^* \subset X^*$. □

Theorem 3.32. *(X, τ) is LCS, then $(X^*, \tau_{w^*})^* = J(X)$*

Proof. $J(X) \subset (X^*, \tau_{w^*})^*$, since the weak* topology on X^* is the weakest one with respect to which $J(X)$ is continuous.

On the other hand, for any $f \in (X^*, \tau_{w^*})^*$, there exists $x_1, \dots, x_n$ in X and positive scalars $\alpha_1, \dots, \alpha_n$, by Theorem 3.15, such that

$$|f(x^*)| \leqslant \sum_{k=1}^{n} \alpha_k |\langle x_k, x^* \rangle|, \quad \text{for all } x^* \in X^*.$$

Thus

$$\ker f \subset \bigcap_{k=1}^{n} \ker(Jx_k).$$

By Lemma 0.2, $f \in \mathrm{span} Jx_1, \cdots, Jx_n$, of course $f \in J(X)$. □

Remark 3.5. Since $J(X) = (X^*, \tau_{w^*})^*$, we can define a weak* topology on X by regarding $J(X)$ as X (Indeed, if $\mathcal{T}$ is the weak* topology on $J(X)$, then $J^{-1}(\mathcal{T})$ is the topology on X induced by J), this topology is exactly τ_w.

Weak closure

Theorem 3.33. *(X, τ) is LCS, K is a convex subset, then the weak closure of K is equal to its original closure, i.e.,*

$$\overline{K} = \overline{K}^{\tau_w}. \tag{3.14}$$

Proof. $\overline{K}^{\tau_w}$ is weakly closed, hence originally closed, so that

$$\overline{K} \subset \overline{K}^{\tau_w}.$$

To obtain the opposite inclusion, if there exists $x_0 \in X$, and

$$x_0 \in \overline{K}^{\tau_w} \backslash \overline{K}$$

By H-B separation theorem, there exist $x^* \in X^*$ and $\gamma \in \mathbb{R}$ such that, for every $x \in K$,

$$\mathrm{Re}\langle x_0, x^* \rangle < \gamma < \gamma + \epsilon < \mathrm{Re}\langle x, x^* \rangle.$$

Thus $|\langle x - x_0, x^* \rangle| \geqslant \epsilon$ for all $k \in K$, i.e.,

$$K \cap V(x_0, x^*, \epsilon) = \varnothing \, .$$

Thus $x_0 \notin \overline{K}^{\tau_w}$, this is a contradiction. $\square$

Corollary 3.34. *(X, τ) is LCS, K is convex. Then K is closed if and only if K is weakly closed.*

Theorem 3.35. *(X, τ) is a metrizable locally convex space. If $\{x_n\}$ is a sequence in X that converges weakly to some $x \in X$, then there is a sequence $\{y_m\}$ in X such that*

(a) each y_m is a convex combination of finitely many x_n, and

(b) $y_m \to x$ originally.

Proof. Let H be the convex hull of the set of all x_n, let K be the weak closure of H. Then $x \in K$. By Theorem 3.33, x is also in the original closure of H. Since the original topology of X is assumed to be metrizable, it follows that there is a sequence $\{y_m\}$ in H that converges originally to x. $\square$

To get a feeling for what is involved here, consider the following example.

Example 3.3. Let K be a compact Hausdorff space,for example, the unit interval on the real line, assume that f and f_n are continuous complex functions on K such that

(a) $f_n(x) \to f(x)$ for every $x \in K$.

(b) $|f_n(x)| \leqslant 1$ for all n and x.

Theorem 3.35 asserts that there are convex combinations of the f_n that converge uniformly to f.

To see this, let $C(K)$ be the Banach space of all complex continuous functions on K, normed by the supremum. Then norm conversence is the same as uniform convergence on K. If μ is any complex Borel measure on K, Lebesgue's dominated convergence theorem implies that

$$\int f_n d\mu \to \int f d\mu.$$

Hence $f_n \to f$ weakly, by the Riesz representation theorem which identifies $C(K)^*$ with the space of all regular complex Borel measures on K.

Annihilators

Annihilator is in an analogue of the orthogonal complement in Hilbert space.

Definition 3.5. (X, τ) is LCS, X^* is the dual of X.

- M is a subspace of X, the **annihilator of** M is defined by

$$M^\perp = \{x^* \in X^* : \langle x, x^* \rangle = 0 \text{ for all } x \in M\} . \tag{3.15}$$

- N is a subspace of X^*, the **annihilator of** N is defined by

$${}^\perp N = \{x \in X : \langle x, x^* \rangle = 0 \text{ for all } x^* \in N\} . \tag{3.16}$$

Thus $M^\perp$ consists of all continuous linear functionals on X vanishing on M, and ${}^\perp N$ is the subset of X on which every member of N vanishes.

It is clear that $M^\perp$ is weak-closed subspace of X, and ${}^\perp N$ is weak*- closed subspace of X^*. The following theorem describes the duality between these two types of annihilators.

Theorem 3.36. *Under the preceding hypotheses,*

$${}^\perp\left(M^\perp\right) = \overline{M}^{\tau_w} = \overline{M}\,, \; \left({}^\perp N\right)^\perp = \overline{N}^{\tau_{w^*}} . \tag{3.17}$$

Proof. Firstly, ${}^{\perp}\left(M^{\perp}\right)$ is weak-closed, so

$$\overline{M}^{\tau_w} \subset {}^{\perp}\left(M^{\perp}\right).$$

If $x \notin \overline{M}^{\tau_w}$, by Theorem 3.27, there exists nonzero continuous linear functional $x^* \in X^*$ vanishing on M and $\langle x, x^* \rangle = 1$. Thus

$$x^* \in M^{\perp} \text{ and, } x \notin {}^{\perp}(M^{\perp}).$$

So the equality is established.

Similarly, $\left({}^{\perp}N\right)^{\perp}$ is weak*-closed, so

$$\overline{N}^{\tau_{w^*}} \subset \left({}^{\perp}N\right)^{\perp}.$$

If $x^* \notin \overline{N}^{\tau_{w^*}}$, by Theorem 3.27 implies the existence of an $x \in {}^{\perp}N$ such that $\langle x, x^* \rangle = 1$. Thus $x^* \left({}^{+}N\right)^2$, Thus

$$x \in^{\perp} N \text{ and, } x^* \notin \left({}^{\perp}N\right)^{\perp}.$$

Now we get the desired equality . $\square$

3.8 Appendix 2: Metrization of TVS

We recall that a topology τ on a set X is said to be metriable if there is a metric d on X which is compatible with τ. In that case, X is first-countable space. This gives a necessary conditionfor metrizability which, for topological vector spaces, turns out to be also sufitieient.

Theorem 3.37. *If X is a topological vector space with a countable local base, then there is a metric d on X such that*

(a) *d is compatible with the topology of X*

(b) *the open balls centered at 0 are balanced, and*

(c) *d is invariant: $d(x+z, y+z) = d(x, y)$ for $x, y, z \in X$.*

If, in addition, X is locally convex, then d can be chosen so as to satisfy that all open balls are convex.

Proof. X has a countable balanced local base $\{V_n\}$ such that for $n \geq 1$

$$V_{n+1} + V_{n+1} + V_{n+1} + V_{n+1} \subset V_n \,.$$

when X is locally convex, this local base can be chosen so that each V_n is also convex.

Let D be the set of all rational numbers r of the form

$$r = \sum_{n=1}^{\infty} c_n(r) 2^{-n}$$

where each of the "digits" $c(r)$ is 0 or 1 and only finitely many are 1. Thus each $r \in D$ satisfies the inequalities $0 \leq r < 1$.

Put $A(r) = X$ if $r \geq 1$; for any $r \in D$, define

$$A(r) = c_1(r)V_1 + c_2(r)V_2 + c_3(r)V_3 + \cdots$$

Note that each of these sums is actually finite. Define

$$\ell(x) = \inf\{r : x \in A(r)\} \,, \; x \in X$$

and

$$d(x, y) = \ell(x - y)\,,\ x, y \in X$$

The proof that this d has the desired properties depends on the inclusions

$$A(r) + A(s) \subset A(r + s)\,,\ r \in D, s \in D \tag{3.18}$$

Thus $\{A(r)\}$ is totally ordered by set inclusion. We claim that

$$\ell(x + y) \leq \ell(x) + \ell(y)\,,\ x, y \in X$$

$\square$

Chapter 4

Opertators and Functionals on Banach Space

4.1 Convergence of operaters

Let X and Y be normed linear spaces over $\mathbb{F}$.

Definition 4.1. $\{A_n\}$ in $\mathcal{B}(X,Y)$ is said to be **uniformly convergent** to $A \in \mathcal{B}(X,Y)$ if

$$\|A_n - A\| \to 0\,. \tag{4.1}$$

In this case, A is called the **uniform limit** of the sequence $\{A_n\}$.

Definition 4.2. $\{A_n\}$ in $\mathcal{B}(X,Y)$ is said to be **strongly convergent** to $A \in \mathcal{B}(X,Y)$ if

$$A_n x \to Ax \text{ for each } x \in X\,. \tag{4.2}$$

In this case, A is called the **strong limit** of the sequence $\{A_n\}$, denoted by $A = s\text{-}\lim A_n$ or $A_n \xrightarrow{s} A$.

Definition 4.3. $\{A_n\}$ in $\mathcal{B}(X,Y)$ is said to be **weakly convergent** to $A \in \mathcal{B}(X,Y)$ if

$$\langle A_n x, y^* \rangle \to \langle Ax, y^* \rangle\,, \text{ for each } x \in X \text{ and } y^* \in Y^*\,. \tag{4.3}$$

In this case A is called the **weak limit** of the sequence $\{A_n\}$, denoted by $A = w\text{-}\lim A_n$ or $A_n \xrightarrow{w} A$.

It is obvious that uniform convergence implies strong convergence, and strong convergence implies weak convergence, but the converse does not hold.

Example 4.1. Consider the sequence $\{T_n\}$ on ℓ^2, where for each n, $T_n : \ell^2 \to \ell^2$ is given by

$$T_n(x_1, x_2, \cdots) = (0, 0, \cdots, 0, x_{n+1}, x_{n+2}, \cdots)$$

Then $T_n \to 0$ strongly, but $\|T_n\| = 1$ for all n, so $\{T_n\}$ does not converge to 0 in the uniform topology.

Example 4.2. Consider the sequence $\{S_n\}$ on ℓ^2, where $S : \ell^2 \to \ell^2$ is given by

$$S(x_1, x_2, \cdots) = (0, x_1, x_2, \cdots)$$

and $S_n = S^n$ for $n \in \mathbb{N}$.

Then $S_n \to 0$ weakly, but $\|S_n x\| = \|x\|$ for all x, thus $\{S_n x\}$ does not converge to 0, i.e., $\{S_n\}$ doesn't converge strongly to 0.

Topologies*

In fact, we can induce some topologies on $\mathcal{B}(X, Y)$ to discribe this serval convergence.

- the operator norm $\|\cdot\|$ deteimines a natural norm topology τ (or unique topology) on $\mathcal{B}(X, Y)$.
- For any $x \in X$, define p_x by

$$p_x(A) = \|Ax\|\,, \text{ for all } A \in \mathcal{B}(X, Y)\,. \tag{4.4}$$

Then $\{p_x\}_{x\in X}$ is a separating family of seminorms on $\mathcal{B}(X,Y)$, inducing a locally convex topology, namely strong topology,denoted by τ_s.

- For any $x \in X$ and $y^* \in Y^*$, define p_{x,y^*} by

$$p_{x,y^*}(A) = \|\langle Ax, y^*\rangle\|\,,\ \text{for all } A \in \mathcal{B}(X,Y)\,. \tag{4.5}$$

  Then $\{p_{x,y^*}\}_{x\in X, y^*\in Y^*}$ is a separating family of seminorms on $\mathcal{B}(X,Y)$, inducing a locally convex topology, namely strong topology, denoted by τ_w.

clearly, $\tau_w \subset \tau_s \subset \tau$. Moreover, we have

(a) $A_n \xrightarrow{s} A$ iff $A_n \to A$ in τ_s and $A_n \xrightarrow{w} A$ iff $A_n \to A$ in τ_w.

(b) τ_s is the weakest topology with respect to which $A \mapsto Ax$ is continuous on $\mathcal{B}(X,Y)$ for all $x \in X$, τ_w is the weakest topology with respect to which $A \mapsto \langle Ax, y^*\rangle$ is continuous on $\mathcal{B}(X,Y)$ for all $x \in X$ and $y^* \in Y^*$.

(c) When $Y = \mathbb{F}$, the scalars field, $\mathcal{B}(X,Y)$ is exactly X^*, then it's easy to find the τ_s, τ_w coinsides with $\sigma(X^*, X)$.

(d) The relative topology on X (不区分 JX) induced by the $\sigma(X^{**}, X^*)$, is $\sigma(X, X^*)$. Thus the weak topology on X can be regarded as a relative topology induced by τ_s or τ_w.

4.2 Applications of Baire's Category

Principle of uniform boundedness

A subset $\mathcal{A}$ of $\mathcal{B}(X, Y)$ is said to be **pointwise bounded** on X, if for each $x \in X$,

$$\sup_{A \in \mathcal{A}} \|Ax\| < \infty\,;$$

be **uniformly** (or **norm**) **bounded** if

$$\sup_{A \in \mathcal{A}} \|A\| < \infty\,.$$

Remark 4.1. In fact, $\mathcal{A}$ is pointwise bounded iff $\mathcal{A}$ is strongly bounded, that is $\mathcal{A}$ is bounded with respect to the strong topology τ_s.

Obviously, $\mathcal{A}$ is uniformly bounded implies it be pointwise bounded. The next theorem asserts that When X is Banach space, the converse is ture.

Theorem 4.1 (Principle of Uniform Boundedness)**.** *Let X be a Banach space and let Y a norned space. If $\mathcal{A} \subseteq \mathcal{B}(X, Y)$ is pointwise bounded, then $\mathcal{A}$ is uniformly bounded.*

Proof. For each $k \in \mathbb{N}$, let

$$E_k = \{x \in X : \|Ax\| \leq k \text{ for all } A \in \mathcal{A}\}\ .$$

since A is continuous, E_k is closed. Note that $X = \cup_{k=1}^{\infty} E_k$. By Baire's category theorem, there is an index m such that $\mathrm{int}(E_m) \neq \varnothing$.

That is, there is an $x_0 \in E_m$ and an $\epsilon > 0$ such that $B\left(x_0, 2\epsilon\right) \subset E_m$. Then for any $\|x\| \leqslant 1$, $x_0 + \epsilon x \in B\left(x_0, 2\epsilon\right)$, so

$$\|A(x_0 + \epsilon x)\| \leqslant m\,, \quad \text{for all } A \in \mathcal{A}\,.$$

Thus

$$\|Ax\| = \frac{\|Ax_0\| + m\|}{\epsilon} \leqslant \frac{2m}{\epsilon}\,, \quad \text{for all } A \in \mathcal{A}\,.$$

So $\mathcal{A}$ is uniformly bounded. $\square$

Corollary 4.2. *X is a Banach space and $N \subseteq X^*$, then N is norm bounded iff for every x in X,*

$$\sup_{x^* \in N} |\langle x, x^* \rangle| < \infty$$

In other words, N is weak-bounded implies N is norm bounded.*

The next corollary doesn't need X is complete, since we use PUB for the Banach space X^*.

Corollary 4.3. *X is a normed space and $M \subset X$, then M is bounded iff for every x^* in X^*,*

$$\sup_{x \in M} |\langle x, x^* \rangle| < \infty$$

In other words, M is weakly bounded implies N is norm bounded.

Using PUB twice, we get

Corollary 4.4. *X is a Banach space and Y is a normed space, if $\mathcal{A} \subset \mathcal{B}(X, Y)$ such that for every x in X and y^* in Y^**

$$\sup_{A \in \mathcal{A}} |\langle Ax, y^* \rangle| < \infty$$

then $\mathcal{A}$ is uniformly bounded.

Remark 4.2. This means if X is Banach space, $\mathcal{A}$ is weakly bounded $\Leftrightarrow$ $\mathcal{A}$ is strongly bounded $\Leftrightarrow$ $\mathcal{A}$ is uniformly bounded .

Proposition 4.5. *X is a Banach space and Y is a normed space, $\{A_n\}$ be sequence in $\mathcal{B}(X, Y)$. If $A_n \to A$ strongly, then $\{A_n\}$ is uniformly bounded and,*

$$\|A\| \leqslant \liminf_{n \to \infty} \|A_n\|\,. \tag{4.6}$$

Proof. By PUB, we have that $\{A_n\}$ is uniformly bounded, and note that

$$|Ax| = \lim_{n\to\infty} |A_n x| \leqslant \liminf_{n\to\infty} \|A_n\| \|x\|$$

for any $x \in X$, Thus

$$\|A\| \leqslant \liminf_{n\to\infty} \|A_n\| \,. \qquad \square$$

A special form of the PUB that is quite useful is the following:

Proposition 4.6. *X is a Banach space and Y is a normed space, $\{A_n\}$ be sequence in $\mathcal{B}(X,Y)$. $\{A_n\}$ is strongly convergent if and only if*

(a) $\{A_n\}$ *is uniformly bounded and,*

(b) $\{A_n x\}$ *converges for all* $x \in X$.

Proof. Necessity is obvious. To show sufficiency, we define $A : X \to Y$ by

$$Ax := \lim_{n\to\infty} A_n x \,.$$

A is linear, clearly. A is continuous by the proof of Proposition 4.5. $\square$

Corollary 4.7 (Banach-Steinhaus)**.** *X, Y are a Banach space, $\{A_n\}$ is a sequence in $\mathcal{B}(X,Y)$. Then $\{A_n\}$ is strongly convergent if and only if*

(a) $\{A_n\}$ *is uniformly bounded and,*

(b) there exists D dense in X, so that $\{A_n x\}$ converges for all $x \in X$.

Proof. It sufficies to show that $\{A_n x\}$ converges for all $x \in X$, and this is why we need Y is complete. For any $x \in X$, and given $\epsilon > 0$, there exists $x' \in D$, depending on ϵ, so that $\|x - x'\| \leqslant \epsilon$. Then

$$\begin{aligned}\|A_{n+p}x - A_n x\| &\leqslant \|A_{n+p}x - A_{n+p}x'\| + \|A_{n+p}x' - A_n x'\| + \|A_n x' - A_n x\| \\ &\leqslant 2\sup_n \|A_n\|\epsilon + \|A_{n+p}x' - A_n x'\|\end{aligned}$$

Since $\{A_n x'\}$ is a Cauchy sequence, it's easy to find that $\{A_n x\}$ is a Cauchy sequence. $\square$

The open mapping theorem

Definition 4.4. Let X and Y be normed linear spaces. $A : X \to Y$. is called an **open mapping** if $A(U)$ is open in Y whenever U is open in X.

Theorem 4.8 (The Open Mapping Theorem)**.** *Let X and Y be Banach spaces and suppose that $A \in \mathcal{B}(X, Y)$. If A is surjective, then A is an open mapping.*

Proof. **Step 1.** Note that for any $x \in X$, $AB_X(x, r) = Ax + rB_X(0, 1)$, We only need to prove that $AB_X(0, 1)$ is open. To see this, it must be shown that there is constant $r > 0$ such that $B_Y(0, r) \subset AB_X(0, 1)$. But this is also sufficient, in deed, for any $y = Ax \in AB_X(0, 1)$, take $\epsilon > 0$ such that $\|x\| + \epsilon < 1$, then

$$B_Y(y, \epsilon) = Ax + B_Y(0, \epsilon) \subset AB_X(0, 1)\,.$$

Thus $AB_X(0, 1)$ is open.

Step 2. We shall prove there is a constant $r > 0$ such that $B_Y(0, r) \subset \overline{AB_X(0, 1)}$.

It is easy to see that $X = \bigcup_{n=1}^{\infty} nB_X(0, 1)$. Since T is surjective,

$$Y = TX = T\left(\bigcup_{n=1}^{\infty} nB_X(0, 1)\right) = \bigcup_{n=1}^{\infty} nTB_X(0, 1) = \bigcup_{n=1}^{\infty} n\overline{TB_X(0, 1)}\,.$$

By Baire's category theorem, there is $m \in \mathbb{N}_+$ such that $\left(m\overline{TB_X(0, 1)}\right)^\circ \neq \varnothing$. This implies that $\left(\overline{TB_X(0, 1)}\right)^\circ \neq \varnothing$. Hence, there is a constant $r > 0$ and an element $y_0 \in Y$ such that $B_Y(y_0, 2r) \subset \overline{TB_X(0, 1)}$. Since $y_0 \in \overline{TB_X(0, 1)}$, it follows, by symmetry, that $-y_0 \in \overline{TB_X(0, 1)}$. Therefore

$$B_Y(0, 2r) = B_Y(y_0, 2r) - y_0 \subset \overline{TB_X(0, 1)} + \overline{TB_X(0, 1)}\,.$$

Since $\overline{TB_x(0, 1)}$ is convex, $\overline{TB_X(0, 1)} + \overline{TB_X(0, 1)} = 2\overline{TB_X(0, 1)}$. Hence, $B_Y(0, 2r) \subset 2\overline{TB_X(0, 1)}$ and, consequently, $B_Y(0, r) \subset \overline{TB_X(0, 1)}$.

Step 3. We shall prove that $B_Y(0,r/2) \subset AB_X(0,1)$.
Take any $y \in B_Y(0,r/2)$. From $B_Y(0,r) \subset \overline{AB_X(0,1)}$, we have

$$B_Y(0,r/2^n) \subset \overline{AB_X(0,1/2^n)}\,,\ n \in \mathbb{N}_+.$$

So there is $x_1 \in B_X(0,1/2)$ such that $y - Ax_1 \in B_Y(0,r/2^2)$. By induction there is a suquence $\{x_n\}$ such that $x_n \in B_X(0,1/2^n)$ and

$$y - A(x_1 + \cdots + x_n) \in B_X(0,1/2^n)\,.$$

Note that $\sum_{n=1}^{\infty} \|x_n\| < 1$, since X is Banach space, there is $z \in B_X(0,1)$ such that $z = \sum_{n=1}^{\infty} x_n$, and $Az = y$. Thus $B_Y(0,r/2) \subset AB_X(0,1)$. □

Remark 4.3. Conversely, if A is an open mapping, then A must be surjective.

From the open mapping theorem, we get the following theorem directly.

Theorem 4.9 (Inverse mapping theorem)**.** *Let X and Y be Banach spaces and assume $A \in \mathcal{B}(X,Y)$ is bijective. Then $A^{-1} \in \mathcal{B}(Y,X)$.*

The closed graph theorem

Definition 4.5. Let X and Y be linear spaces and $A : X \to Y$. The graph of A, denoted by $\mathcal{G}(A)$, is the subset of $X \oplus_1 Y$ given by

$$\mathcal{G}(A) = \{(x, Ax) : x \in X\}\,.$$

Obviously, $\mathcal{G}(A)$ is a linear subspace of $X \oplus_1 Y$.

Theorem 4.10 (The Closed Graph Theorem)**.** *Let X and Y be Banach spaces and $A \in \mathcal{L}(X,Y)$ has a closed graph. Then $A \in \mathcal{B}(X,Y)$.*

Proof. Since $X \oplus_1 Y$ is a Banach space and $\mathcal{G}(A)$ is closed, $\mathcal{G}(A)$ is a Banach space. Define $P : \mathcal{G}(A) \to X$ by $P(x \oplus Ax) = x$. It is easy to check that P is bounded and bijective. By the inverse mapping theorem, $P^{-1} : X \to \mathcal{G}(A)$ is continuous. Thus $A : X \to Y$ is the composition of the continuous map $P^{-1} : X \to \mathcal{G}(A)$ and the continuous map of $\mathcal{G}(A) \to Y$ defined by $x \oplus Ax \mapsto Ax$. A is therefore continuous. $\square$

EXERCISE

Exercise 4.1. Let $A \in \mathcal{B}(X, Y)$. Then $A^{-1} : \operatorname{ran}(A) \to X$ exists and in $\mathcal{B}(\operatorname{ran} A, X)$ if and noly if there is a constant $m > 0$ such that for each $x \in X, \|Ax\| \geqslant m\|x\|$.

4.3 Linear functionals

We have pointed that, a subspace in X is a hyperplane if and only if it is the kernel of a non-zero linear functional. Two linear functionals have the same kernel if and only if one is a non-zero multiple of the other.

Hyperplanes in a normed space fall into one of two categories.

Proposition 4.11. *If X is a normed space and M is a hyperplane in X, then either M is closed or M is dense.*

Proof. Note that $\dim(X/\overline{M}) \leqslant \dim(X/M)$. $\square$

Example 4.3. To get an example of a dense hyperplane, consider normed linear space c_0. Denote e_n as the element of c_0 such that $e_n(k) = 0$ if $k \neq n$ and $e_n(n) = 1$. Let $x_0(n) = 1/n$ for all n, so $x_0 \in c_0$ and $\{x_0, e_1, e_2, \ldots\}$ is a linearly independent set in c_0. Let $\mathscr{B} =$ a Hamel basis in c_0 which contains $\{x_0, e_1, e_2, \ldots\}$. Put $\mathscr{B} = \{x_0, e_1, e_2, \ldots\} \cup \{b_i : i \in I\}$. Define $f : c_0 \to \mathbb{F}$ by

$$f\left(\alpha_0 x_0 + \sum_n \alpha_n e_n + \sum_i \beta_i b_i\right) = \alpha_0.$$

(Remember that in the preceding expression at most a finite number of the α_n and β_i are not zero). Since $e_n \in \ker f$ for all $n \geqslant 1$, ker f is dense but clearly ker $f \neq c_0$.

Theorem 4.12. *If X is a normed space and $f : X \to \mathbb{F}$ is a linear functional, then f is continuous if and only if $\ker f$ is closed.*

Proof. If f is continuous, $\ker f = f^{-1}(\{0\})$ and so $\ker f$ must be closed. Assume now that $\ker f$ is closed and let $Q : X \to X/\ker f$ be the natural map. Let $T : X/\ker f \to \mathbb{F}$ be an isomorphism, both Q, T are continuous.(Why?) Thus, if $g = T \circ Q : X \to \mathbb{F}, g$ is continuous and $\ker f = \ker g$. Hence $f = \alpha g$ for some α in $\mathbb{F}$ and so f is continuous. $\square$

Second proof. If there exises $\{x_n\}$ and $\epsilon > 0$, so that $\|x_n\| \to 0$ but $|f(x_n)| \geqslant \epsilon$ for all n. Consider

$$\frac{x_n}{f(x_n)} - \frac{x_1}{f(x_1)} \in \ker f \, .$$

Since $\ker f$ is closed and $\frac{x_n}{f(x_n)} \to 0$, we deduce that $\frac{x_1}{f(x_1)} \in \ker f$, which is absurd. □

Hahn-Banach Theorem in Normed linear space

Theorem 4.13 (Hahn-Banach)**.** *If X is a normed space, Y is a subspace in X, and $f : Y \to \mathbb{F}$ is a bounded linear functional, then there is an F in X^* such that $F|_Y = f$ and $\|F\| = \|f\|$.*

Proof. Using Theorem 1.17 with $p(\cdot) = \|f\|\|\cdot\|$. □

Note that Y don't need to be closed. In fact, without using Hahn-Banach theorem we can prove there is a bounded linear functional F on $\overline{Y}$ such that $F|_Y = f$ and $\|F\| = \|f\|$.

Corollary 4.14. *If X is a normed space, $\{x_1, x_2, \ldots, x_a\}$ is a linearly independent subset of X, and $\alpha_1, \alpha_2, \ldots, \alpha_d$ are arbitrary scalars in $\mathbb{F}$, then there is an f in X^* such that $f(x_j) = \alpha_j$ for $1 \leqslant j \leqslant d$.*

Proof. Let Y = the linear span of $x_1, \ldots, x_d$ and define $g : Y \to \mathbb{F}$ by $g\left(\sum_j \beta_j x_j\right) = \sum_j \beta_j \alpha_j$. So g is linear. since Y is finite dimensional, g is continuous. Let f be a continuous extension of g to X. □

Corollary 4.15. *X is normed space. For any $x \in X$,*

$$\|x\| = \sup_{f \in X^*, \|f\| \leqslant 1} |\langle f, x \rangle| \, . \tag{4.7}$$

Moreover, this supremum is attained.

Proof. Obviously,

$$\sup_{f \in X^*, \|f\| \leqslant 1} |\langle f, x\rangle| \leqslant \|x\| \,.$$

On the other hand, define $g : \mathrm{span}\{x\} \to \mathbb{F}$ by

$$g(\beta x) = \beta \|x\|, \text{ for all } \beta \in \mathbb{F}\,.$$

Then $g$is bounded and $\|g\| = 1$. By H-B extension theorem, there is f in X^* such that $\|f\| = 1$ and $f(x) = g(x) = \|x\|$. ☐

Corollary 4.16. *X, Y are normed space. $T \in \mathcal{B}(X, Y)$, then*

$$\|T\| = \sup_{\|x\| \leqslant 1, \|y^*\| \leqslant 1} |\langle Tx, y^*\rangle| \,. \tag{4.8}$$

Corollary 4.17. *If X is a normed space, M is a closed subspace of X, $x_0 \notin M$. Then there exists f in X^* such that*

(a) $\|f\| = 1$.

(b) $f(x_0) = \mathrm{dist}(x_0, M)$.

(c) $f(x) = 0$ *for all x in M.*

Proof. Let $Q : X \to X/M$ be the natural map. since $\|[x_0]\| = d$, by the preceding corollary there is a g in $(X/M)^*$ such that $g([x_0]) = d$ and $\|g\| = 1$. Let $f = g \circ Q : X \to \mathbb{F}$. Then f is continuous, $f(x) = 0$ for x in M, and $f(x_0) = 1$. Also, $|f(x)| = |g(Q(x))| \leqslant \|Q(x)\| \leqslant \|x\|$; hence $\|f\| \leqslant 1$. On the other hand, $\|g\| = 1$ so there is a sequence $\{x_n\}$ such that $|g([x_n])| \to 1$ and $\|[x_n]\| < 1$ for all n. Without loss of generality, we can assume $\|x_n\| < 1$. Then $|f(x_n)| = |g(x_n)| \to 1$, so $\|f\| = 1$. ☐

4.4 Dual Spaces

We shall denote by X^* the set of all countinuous linear functionals on X. We call X^* the **dual** of X. As a corollary of Proposition 1.3, , X^* is Banach space.

Remark 4.4. It should be emphasized that we did NOT assume X is complete. In fact, if $\widehat{X}$ is its completion, then X^* and $\widehat{X}^*$ are isometrically isomorphic.

Proposition 4.18. *X^* is isometricly isomorphic to a closed subspace of $C_b(B_X)$, where B_X is the closed unit ball in X.*

Proof. To see this, if $f \in X^*$, denote $f|_{B_X}$ as the restriction of f to B_X. Note that $\rho : X^* \to C_b(B_X)\,; f \mapsto f|_{B_X}$ is a linear isometry embdding. $\square$

Theorem 4.19. *X^* is separable, then X is also separable.*

Proof. Let $S = S_{X^*} = \{x^* : \|x^*\| = 1\}$ be the unit sphere in X^*. Then S is separable. Let $\{x_n^*\}$ be a countable dense subset of S.

Hence, for each $n \in \mathbb{N}$ there is an element $x_n \in X$ such that

$$\|x_n\| = 1 \text{ and } |\langle x_n, x_n^*\rangle| > \frac{1}{2}\,.$$

Then $\overline{\text{span}}\{x_n\} = X$, this is because $\ell \in X^*$ vanishes at $\{x_n\}$ implies $\ell = 0$ (Using Corollary 4.17). Thus X is separable. $\square$

Example 4.4. ℓ^1 is separable but $(\ell^1)^* = \ell^\infty$ is not. To see this, note that any $x \in (0,1)$ has a *binary representation*, denote by $\{b_n(x)\}$, or in other words

$$0.b_1(x)b_2(x)\cdots$$

where $b_n(x) \in \{0,1\}$. Then $\{b_n(x)\} \in \ell^\infty$ and for $n \neq m$,

$$\|b_n(x) - b_m(x)\|_\infty = 1\,.$$

Thus ℓ^∞ is not separable.

Represent of dual space

X and X' are metric spaces. A map $M : X \to X'$ is called an **isometry** if for any $x, y \in X$ one has $d(Mx, My) = d(x, y)$.

Particularly, if X, X' normed linear space, and the distance are induced by norm, then M is **isometry** iff $\|Mx\| = \|x\|$ for all $x \in X$.

Obviously, an isometry is automatically injective, and it can not be surjective.

Example 4.5. Definite $S : \ell^2 \to \ell^2$ by $S(\alpha_1, \alpha_2, \ldots) = (0, \alpha_1, \alpha_2, \ldots)$. Then S is an isometry that is not surjective.

Two linear spaces X and Y over the same field $\mathbb{F}$ are said to be **isomorphic** if there is a bijective linear operator $T \in \mathcal{L}(X, Y)$. If in addition, T is an isometry, then we say that T is an **isometry isomorphism**. In this case, X and Y are said to be isometrically isomorphic and we write $X \cong Y$.

Example 4.6. $c_0^* \cong \ell_1$.

To prove this, let $w = (w_n) \in \ell_1$ and define $\Phi : \ell_1 \to c_0^*$ by

$$\langle \Phi w, x \rangle = \sum_{n=1}^{\infty} x_n w_n\,, \quad \text{for any } x = (x_n) \in c_0$$

It's easy to show that Φw is a bounded linear functional on c_0 and

$$\|\Phi w\| = \|w\|_1.$$

To show that Φ is a surjective, consider a Schauder basis for c_0 is (e_n), where $e_n = (\delta_{nm})$ has 1 in the n-th position and zeroes elsewhere. Let $f \in c_0^*$ and $x = (x_n) \in c_0$. Then $x = \sum_{n=1}^{\infty} x_n e_n$ and therefore

$$f(x) = \sum_{n=1}^{\infty} x_n f(e_n)$$

Take any $k \in \mathbb{N}$, let

$$x_n^{(k)} = \begin{cases} |f(e_n)|/f(e_n), & n \leqslant k \text{ and } f(e_n) \neq 0. \\ 0, & \text{otherwise}. \end{cases}$$

Then $x^{(k)} = (x_n^{(k)}) \in c_0$ and $\|x^{(k)}\| = 1$. So $f(x^{(k)}) = \sum_{n=1}^{k} |f(e_n)| \leqslant \|f\|$. So we have $\sum_{n=1}^{\infty} |f(e_n)| < \infty$, $(f(e_n)) \in \ell_1$. Therefore, $\Phi(f(e_n)) = f$ and Φ is surjective.

Example 4.7. $L^p(\mu)^* \cong L^q(\mu)$.

Specifically, (X, Ω, μ) is a measure space and $p \in (1, \infty)$, $1/p + 1/q = 1$. For any $g \in L^q(X, \Omega, \mu)$, define $\ell_g : L^p(\mu) \to \mathbb{F}$ by

$$\ell_g(f) = \int fg d\mu$$

Then $\ell_g \in L^p(\mu)^*$ and the map $g \mapsto \ell_g$ defines an isometric isomorphism of $L^q(\mu)$ onto $L^p(\mu)^*$.

This theorem and the next have been proved in courses in measure and integration, we omit the proof.

Example 4.8. $L^1(\mu)^* \cong L^\infty(\mu)$, where μ must be σ-finite.

Specifically, (X, Ω, μ) is a σ-finite measure space. For any $g \in L^\infty(X, \Omega, \mu)$, define $\ell_g : L^1(\mu) \to \mathbb{F}$ by

$$\ell_g(f) = \int fg d\mu \ .$$

Then $\ell_g \in L^1(\mu)^*$ and the map $g \mapsto \ell_g$ defines an isometric isomorphism of $L^\infty(\mu)$ onto $L^1(\mu)^*$.

Example 4.9 (Riesz representation theorem)**.** $C_0(X)^* \cong \mathcal{M}(X)$, where X is a LCH[1].

[1]Locally Compact Hausdorff (topological space)

Specifically, X is a LCH, $M(X)$ denotes the space of all $\mathbb{F}$-valued regular Borel measures on X with the total variation norm. For any $\mu \in \mathcal{M}(X)$, define $F_\mu : C_0(X) \to \mathbb{F}$ by

$$\ell_\mu(f) = \int f d\mu \ .$$

Then $\ell_\mu \in C_0(X)^*$ and the map $\mu \to \ell_\mu$ is an isometric isomorphism of $\mathcal{M}(X)$ onto $C_0(X)^*$.

There are special cases of these theorems that deserve to be pointed out

- $l_p^* \cong l_q$, $l_1^* \cong l_\infty$.
- $c_0^* \cong l_1$. In fact, $c_0 = C_0(\mathbb{N})$, where $\mathbb{N}$ is given the discrete topology, and $l_1 = M(\mathbb{N})$.

The Dual of Subspace and Quotient Space

Let X be a normed space and $M \leqslant X$. If $f \in X^*$, then $f|_M$, the restriction of f to M, belongs to M^* and $\|f|_M| \leqslant \|f\|$. Acording to the Hahn-Banach Theorem, every bounded linear functional on M is obtainable as the restriction of a functional from X^*. In fact, more can be said.

Let M be a subset of a normed linear space X. The **annihilator** of M, denoted by ${}^\perp M$, is the set

$$M^\perp = \{x^* \in X^* : \langle x, x^*\rangle = 0 \text{ for all } x \in M\}$$

Let N be any subset of X^*, then the **annihilator** of N, denoted by $N^\perp$, is the set

$$N^\perp = \{x \in X : \langle x, x^*\rangle = 0 \text{ for all } x^* \in N\}$$

It is easy to observe that $M^\perp$ and ${}^\perp N$ are closed subspace of X^* and X.

Theorem 4.20. *M is a subspace of X, N is a subspace of X^*, then*

$$^\perp(M^\perp) = \overline{M}\,,\ \ ({}^\perp N)^\perp = \overline{N}^{\sigma(X^*,X)}\,. \tag{4.9}$$

Proof. Use H-B theorem. $\square$

Theorem 4.21. *M is a closed linear subspace of X. Then $M^* \cong X^*/M^\perp$. Moreover, the map $\rho : X^*/M^\perp \to M^*$ defined by*

$$\rho\left(\tilde{f}\right) = f|_M$$

is an isometric isomorphism.

Proof. $\square$

Theorem 4.22. *If $M \leqslant X$ and $Q : X \to X/M$ is the natural map, then $\rho(f) = f \circ Q$ defines an isometric isomorphism of $(X/M)^*$ onto $M^\perp$.*

Proof. $\square$

4.5 Bidual space and Reflexivity

Definition 4.6. The dual space of X^* is called the **second dual space** or **bidual space** of X, and denote as X^{**}.

Let $(X, \|\cdot\|)$ be a normed linear space over $\mathbb{F}$ and $x \in X$. Define a functional $Jx : X^* \to \mathbb{F}$ by

$$\langle x^*, Jx\rangle = \langle x, x^*\rangle \text{ for all } x^* \in X^* .$$

It is easy to verify that $Jx \in X^{**}$, and $|Jx| = \|x\|$.

It now follows that we can define a map

$$J : X \to X^{**},\ x \mapsto Jx. \tag{4.10}$$

Obviously, J_X is linear,then J_X is a linear isometry of X into its bidual X^{**}, which is called the **canonical** or **natural embedding** of X into its bidual X^{**}. This shows that we can identify X with the subspace JX of X^{**}.

Definition 4.7. $(X, \|\cdot\|)$ is said to be **reflexive** if the canonical embedding is surjective, i.e., $JX = X^{**}$.

Remark 4.5. (a) X isreflexive implies that the canonical embedding J is a isometrical isomorphism to X^{**}, hence X is a Banach space.

(b) Banach space X that is isometriccally isomorphic to X^{**} may be NOT reflexive. See *R.C. James [1951]. A non-reflexive Banach space isometric with its second conjugate space. Proc. Nat. Acad. Sci. USA, 37, 174-177.*

Example 4.10. Every finite-dimensional normed linear space is reflexive.

Example 4.11. For $1 < p < \infty$, $L^p(\mu)$ is reflexive. See Example 4.7.

Example 4.12. c_0, and $C[0,1]$ both are non-reflexive.

(a) $c_0^* = l_1$, so $c_0^{**} = (l_1)^* = l_\infty$. With these identifications, the natural map $c_0 \to c_0^{**}$ is precisely the inclusion map $c_0 \to l_\infty$.

(b) Note that $C[0,1]$ is separable: every continuous function an be approximate by piecewise linear functions with rational nodes and rational ordinates. On the other hand, $C[0,1]^*$ is not separable; the linear functionals ℓ_s defined by

$$\ell_s(f) = f(s), \quad -1 \le s \le 1$$

are clearly each bounded by 1, and equally clearly

$$|\ell_s - \ell_t| = 2 \quad \text{ for } s \neq t$$

since the $\{\ell_s\}$ form a nondenumerable collection, $C[0,1]^*$ cannot contain a dense denumerable subset. It follows now that $C[0,1]^{**} \neq C[0,1]$.

Lemma 4.23. *Let M be a closed subspace of X and $J_X : X \to X^{**}$ and $J_M : M \to M^{**}$ be the natural maps. Let $i : M \to X$ is the inclusion map, then there exists a linear isometry embedding $\Phi : M^{**} \to X^{**}$ such that the following diagram commutes.*

$$\begin{tikzcd}
X \arrow[r, "J_X"] & X^{**} \\
M \arrow[u, "i"] \arrow[r, "J_M"] & M^{**} \arrow[u, "\Phi"']
\end{tikzcd}$$

Proof. For any $y^{**} \in M^{**}$, define

$$\langle x^*, \Phi y^{**} \rangle = \langle x^*|_M, y^{**} \rangle\,, \quad \text{for any } x^* \in X^*\,.$$

Then it's easy to check that Φ is a linear isometry. Then for any $y \in M$ and $x^* \in X^*$,

$$\langle x^*, \Phi(J_M y) \rangle = \langle x^*|_M, J_M y \rangle = \langle y, x^*|_M \rangle = \langle y, x^* \rangle = \langle x^*, J_X y \rangle\,.$$

So the diagram commutes. □

Theorem 4.24. *A closed linear M of a reflexive space X is reflexive.*

Proof. Using the the lemma above, for any $y^{**} \in M^{**}$, there is a $y \in X$ such that $J_X y = \Phi y^{**}$. We only need to show that $y \in M$.

If not, there is $x^* \in M^\perp$, $\|x^*\| = 1$ and $\langle y, x^* \rangle > 0$. But then $x^*|_M = 0$ so that

$$\langle y, x^* \rangle = \langle x^*, J_X y \rangle = \langle x^*, \Phi y^{**} \rangle = \langle x^*|_M, y^{**} \rangle = 0\,.$$

This is a contradiction. □

Theorem 4.25. *A Banach space X is reflexive iff its dual X^* is reflexive.*

Proof. Assume that X is reflexive. Let $J_X : X \to X^{**}$ and $J_{X^*} : X^* \to X^{***}$ be the canonical embeddings of X and X^* respectively. We must show that J_{X^*} is surjective. To that end, let $\xi \in X^{***}$ and consider the following communicate diagram to define a functional x^* on X by $x^* = \xi \circ J_X$.

$$\begin{tikzcd}
X \arrow[r, "J_X"] \arrow[dr, "x^*"'] & X^{**} \arrow[d, "\xi"] \\
 & \mathbb{F}
\end{tikzcd}$$

It is obvious that x^* is linear since both ξ and J_X are linear. Also, for each $x \in X$,

$$|\langle x, x^* \rangle| = |\langle J_X x, \xi \rangle| \leqslant \|\xi\| \|x\| \,,$$

so that $x^* \in X^*$. We konw show that $J_{X^*} x^* = \xi$. In fact, for any $J_X x \in X^{**}$,

$$\langle J_X x, J_{X^*} x^*, \rangle = \langle x^*, J_X x \rangle = \langle x, x^* \rangle = \langle J_X x, \xi \rangle \,.$$

Suppose now that X^* is reflexive. Then the canonical embedoing $J_{X^*} : X^* \to X^{***}$ is surjective. If $J_X X \neq X^{**}$, take $x^{**} \in X^{**} \backslash J_X X$. Since X is Banach space, $J_X X$ is a closed subspace of X^{**}, it follows from Corollary 4.17 that there is a functional $J_{X^*} x^* \in X^{***}$ such that $\|x^*\| = 1$, and $\langle x^{**}, J_{X^*} x^* \rangle > 0$ and

$$\langle J_X x, J_{X^*} x^* \rangle = 0 \;\; \text{for all } x \in X \,.$$

Hence, for each $x \in X$.

$$0 = \langle J_X x, J_{X^*} x^* \rangle = \langle x^* . J_X x \rangle = \langle x, x^* \rangle$$

Thus $x^* = 0$, which is a contradiction. $\square$

4.6 Weak topologies

We have made the point that a norm on a vector space X induces a metric. A metric, in turn, induces a topology on X called the norm topology, and X^* also has a natural norm topology. In this section we investigate some of the properties of weak topology, defined in Section 3.7, on normed sapce X and weak* topology on X^*.

Proposition 4.26. $(X, \|\cdot\|)$ *is normed sapce, let* τ *be the norm topology.*

(a) $\sigma(X, X^*) \subset \tau$

(b) $\sigma(X, X^*) = \tau$ *if and only if* X *is finite-dimensional.*

Proof. (a) is obvious.

To show (b), when X is finite-dimensional, we ahve proved that all the vector topology on X are the same one, see Corollary 3.18, so $\tau_w = \tau$.

Conversely, assume that X is infinite-dimensional. The (open) unit ball

$$B(0,1) = \{x : \|x\| < 1\},$$

is open in τ, but we claim that $B(0,1) \notin \tau_w$. If not, there exists a finite subset Φ of X^* and $\epsilon > 0$ so that

$$V(\Phi, \epsilon) \subset B(0,1).$$

Thus,

$$\bigcap_{x^* \in \Phi} \mathrm{ker} x^* \subset B(0,1).$$

Since X is infinite-dimensional, $\cap_{x^* \in \Phi} \mathrm{ker} x^* \neq \{0\}$ is a subspace of X, this is a contradiction. $\square$

Note that there are three topologies on X^*: the norm topology, the weak topology, denoted by $\sigma(X^*, X^{**})$, and the weak* topology, denoted by $\sigma(X^*, X)$.

Proposition 4.27. *Let τ^* denote the norm topology on X^*. Then*

(a) $\sigma(X^*, X) \subset \sigma(X^*, X^{**}) \subset \tau^*$

(b) $\sigma(X^*, X^{**}) = \tau^*$ *iff X is finite-dimensional.*

(c) $\sigma(X^*, X) = \sigma(X^*, X^{**})$ *if and only if X is reflexive.*

Proof. (a), (b) is obvious.

To show (c), " if " part is obvious. On the other hand, note that

$$(X^*, \sigma(X^*, X))^* = J(X) \text{ and, } (X^*, \sigma(X^*, X^{**}))^* = X^{**}.$$

Therefore if $\sigma(X^*, X) = \sigma(X^*, X^{**})$, then $J(X) = X^{**}$, X is reflexive. □

Convergence

In Example 3.1, we have shown that

- A net $\{x_i\}$ in X is convergent to some $x \in X$ in τ_w if and only if

$$\langle x_i, x^* \rangle \to \langle x, x^* \rangle \text{ , for all } x^* \in X^*.$$

- A net $\{x_i^*\}$ in X^* is convergent to some $x^* \in X$ in τ_{w^*} if and only if

$$\langle x, x_i^* \rangle \to \langle x, x^* \rangle \text{ , for all } x \in X.$$

However, in this subsection, we care more about convergent sequence. To avoid any confusion we shall sometimes say, $x_n \to x$ weakly in $\sigma(X, X^*)$, or $x_n^* \to x^*$ weakly in $\sigma(X, X^*)$. In ovder to be totuly clear we sonil sometimes empasize norm convergence by saying, "$x_n \to x$ in norm," meaning that $\|x_n - x\| \to 0$.

Example 4.13. X is finite-dimensional, the weak topology $\sigma(X, X^*)$ and the usual topology are the same. In particular, a sequence (x_n) converges weakly if and only if it converges strongly.

Example 4.14. If a sequence in l^1 converges weakly, it converges in norm.

Aproof can be found in [J.Conway] chapter V. 5.2. Proposition.

Example 4.15. In $L^2[0, 2\pi]$, let $x_n = x_n(t) = \sin nt$. By Riemann-Lebesgue theorem, for any $x \in L^2[0, 2\pi]$,

$$\langle x, x_n \rangle = \int_0^{2\pi} x(t) sin(nt)\, dt \to 0$$

This is $x_n \xrightarrow{w} 0$, but $\|x_n\| = \pi$ for all n, x_n doesn't converge to 0 in norm.

Generally, H is a Hilbert space and $\{e_n\}$ is a orthonomal systerm, by Bessel's inequality we konw $e_n \xrightarrow{w} 0$ but $\|e_n\| = 1$ for all n.

Proposition 4.28. *$\{x_n\}$ is a sequence in X.*

(a) If $x_n \to x$ weakly in $\sigma(X, X^\star)$, then $\{x_n\}$ is (norm) bounded and,

$$\|x\| \leqslant \liminf_{n\to\infty} \|x_n\| . \tag{4.11}$$

(b) If $x_n \to x$ weakly in $\sigma(X, X^\star)$ and if $x_n^* \to x^*$ in norm, then

$$\langle x_n, x_n^* \rangle \to \langle x, x^* \rangle . \tag{4.12}$$

Proof. To show (a), by PUB, we have that $\{x_n\}$ is bounded, and note that

$$|\langle x, x^* \rangle| = \lim_{n\to\infty} |\langle x_n, x^* \rangle| \leqslant \liminf_{n\to\infty} \|x_n\| \|x^*\|$$

for any $x^* \in X^*$, Thus

$$\|x\| \leqslant \liminf_{n\to\infty} \|x_n\| .$$

To show (b), note that

$$\begin{aligned} |\langle x_n, x_n^* \rangle - \langle x, x^* \rangle| &\leqslant |\langle x_n, x_n^* \rangle - \langle x_n, x^* \rangle| + |\langle x_n, x^* \rangle - \langle x, x^* \rangle| \\ &\leqslant \|x_n^* - x^*\| \sup_n \|x_n\| + |\langle x_n, x^* \rangle - \langle x, x^* \rangle| . \end{aligned}$$ □

By using Banach-Steinhaus theorem, we get the following result, which is useful when we discussing sequential compactness.

Theorem 4.29.

- *X is normed sapce, a sequence $\{x_n\}$ in X is weakly convergent iff*
  - *(a) $\{x_n\}$ is uniformly bounded and,*
  - *(b) there exists $N \subset X^*$ is dense, so that $\{\langle x_n, x^* \rangle\}$ converges for all $x^* \in N$.*
- *X is Banach sapce, a sequence $\{x_n^*\}$ in X is weakly convergent iff*
  - *(a) $\{x_n^*\}$ is uniformly bounded and,*
  - *(b) there exists $M \subset X$ is dense, so that $\{\langle x, x_n^* \rangle\}$ converges for all $x \in M$.*

Compact, sequentially compact sets in weak* topology

Observe that $X^* \subset \mathbb{F}^X = \prod_X \mathbb{F}$ and that the weak* topology $\sigma(X^*, X)$ on X^* is the relative topology on X^* induced by the product topology on $\prod_X \mathbb{F}$.

Theorem 4.30 (Alaoglu's theorem)**.** *$(X, \|\cdot\|)$ is a normed space. Then the closed unit ball in X^* is weak*-compact, i.e.,*

$$B_{X^*} = \{x^* \in X^* : \|x^*\| \leqslant 1\}$$

is compact with respect to the topology $\sigma(X^*, X)$.

Proof. **Step 1**. For each $x \in X$, let

$$D_x = \{\lambda \in \mathbb{F} : |\lambda| \leqslant \|x\|\}.$$

Then, for each $x \in X, D_x$ is a closed interval in $\mathbb{R}$ or a closed disk in $\mathbb{C}$. Equipped with the natural topology, D_x is compact for each $x \in X$. Let

$$D = \prod_{x \in X} D_x \, .$$

By *Tychonoff's theorem*, D is compact with respect to the product topology.The points of D are just $\mathbb{F}$-valued functions (not to be linear) f on X such that $|f(x)| \leqslant \|x\|$ for each $x \in X$. Obviously, we have

$$B_{X^*} \subset D \, .$$

Step 2. We observe that the topology that D induces on $B\,(X^*)$ is precisely the weak* topology on B_{X^*}. It remains to show that B_{X^*} is a closed subset of D.

To this end, let $\{x_i^*\}$ be a net in B_{X^*} and $x_i^* \to x^* \in D$ in the product topology. Then $\langle x, x_i^* \rangle \to \langle x, x^* \rangle$ for all $x \in X$. Thus

$$\begin{aligned} \langle \alpha x + \beta y, x^* \rangle &= \lim_i \langle \alpha x + \beta y, x_i^* \rangle \\ = \lim_i \alpha \langle x, x_i^* \rangle + \beta \langle y, x_i^* \rangle &= \alpha \langle x, x^* \rangle + \beta \langle x, x^* \rangle \, . \end{aligned}$$

for all x, y in X and α, β in $\mathbb{F}$, thus x^* is linear. Since

$$|\langle x, x^* \rangle| = \lim_i |\langle x, x_i^* \rangle| \leqslant \|x\|$$

for all $x \in X$, x^* is continuous, and $\|x^*\| \leqslant 1$. That is,

$$x^* \in B_{X^*}.$$

Therefore B_{X^*} is closed in D and hence compact. $\square$

Theorem 4.31. *X is a separable Banach space. Then B_{X^*} is weak*-sequentially compact set.*

Proof. Assume $\{x_n\}$ is dense in X. Given any sequence $\{x_n^*\}$ in B_{X^*}, by *diagonal process*, we can select a subsequence $\{y_n^*\}$ of $\{x_n^*\}$ so that

$$\lim_{n\to\infty} \langle x_k, y_n\rangle \text{ exists for all } x_k\,.$$

By Theorem 4.29, $\{y_n^*\}$ is converges in weak* topology. □

Remark 4.6. In fact, X is a Banach space. Then B_{X^*} is metrizable in the weak* topology iff X is separable.

Compact, sequentially compact sets in weak topology : reflexivity revisited

Theorem 4.32 (Goldstine). $(X, \|\cdot\|)$ *is normed space, J is the natural embeding of X into X^{**}. Let B_X and $B_{X^{**}}$ be the closed unit ball in X and X^{**}, respectively. Then JB_X is dense in $B_{X^{**}}$ relative to the weak* topology $\sigma(X^{**}, X^*)$ on X^{**}, that is,*

$$\overline{JB_X}^{\sigma(X^{**},X^*)} = B_{X^{**}}\,.$$

Proof. Clearly, $JB_X \subset B_{X^{**}}$. By Alaoglu's theorem, $B_{X^{**}}$ is compact with respect to $\sigma(X^{**}, X^*)$, since weak* topology is Hausdorff, so $B_{X^{**}}$ is closed, thus we have

$$\overline{JB_X}^{\sigma(X^{**},X^*)} \subset B_{X^{**}}\,.$$

On the other hand, if there exists $x_0^{**} \notin \overline{JB_X}^{\sigma(X^{**},X^*)}$, note that $\overline{JB_X}^{\sigma(X^{**},X^*)}$ is balanced closed convex set, by Theorem 3.28, there exists $x^* \in X^*$ (Indeed, there we used Theorem 3.32 wo deduce that $(X^{**}, \sigma(X^{**}, X^*))^* = X^*$), so that

$$|\langle x, x^*\rangle| \leqslant 1 < |\langle x^*, x_0^{**}\rangle|\ ,\ \text{for all } x \in B_X\,.$$

Thus $\|x^*\| \leqslant 1$, and $\|x_0^{**}\| > 1$, so $x_0^{**} \notin B_{X^{**}}$. So

$$B_{X^{**}} \subset \overline{JB_X}^{\sigma(X^{**},X^*)}\,.$$

Thus we get the equality. □

Corollary 4.33. *X is a normed space, J is the canonical embedding of X into X^{**}. Then $J(X)$ is dense in X^{**} relative to the weak* topology $\sigma(X^{**}, X^*)$ on X^{**}. That is,*

$$\overline{JX}^{\sigma(X^{**},X^*)} = X^{**}.$$

Proof. Note that

$$\begin{aligned} \overline{JX}^{\sigma(X^{**},X^*)} &= \overline{\bigcup_{n=1}^{\infty} nJB_X}^{\sigma(X^{**},X^*)} = \bigcup_{n=1}^{\infty} n\overline{JB_X}^{\sigma(X^{**},X^*)} \\ &= \bigcup_{n=1}^{\infty} nB_{X^{**}} = X^{**}. \end{aligned}$$

□

Weak compactness and reflexity

Theorem 4.34 (Kakutani's theorem)**.** *X is a normed space, B_X is the closed unit ball. Then B_X is weakly compact if and only if X is reflexive.*

Proof. First of all we show that

$$J : (X, \sigma(X, X^*)) \to (J(X), \sigma(X^{**}, X^*)|_{J(X)})$$

is a linear homeomorphism. Obviously J is linear bijection. To see J is homeomorphism, $\{x_i\}$ converges to $x \in X$ with respect to $\sigma(X, X^*)$, $\Leftrightarrow$

$$\langle x_i, x^* \rangle \to \langle x, x^* \rangle, \text{ for all } x^* \in X^*.$$

$\Leftrightarrow$

$$\langle x^*, Jx_i \rangle \to \langle x^*, Jx \rangle, \text{ for all } x^* \in X^*.$$

$\Leftrightarrow$ Jx_i converges to Jx with respect to $\sigma(X^{**}, X^*)|_{J(X)}$. Therefore

- If B_X is weakly compact, since J is continuous, JB_X is cpmpact in X^{**}. By Goldstine theorem, $JB_X = B_{X^{**}}$. Thus $JX = X^{**}$.

- On the other hand, if $J(X) = X^{**}$, note that $B_{X^{**}} = J(B_X)$ is compact with respect to $\sigma(X^{**}, X^*)$. Since J is a linear homeomorphism, B_X is compact with respect to $\sigma(X, X^*)$. $\square$

Corollary 4.35. *X is a reflexive. Let $K \subset X$ be a norm bounded, closed, and convex subset of X. Then K is weakly compact.*

Proof. K is closed and convex, so K is weakly closed subset of a weakly compact set, thus K is compact. $\square$

Weak sequential compactness and reflexity

In connection with the compactness properties of reflexive spaces we also have the following two results about sequential compactness.

Theorem 4.36. *Let X be a reflexive Banach space, B_X is the closed unit ball, then B_X is weakly sequentially compact.*

Proof. Given a sequence $\{x_n\}$ in B_X, let M be the closed subspace spanned by $\{x_n\}$, i.e.,

$$M \coloneqq \overline{\text{span}}\{x_n\}$$

Then M is separable and M is reflexive, by Theorem 4.24 or Exercise 4.2, M is reflexive. Since $Y^{**} = Y$ is separable, by Theorem 4.19, Y^* is separable. Let $\{y_n^*\}$ dense in Y^*. By the *diagonal process*, we can select a subspace $\{z_n\}$ of $\{x_n\}$ so that

$$\lim_{n\to\infty} \langle z_n, y_k \rangle \text{exists for all} k \in \mathbb{N}^+ .$$

Thus we get $\{z_n\}$ is weakly convergent by Theorem 4.29. $\square$

Remarkblely, the converse of Theorem 4.36 is also true, namely the following

Theorem 4.37 (Eberlein-Smulian theorem)**.** *X is a Banach space, B_X is the closed unit ball. If B_X is weakly sequentially compact, then X is reflexive.*

Proof. The proof is rather delicate and is omitted, see, e.g., [K. Yosida Functional Analysis. □

Remark 4.7. In order to clarify the connection between Theorem 4.36 and Theorem 4.37, it is useful to recall the following facts:

(a) If X is a metric space, then compactness is equivalent to sequential sompactness. But in infinite-dimensional spaces, the weak topology is never metrizable, so this result is not trival.

(b) There exist compact topological spaces X and some sequences in X without any convergent subsequence. A typical example is $X = B_{X^*}$, which is compact in the topology $\sigma(X^*, X)$; when $X = \ell^\infty$ it is easy to construct a sequence in X without any convergent subsequence.(see [H.Brezis] exercise 3.18)

(c) If X is a topological space with the property that every sequence admits a convergent subsequence, then X need not be compact.

We give an application of this sequential compactness.

Theorem 4.38. *Let X be a reflexive Banach space, K a closed, convex subset of X, For each x in X, there is a point k of K so that*

$$\|x - k\| = dist(x, K) = \inf_{k \in K} \|x - k\| .$$

Remark 4.8. Note that such k may not be unique!

Proof. Without loss of generality, assume $x = 0 \notin K$, then there is $\{k_n\}$ in K so that

$$\|k_n\| \to d = dist(x, K) > 0 .$$

Then there exists some $k \in \overline{K}^{\sigma(X,X^*)} = K$ so that

$$k_n \xrightarrow{w} k\,.$$

Note that

$$\|k\| \leqslant \liminf_{n\to\infty} \|k_n\| = d$$

So we get the desired point. □

EXERCISE

Exercise 4.2. Using Kakutani's theorem wo show that:

(a) A Banach space X is reflexive iff its dual X^* is reflexive.

(b) A closed linear subspace M of a reflexive space X is reflexive.

4.7 Uniformly convex space

A norm is called *strictly subadditive* if

$$\|x+y\| \leqslant \|x\| + \|y\|$$

in strict inequality holds except when x or y is a nonnegative multiple of the other. Furthermore for each of these norms the condition holds uniformly, in the following sense:

For any pair of unit vectors x, y, the norm of $(x+y)/2$ is strictly less than 1 by an amount that depends only on $\|x-y\|$. More explicitly, there is an increasing function $\epsilon\,(r)$ defined for positive $r \in [0,2]$,

$$\epsilon(r) > \epsilon(0) = 0, \text{ for any } r > 0\,, \quad \lim_{r\to 0} \epsilon\,(r) = 0 \tag{4.13}$$

such that for all x, y such that $\|x\| \leqslant 1, \|y\| \leqslant 1$, the inequality

$$\left\| \frac{x+y}{2} \right\| \leqslant 1 - \epsilon(\|x-y\|) \tag{4.14}$$

holds.

Definition 4.8. $(X, \|\cdot\|)$ is called **uniformly convex**, if the norm satisfies (4.14) for all vectors x, y in B_X, where $\epsilon(r)$ is some function satisfying (4.13).

Some other textbooks defined uniformly convex space by : any $\epsilon > 0$, there exists some $\delta > 0$, depending on ϵ so that

$$\left\| \frac{x+y}{2} \right\| \leq 1 - \delta$$

for all x and y in B_X that $\|x-y\| \geqslant \epsilon$.

The uniform convexity is a geometric property of the unit ball: if we slide a rule of length $\varepsilon > 0$ in the unit ball, then its midpoint must stay within a ball of radius $(1-\delta)$ for some $\delta > 0$. In particular, the unit sphere must be "round" and cannot include any line segment.

Example 4.16. Let $X = \mathbb{R}^2$. The norm $\|x\|_2 = \sqrt{|x_1|^2 + |x_2|^2}$ is uniformly convex, while the norm $\|x\|_1 = |x_1| + |x_2|$ and the norm $\|x\|_\infty = \max(|x_1|, |x_2|)$ are not uniformly convex. This can be easily seen by staring at the unit balls, as shown in Figure 4.1.

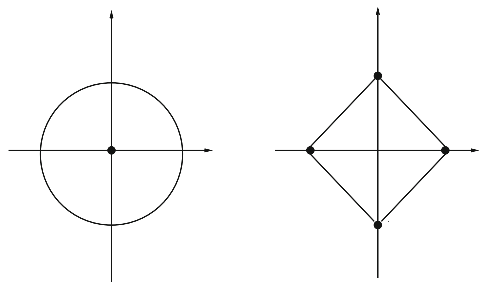

Figure 4.1: Unit ball B_X for $\|\cdot\|_2$ and $\|\cdot\|_1$

Example 4.17. L^p space is uniformly convex for $1 < p < \infty$.

Convergence

Theorem 4.39. *X is a uniformly convex space, x_n, x in X. Then x_n converges to x if and only if $\|x_n\| \to \|x\|$ and x_n weakly convergent to x.*

Proof. Without loss of generality, we assume $\|x\| > 0$, and $\|x_n\| > 0$ for all n. Note that

$$\epsilon\left(\left\|\frac{x_n}{\|x_n\|} - \frac{x}{\|x\|}\right\|\right) \leqslant 1 - \frac{1}{2}\left\|\frac{x_n}{\|x_n\|} + \frac{x}{\|x\|}\right\| .$$

To show $\|x_n - x\| \to 0$, it suffices to show that

$$\epsilon\left(\left\|\frac{x_n}{\|x_n\|} - \frac{x}{\|x\|}\right\|\right) \to 0\,, \tag{4.15}$$

since $\epsilon(\cdot)$ is incresing, continuous at 0, $\epsilon(r) > 0$ for any positive r, and $\|x_n\| \to \|x\|$. To show (4.15), it suffices to show that

$$\left\| \frac{x_n + x}{2} \right\| \to \|x\| .$$

Since $x_n \xrightarrow{w} x$, we have $(x_n + x)/2 \xrightarrow{w} x$ and hence

$$\|x\| \leqslant \liminf_{n\to\infty} \left\| \frac{x_n + x}{2} \right\|$$

On the other hand,

$$\limsup_{n\to\infty} \left\| \frac{x_n + x}{2} \right\| \leqslant \limsup_{n\to\infty} \frac{\|x_n\| + \|x\|}{2} = \|x\| .$$

So we get the desired result. □

Best approximation

Theorem 4.40. *Let X be a uniformly convex space. Let K be a complete, convex subset of X. Then each $x \in X$ has a unique best approximation in K, i.e. there is a unique point $k \in K$ satisfying*

$$\|x - k\| = dist(x, K) \coloneqq \inf_{k\in K} \|x - k\| . \tag{4.16}$$

Proof. Without loss of generality, assume $x = 0 \notin K$, then there is $\{k_n\}$ in K so that

$$\|k_n\| \to d = dist(x, K) > 0 .$$

Assume $\|k_n\| > 0$ for all n, denote

$$x_n = \frac{k_n}{\|k_n\|}$$

Then

$$\epsilon\left(\|x_n - x_m\|\right) \leqslant 1 - \left\| \frac{x_n + x_m}{2} \right\| .$$

Note that

$$\begin{aligned}\frac{x_n + x_m}{2} &= \frac{k_n}{2\|k_n\|} + \frac{k_m}{2\|k_m\|} \\ &= \left(\frac{1}{2\|k_n\|} + \frac{1}{2\|k_m\|}\right)[ck_n + (1-c)k_m]\end{aligned}$$

where c is a constant in $(0,1)$. Since $ck_n + (1-c)k_m \in K$.

$$\frac{x_n + x_m}{2} \geqslant \frac{d}{2}\left(\frac{1}{\|k_n\|} + \frac{1}{\|k_m\|}\right)$$

Thus

$$\lim_{n,m\to\infty} \epsilon\left(\|x_n - x_m\|\right) \to 0\,.$$

By the properties of $\epsilon(\cdot)$, we konw

$$\lim_{n,m\to\infty} \|x_n - x_m\| \to 0\,.$$

Since $\|k_n\| \to d$, then

$$\lim_{n,m\to\infty} \|k_n - k_m\| \to 0\,.$$

Thus $\{k_n\}$ is a Cauchy sequence, sence K is complete, there is some $k \in K$ so that $k_n \to k$, thus $\|k\| = d$.

If there exists $k' \in K$ so that $\|k'\| = d$, then

$$\epsilon\left(\left\|\frac{k-k'}{d}\right\|\right) \leqslant 1 - \left\|\frac{k+k'}{2d}\right\| \leqslant 0\,.$$

since $k + k'/2 \in K$, so $k = k'$. □

Remark 4.9. By Theorem 4.41, and Theorem 4.38, the existence of k is obvious, the uniform convexity guarantee the uniqueness.

Uniform convexity and reftexity

Theorem 4.41 (Milman-Pettis)**.** *Every uniformly convex Banach space is reftexive.*

Proof. The proof can be found in Haim Brezis, *Functional Analysis, Sobolev Spaces and Partial Differential Equations*, Theorem 3.31. □

REMARK 4.10. Uniform convexity is a *geometric property* of the norm; an equivalent norm need not be uniformly convex. On the other hand, reflexivity is a *topological property*: a reflexive space remains reflexive for an equivalent norm. It is a striking feature of this theorem that a geometric property implies a topological property. Uniform convexity is often used as a tool to prove reflexivity; but it is not the ultimate tool, since there are some weird reflexive spaces that admit no uniformly convex equivalent norm!

4.8 Projections, Direct Sum Decomposition

Let X be a normed space, $P \in \mathcal{B}(X)$ is called a **projection** if

$$P^2 = P\,.$$

It's easy to find that

Proposition 4.42.

(a) *P is an projection if and only if $I - P$ is an projection.*

(b) $\ker P = \operatorname{ran}(I - P)$, $\operatorname{ran}(P) = \ker(I - P)$ *and both* $\operatorname{ran}(P)$ *and* $\ker(P)$ *are closed subspaces of X.*

(c) $X = \operatorname{ran} P \oplus \ker(P)$.

Proof. To show (a), observe that

$$(I - P)^2 = I - 2P + P^2 = I - 2P + P = I - P\,,$$

thus $I - P$ is also an projection.

To show (b), since P is continuous, $\ker(P)$ is a closed subspace of X. Also,

$$x \in \operatorname{ran}(P) \Leftrightarrow Px = x \Leftrightarrow (I - P)x = 0 \Leftrightarrow x \in \ker(I - P).$$

Similarly, $\operatorname{ran}(I - P) = \ker P$.

To show (c), note that $x = Px + (I - P)x$. □

From this proposition, we say P is the projection from X to $\operatorname{ran}P$. Moreover, M is a closed subspace of X, we say there exists a peojection P from X to M, if there exists projection P such that $\operatorname{ran}P = M$. When does the projection exist ?

Theorem 4.43. *X is a Banach space. If M, N are two closed subspace of , such that $X = M \oplus N$. then there is an projection P such that*

$$\operatorname{ran} P = M \ \textit{and}\ \ker P = N\,.$$

Proof. For each $x \in X$, there is a unique composition

$$x = m + n\,.$$

where $m \in M$ and $n \in N$. Define P by

$$Px := m\,.$$

P is linear, since the composition is unique. Clearly,$P^2 = P$. By closed graph theorem, it suffices to show P has a closed graph.

Assume $x_n \to x$ and $Px_n \to m$, then

$$x_n - Px_n \to x - m\,.$$

Sicne M, N are closed, $m \in M$ and $x - m \in N$. Thus we have $Px = m$. $\square$

Definition 4.9. X is a normed space, a closed subspace M is said to be **complemented in** X, if there exists a closed subspace N such that

$$X = M \oplus N\,.$$

Thus a projection operator is equivalent to a direct sum decomposition of X. At present we need only the following simple facts.

Lemma 4.44. *M is a closed subspace of normed space X. If*

$$\text{dim}X < \infty \textit{ or } \text{codim}X < \infty\,,$$

there is a projection P form X to M, i.e., M is complemented.

Proof. Case 1. Assume $\text{dim}M = n$, and

$$\{e_1, e_2, \cdots, e_n\}$$

a base for M. For each $x \in M$, there exists unique $\{c_1, \cdots, c_n\}$, depending on x, so that

$$x = c_1(x)e_1 + \cdots + c_n(x)e_n\,.$$

Clearly, $c_k(\cdot)$ is continuous linear functional on M. By H-B extension theorem, it can be extented on X as a continuous linear functional. Then define $P : X \to X$ by

$$Px = \sum_{i=1}^{n} \alpha_i(x) e_i \,,$$

and it's easy to check that P is a projection form X to M.

Case 2. Let $\dim X/M = n$, and

$$\{\widetilde{e}_1, \cdots, \widetilde{e}_n\}$$

a base for X/M. Pick any $e_i \in \widetilde{e}_i$, then

$$\{e_1, e_2, \cdots, e_n\}$$

is linear independent. Let

$$N = \operatorname{span}\{e_1, e_2, \cdots, e_n\} \,,$$

then $X = M \oplus N$, and N is closed. □

4.9 Adjoints

We shall now associate with each $T \in \mathcal{B}(X, Y)$ its adjoint, an operator $T^* \in \mathcal{B}(Y^*, X^*)$, and will see how certain properties of T are reflected in the behavior of T^*.

If X and Y are finite-dimensional, every $T \in \mathcal{B}(X, Y)$ can be represented by a matrix $[T]$; in that case, $[T^*]$ is the transpose of $[T]$, provided that the various vector space bases are properly chosen. No particular attention will be paid to the finite-dimensional case in what follows, but historically linear algebra did provide the background and much of the motivation that went into the construction of what is now known as operator theory.

Suppose X and Y are normed spaces. To each $T \in \mathcal{B}(X, Y)$ corresponds a unique $T^* \in \mathcal{B}(Y^*, X^*)$ that satisfies

$$\langle Tx, y^* \rangle = \langle x, T^* y^* \rangle \tag{4.17}$$

for all $x \in X$ and all $y^* \in Y^*$. In fact, $T^* y^* = y^* \circ T$, clearly $T^* \in \mathcal{B}(Y^*, X^*)$. T^* is called the **adjoint** of T. Also, one can show that $T^* : (Y^*, \sigma(Y^*, Y)) \to (X^*, \sigma(X^*, X))$ is continuous.

$$\begin{tikzcd}
X \arrow[r, "T"] \arrow[dr, dashed, "T^\star y^\star"'] & Y \arrow[d, "y^\star"] \\
& \mathbb{F}
\end{tikzcd}$$

Figure 4.2: The adjoint of T

Proposition 4.45. *X and Y are normed spaces, $T \in \mathcal{B}(X, Y)$. Then $T^* \in \mathcal{B}(Y^*, X^*)$ and $\|T^*\| = \|T\|$.*

Proof. To show $\|T^*\| = \|T\|$, note that

$$\|T\| = \sup_{x \in B_X, y^* \in B_{Y^*}} |\langle Tx, y^* \rangle| = \sup_{x \in B_X, y^* \in B_{Y^*}} |\langle x, T^* y^* \rangle| = \|T^*\| . \qquad \square$$

Remark 4.11. By this property, $T \mapsto T^{**}$ is an linear isometic embedding form $\mathcal{B}(X,Y)$ to $\mathcal{B}(X^{**}, Y^{**})$. And we can regard T^{**} as a cintinuous extension of T on X^{**}, in fact,

$$T^{**} \circ J_X = J_Y \circ T \tag{4.18}$$

Obviously, if $A, B \in \mathcal{B}(X,Y)$, and α, β are scalars, then

$$(\alpha A + \beta B)^* = \alpha A^* + \beta B^*$$

If $A \in \mathcal{B}(X,Y)$ and let $B \in \mathcal{B}(Y,Z)$, then

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

Theorem 4.46. *Suppose X and Y are normed spaces, and $T \in \mathcal{B}(X,Y)$. Then we have*

$$\operatorname{ran}(T)^{\perp} = \ker(T^*) \qquad \textit{and} \qquad {}^{\perp}\operatorname{ran}(T^*) = \ker(T) .$$

Proof. Note that

$$\begin{aligned} y^* \in \operatorname{ran}(T)^{\perp} &\Leftrightarrow \langle Tx, y^* \rangle = 0 \text{ for all } x \\ &\Leftrightarrow \langle x, T^* y^* \rangle = 0 \text{ for all } x \Leftrightarrow T^* y^* = 0 , \end{aligned}$$

and

$$\begin{aligned} x \in {}^{\perp}\operatorname{ran}(T^*) &\Leftrightarrow \langle x, T^* y^* \rangle = 0 \text{ for all } y^* \\ &\Leftrightarrow \langle Tx, y^* \rangle = 0 \text{ for all } y^* \Leftrightarrow Tx = 0 , \end{aligned} \qquad \square$$

By Theorem 4.20, we have

Corollary 4.47.

$$\overline{\operatorname{ran}(T)} = {}^{\perp}\ker(T^*) \qquad \textit{and} \qquad \overline{\operatorname{ran}(T^*)}^{\sigma(X,X^*)} = \ker(T)^{\perp}\,.$$

Corollary 4.48. *(a)* *T is injective iff $\operatorname{ran}(T^*)$ is weak***-dense in X^*,*

(b) *$\operatorname{ran}(T)$ is dense in Y iff T^* is injective.*

Many of the nontrivial properties of adjoints depend on the complete ness of X and Y (the open mapping theorem will play an important role). For this reason, it will be assumed throughout that X and Y are Banach spaces.

Theorem 4.49. *X and Y are Banach spaces, and $T \in \mathcal{B}(X,Y)$. Then T is invertible if and only if T^* is invertibe, in this case*

$$(T^*)^{-1} = (T^{-1})^*\,. \tag{4.19}$$

Proof. If T is invertible, then $T^{-1} \in \mathcal{B}(Y,X)$, so

$$(T^{-1})^*T^* = (TT^{-1})^* = I_Y^* = I_{Y^*}\,,$$

and

$$T^*(T^{-1})^* = (T^{-1}T)^* = I_X^* = I_{X^*}\,.$$

Thus T^* is invertible, and $(T^*)^{-1} = (T^{-1})^*$.

If T^* is invertible, by the preceeding proof then T^{**} is, too. By the inverse mapping theorem, T^{**} is a linear hemeomorphism, thus T is injection. On the other hand, since X is Banach space, $J_X(X)$ is Banach space, then $T^{**}J_X(X) = J_Y(TX)$ is a closed subspace in $J_Y(Y)$. Thus TX is a closed subspace of Y. By Corollary 4.47

$$\operatorname{ran}(T) = {}^{\perp}\ker(T^*) = Y\,. \qquad \square$$

Example 4.18. Let $(X, \mathcal{F}, \mu)$ and the multiplication operator M_ϕ on L^p, where $1 \leqslant p < \infty$, be as in Example 1.14. Let $\frac{1}{p} + \frac{1}{q} = 1$, then $M_\phi^* : L^q \to L^q$ is given by

$$M_\phi^* f = \phi f \text{ for all } f \in L^q .$$

In words, the adjoint of multiplication operator with simble ϕ on L^p is the multiplication operator with simble ϕ on L^q.

Example 4.19. Let the integral operator K on L^p, where $1 \leqslant p < \infty$, and kernel k be as in Example 1.15. Let $\frac{1}{p} + \frac{1}{q} = 1$, then $K^* : L^q \to L^q$ is the integer operator with kernel $k^*(x, y) := k(y, x)$.

The closed range theorem

Lemma 4.50. *X and Y are Banach spaces, $T \in \mathcal{B}(X, Y)$. Then T is surjective if and only if T^* is a linear homeomorphism between Y^* and* $\operatorname{ran}(T^*)$*.*

Proof. Necessity. By Theorem 4.46 we have $\ker(T^*) = \operatorname{ran}(T)^\perp = \{0\}$. So T^* is a injection. To show $(T^*)^{-1} : \operatorname{ran}(T^*) \to Y^*$ is continuous, it suffices to show that there exists some $M > 0$ so that

$$M \|T^* y^*\| \geqslant \|y^*\| , \text{ for all } y^* \in Y^* .$$

That is

$$M \sup_{x \in T(B_X)} |\langle x, y^* \rangle| \geqslant \sup_{y \in B_Y} |\langle y, y^* \rangle| , \text{ for all } y^* \in Y^* . \tag{4.20}$$

By open mapping theorem, there exists some $r > 0$, such that

$$r B_Y \subset T(B_X)$$

Therefore,

$$\sup_{y \in B_Y} |\langle y, y^* \rangle| = \frac{1}{r} \sup_{y \in r B_Y} |\langle y, y^* \rangle| \leqslant \frac{1}{r} \sup_{x \in T(B_X)} |\langle x, y^* \rangle|$$

for all $y^* \in Y^*$. So we get the desired result.

Sufficiency. We show that T is an open mapping, then, clearly, T is a surjective. By the proof of Theorem 4.8, it suffices to show that there exists some $r > 0$ so that

$$B_Y(0,r) \subset \overline{T(B_X(0,1))}\,.$$

Note that $T(B_X(0,1))$ is a balanced convex set, by Theorem 3.28, we have that for any $y \notin \overline{T(B_X(0,1))}^{\sigma(X,X^*)} = \overline{T(B_X(0,1))}$, there exists some $y^* \in Y^*$ so that

$$|\langle Tx, y^*\rangle| \leqslant 1 < \langle y, y^*\rangle\,, \text{ for any } x \in B_X(0,1)\,.$$

Therefore,

$$\|T^*y^*\| < \langle y, y^*\rangle \leqslant \|y\|\|y^*\|$$

Since $(T^*)^{-1}$ is continuous, there exises some $M > 0$ so that

$$M\|T^*y^*\| \geqslant \|y^*\|\,, \text{ for all } y^* \in Y^*\,.$$

Thus

$$\|y\| \geqslant \frac{1}{M}$$

We have

$$B_Y(0,1/M) \subset \overline{T(B_X(0,1))}\,. \qquad \square$$

The following consequence is useful in applications.

Theorem 4.51 (Surjection)**.** *X,Y are Banach spaces. $T \in \mathcal{B}(X,Y)$. Then*

$$\operatorname{ran}(T) = Y \iff \ker(T^*) = \{0\} \textit{ and } \operatorname{ran}(T^*) \textit{ is norm closed}\,.$$

Proof. Sufficiency. By the inverse mapping theorem, $\operatorname{ran}(T^*)$ is norm-closed implies T^* is linear homeomorphism. So by Lemma 4.50, T is a surjection.

Necessity. By Lemma 4.50, T is a surjection implies T^* is a linear homeomorphism between Y^* and $\operatorname{ran}(T^*)$, so T^* is injection and $\operatorname{ran}(T^*)$ is norm-closed. $\square$

Theorem 4.52 (Closed range theorem)**.** *X and Y are Banach spaces and $T \in \mathcal{B}(X,Y)$. Then each of the following three conditions implies the other two:*

(a) $\operatorname{ran}(T)$ *is closed in* Y *;*

(b) $\operatorname{ran}(T^*)$ *is weak*-closed in* X^* *;*

(c) $\operatorname{ran}(T^*)$ *is norm-closed in* X^* *.*

Proof. (a) $\Rightarrow$ (b). Note that

$$\overline{\operatorname{ran}(T^*)}^{\sigma(X,X^*)} = \ker(T)^{\perp}\,,$$

it suffices to show that for any $x^* \in \ker(T)^{\perp}$, there exists $y^* \in Y^*$ so that $x^* = T^*y^*$, that is,

$$\langle x, x^* \rangle = \langle x, T^*y^* \rangle = \langle Tx, y^* \rangle \quad \text{for all } x \in X\,.$$

Define y^* on $\operatorname{ran}(T)$ by

$$\langle Tx, y^* \rangle = \langle x, x^* \rangle \text{ for all } x \in X\,.$$

One can show that y^* is well-defined since $x^* \in \ker(T)^{\perp}$, and y^* is bounded linear functional on $\operatorname{ran}(T)$. The desired result follows fromH-B extension theorem.

(b) $\Rightarrow$ (c). Trival.

(c) $\Rightarrow$ (a). Without loss of generality, let $\overline{\operatorname{ran}(T)} = Y$, or we let $Z = \overline{\operatorname{ran}(T)}$ and let $S \in \mathcal{B}(X,Z)$ with $Sx = Tx$ for each x, then $\operatorname{ran}(S) = \operatorname{ran}(T)$. By H-B extension theorem, one can see that $\operatorname{ran}(S^*) = \operatorname{ran}(T^*)$.

Since $\operatorname{ran}(T)^{\perp} = \ker(T^*)$, ${}^{\perp}\ker(T^*) = \overline{\operatorname{ran}(T)} = Y$, hence

$$\ker(T^*) = \{0\}\,.$$

Thus by Theorem 4.51, $\operatorname{ran}(T) = Y$. $\square$

4.10 Compact Operators

Chapter 5

Operators on Hilbert Space

A large area of current research interest is centered around the theory of operators on Hilbert space. Several other chapters in this book will be devoted to this topic.

There is a marked contrast here between Hilbert spaces and the Banach spaces that are studied in the next chapter. Essentially all of the information about the geometry of Hilbert space is contained in the preceding chapter. The geometry of Banach space lies in darkness and has attracted the attention of many talented research mathematicians. However, the theory of linear operators (linear transformations) on a Banach space has very few general results, whereas Hilbert space operators have an elegant and well-developed general theory. Indeed, the reason for this dichotomy is related to the opposite status of the geometric considerations. Questions concerning operators on Hilbert space don't necessitate or imply any geometric difficulties.

In addition to the fundamentals of operators, this chapter will also present an interesting application to qifferential equations in Section 6.

5.1 The Adjoint of an Operator

Definition 5.1. If H and K are Hilbert spaces, a function $u : H \times K \to \mathbb{F}$ is a **sesquilinear form** if for any h, g in H, k, f in K, and α, β in $\mathbb{F}$,

(a) $u(\alpha h + \beta g, k) = \alpha u(h, k) + \beta u(g, k)$.

(b) $u(h, \alpha k + \beta f) = \bar{\alpha} u(h, k) + \bar{\beta} u(h, f)$.

The prefix "sesqui" is used because the function is linear in one variable but (for $\mathbb{F} = \mathbb{C}$) only conjugate linear in the other. "Sesqui" means "one-and -a-half."

A sesquilinear form is **bounded** if there is a constant M such that $|u(h, k)| \leqslant M\|h\|\|k\|$ for all h in H and k in K. When u is bounded, we say

$$\|u\| \coloneqq \sup_{\|h\|\leqslant 1, \|k\|\leqslant 1} |u(h, k)| \tag{5.1}$$

is the **norm** of u, and u is bounded if and only if u is continous. In fact, "if" part is clearly, if u is not bounded, for any n, there exists $h_n \in H$ and $k_n \in K$ such that

$$\|h_n\| = \|k_n\| = 1\,, |u(h_n, k_n)| \geqslant n\,.$$

Then h_n/n and k_n/n tend to 0, but $|u(h_n/n, k_n/n)| \geqslant 1$ is a contradiction.

Example 5.1. Sesquilinear forms are used to study operators.

(a) If $A \in \mathcal{B}(H, K)$, then $u(h, k) \coloneqq \langle Ah, k\rangle$ is a bounded sesquilinear form, and $\|u\| = \|A\|$.

(b) Also, if $B \in \mathcal{B}(K, H)$ $u(h, k) \coloneqq \langle h, Bk\rangle$ is a bounded sesquilinear form, and $\|u\| = \|B\|$.

A natural question is that are there any more bounded sesquilinear form? Are these two forms related?

Theorem 5.1. *$u : H \times K \to \mathbb{F}$ is a bounded sesquilinear form. Then there are unique operators $A \in \mathcal{B}(H, K)$ and $B \in \mathcal{B}(K, H)$ such that*

$$u(h, k) = \langle Ah, k\rangle = \langle h, Bk\rangle \tag{5.2}$$

for all $h \in H$ and $k \in K$ and $\|u\| = \|A\| = \|B\|$.

Proof. Only the existence of A will be shown. For each h in H, define a functonal on K, namely L_h, by

$$L_h(k) = \overline{u(h, k)}\,, \text{ for all } k \in K\,.$$

Then L_h is linear and,

$$|L_h(k)| \leqslant \|u\|\|h\|\|k\|.$$

By the Riesz representation theorem, there is a unique vector f_h in H such that $\langle k, f_h\rangle = L_h(k) = \overline{u(h, k)}$ and $\|f_h\| \leqslant \|u\|\|h\|$.

Define $A : H \to K$; $h \mapsto f_h$. A is linear, using the uniqueness part of the Riesz representation theorem. Also, $\langle Ah, k\rangle = \overline{\langle k, Ah\rangle} = \overline{\langle k, f_h\rangle} = u(h, k)$. From Example 5.1 we have $\|A\| = \|u\|$.

If $A_1 \in \mathcal{B}(H, K)$ and $u(h, k) = \langle A_1 h, k\rangle$, then $\langle Ah - A_1 h, k\rangle = 0$ for all k, thus $Ah - A_1 h = 0$ for all h. So, A is unique. $\square$

Definition 5.2. If $A \in \mathcal{B}(H, K)$, then the unique operator B in $\mathcal{B}(K, H)$ satisfying (5.2) is called the **adjoint** of A, and is denoted by $B = A^*$.

Remark 5.1. we have defined the adjoint of A before, by $A' : K^* \to H^*$, $k^* \mapsto k^* \circ A$. In fact, defice $\phi_H : H \to H^*$, $h \mapsto \langle \cdot, h\rangle$ and the same is ϕ_K, then the following diagram is commutative.

Example 5.2. P is a projection operater on H, then $P^* = P$.

Example 5.3. U is an isometric isomorphism between H and K , then $U^* = U^{-1}$.

$$
\begin{tikzcd}
H \arrow[r, "\varphi_H"] & H^{\star} \\
K \arrow[u, "A^{\star}"] \arrow[r, "\varphi_K"] & K^{\star} \arrow[u, "A'"']
\end{tikzcd}
$$

Figure 5.1: Two "adjoints"

From now on we will examine and prove results for the adjoint of operators in $\mathcal{B}(H)$. Often, as in the next proposition, there are analogous results for the adjoint of oerators in $\mathcal{B}(H,K)$. This simplification is justified, however, by the cleaner statements that result. Also, the intersted reader will have no trouble formulating the more genéral statement when it is needed. The proof of the next proposition is left as an exercise.

Proposition 5.2. *If $A, B \in \mathcal{B}(H)$ and $\alpha, \beta \in \mathbb{F}$, then:*

(a) $(\alpha A + B)^* = \bar{\alpha} A^* + B^*$.

(b) $(AB)^* = B^* A^*$.

(c) $A^{**} := (A^*)^* = A$.

(d) *A is invertible iff A^* is invertible and,* $(A^*)^{-1} = \left(A^{-1}\right)^*$.

In the end of the section, We conclude with a very important, though easily proved, result.

Theorem 5.3. *If $A \in \mathcal{B}(H)$, then*

$$(\operatorname{ran} A)^{\perp} = \ker A^* \ \textit{and}\ \overline{\operatorname{ran} A} = (\ker A^*)^{\perp} \ . \tag{5.3}$$

Proof. $h \in \ker A \Leftrightarrow Ah = 0 \Leftrightarrow \langle Ah, h' \rangle = 0$ for any $h' \in H \Leftrightarrow \langle h, A^* h' \rangle = 0$ for any $h' \in H \Leftrightarrow h \in (\operatorname{ran} A^*)^{\perp}$. So $\ker A = (\operatorname{ran} A^*)^{\perp}$. Thus $(\ker A)^{\perp} = (\operatorname{ran} A^*)^{\perp\perp} = \overline{\operatorname{ran} A^*}$. Note that $A^{**} = A$, so $\overline{\operatorname{ran} A} = (\ker A^*)^{\perp}$. $\square$

Let $A = A^*$ then we have $(\operatorname{ran} A^*)^\perp = \ker A$ and $(\ker A)^\perp = \overline{\operatorname{ran} A^*}$.

Corollary 5.4. *$A \in \mathcal{B}(H)$, then $H = \ker A \oplus \overline{\operatorname{ran} A}$.*

Examples

Example 5.4. If an operator on $\mathbb{F}^d$ is presented by a matrix, then its adjoint is represented by the conjuagate transpose of the matrix.

Example 5.5. Let (X, Ω, μ) be a σ-finite measure space and let M_ϕ be the multiplication operator with symbol ϕ. Then M_ϕ^* is $M_{\overline{\phi}}$, the multiplication operator with symbol $\overline{\phi}$.

Example 5.6. Let (X, Ω, μ) be a σ-finite measure space and K is the integral operator with kernel $k(x, y)$, then K^* is the integral operator with kernel $k^*(x, y) \equiv \overline{k(y, x)}$.

Example 5.7. If $S : l^2 \to l^2$ is defined by $S\,(\alpha_1, \alpha_2, \ldots) = (0, \alpha_1, \alpha_2, \ldots)$, which is called the *unilateral shift*, then S is an isometry and $S^*\,(\alpha_1, \alpha_2, \ldots) = (\alpha_2, \alpha_3, \ldots)$, and it is called the *backward shift*.

The operation of taking the adjoint of an operator is, as the reader may have seen from the examples above, analogous to taking the conjugate of a complex number. It is good to keep the analogy in mind, but do not become too religious about it.

5.2 Three operators on Hilbert space

Definition 5.3. H is a Hilbert space, $A \in \mathcal{B}(H)$.

- A is called **self-adjoint** if $A^* = A$.
- A is called **normal** if $AA^* = A^*A$.
- A is called **unitary** if A is isometric isomorphism.

In the analogy between the adjoint and the complex conjugate, self-adjoint operators become the analogues of real numbers and, unitaries are the analogues of complex numbers of modulus 1 . Normal operators, as we shall see, are the true analogues of complex numbers. Notice that self-adjoint and unitary operators are normal.

Example 5.8.

- Every multiplication operator M_ϕ is normal; M_ϕ is self-adjoint if and only if ϕ is real-valued; M_ϕ is unitary if and only if $|\phi| = 1$ μ-a.e.
- An integral operator K with kernel k is self-adjoint if and only if $k(x,y) = \overline{k(y,x)}$, $\mu \times \mu$- a.e..
- Obviously, the unilateral shift is not normal.

Self-adjoint operator

Definition 5.4. A bivariate function $u : H \times H \to \mathbb{F}$ is called **heimitian** if, $u(h_1, h_2) = \overline{u(h_2, h_1)}$, for any $h_1, h_2 \in H$.

From Theorem 5.1 we konw there is a one-to-one correspondence between continuous sesquiliear hermitian form and self-adjoint operators. In fact, if u is a continuous sesquiliear hermitian form, then there exists unique $A \in \mathcal{B}(H)$, which is self-adjoint, so that

$$u(h_1, h_2) = \langle Ah_1, h_2 \rangle .$$

Lemma 5.5 (Polar indenity)**.** *u is a sesquiliear form on $\mathbb{C}$-Hilbert space H. Then the following "polar indentity" holds for any $h_1, h_2 \in H$.*

$$
\begin{aligned}
u(h_1, h_2) = &\frac{u(h_1+h_2, h_1+h_2) - u(h_1-h_2, h_1-h_2)}{4} \\
&+ i \times \frac{u(h_1+ih_2, h_1+ih_2) - u(h_1-ih_2, h_1-ih_2)}{4}\,.
\end{aligned}
\tag{5.4}
$$

Remark 5.2. If u is a sesquiliear hermitian form on $\mathbb{F}$-Hilbert space H, the "polar indentity" holds without doubt.

Proposition 5.6. *u is a sesquilinear form on $\mathbb{C}$-Hilbert space H. Then u is hermitian if and only if*

$$u(h, h) \in \mathbb{R}\,, \quad \textit{for all } h \in H\,. \tag{5.5}$$

Particularly, $A \in \mathcal{B}(H)$, then A is self-adjoint if and only if $\langle Ah, h\rangle \in \mathbb{R}$ for any h in H.

Proof. Using the "polar identity", it's easy to check the proposition. □

Example 5.9 (Counterexample)**.** The preceing proposition is false if it is only assumed that H is an $\mathbb{R}$-Hilbert space. For example, if $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ on $\mathbb{R}^2$, then $\langle Ah, h\rangle = 0$ for all h in $\mathbb{R}^2$. However, A^* is the transpose of A and so $A^* \neq A$. Indeed, for any operator A on an $\mathbb{R}$-Hilbert space, $\langle Ah_1, h_2\rangle \in \mathbb{R}$.

Theorem 5.7. *$A \in \mathcal{B}(H)$ and A is self-adjoint. Then*

$$\|A\| = \sup_{\|h\|=1} |\langle Ah, h\rangle|\,. \tag{5.6}$$

Proof. Let $u(h_1, h_2) = \langle Ah_1, h_2\rangle$ and $v(h_1) = u(h_1, h_1)$ for any $h_1, h_2 \in H$. We only need to prove that

$$M := \sup_{\|h\|=1} |v(h)| \geqslant \|A\| = \|u\|\,.$$

For any $h_1, h_2 \in H$ such that $\|h_1\|, \|h_2\| \leqslant 1$. Using "polar identity" we konw

$$\operatorname{Re} u(h_1, h_2) = \frac{v(h_1 + h_2) - v(h_1 - h_2)}{4}$$

So

$$|\operatorname{Re} u(h_1, h_2)| \leqslant \frac{M}{4}(\|h_1 + h_2\|^2 + \|h_1 - h_2\|^2) = \frac{M}{2}(\|h_1\|^2 + \|h_2\|^2) \leqslant M .$$

Let $u(h_1, h_2) = |u(h_1, h_2)|e^{i\theta}$. So $u(e^{-i\theta}h_1, h_2) = |u(h_2, h_2)|$ is real. Thus,

$$|u(h_1, h_2)| = u(e^{-i\theta}h_1, h_2) = \operatorname{Re} u(e^{-i\theta}h_1, h_2) \leqslant M .$$

Therefore, $\|A\| = \|u\| \leqslant M$. □

Corollary 5.8. *$A \in \mathcal{B}(H)$ is self-adjoint.If $\langle Ah, h\rangle = 0$ for all h, then $A = 0$.*

Normal operator

Theorem 5.9. *$A \in \mathcal{B}(H)$, then A is normal if and only if*

$$\|Ah\| = \|A^*h\| \ \text{ for all} h \in H. \tag{5.7}$$

If H is a $\mathbb{C}$-Hilbert space and $A \in \mathcal{B}(H)$, then let

$$B = \frac{A + A^*}{2}\ , \text{ and } C = \frac{A - A^*}{2i}\ . \tag{5.8}$$

Then $B < C$ are self-adjoint and,

$$A = B + iC\,.$$

The operators B and C are called, respectively, the real and imaginary parts of A.

Proposition 5.10. *H is a $\mathbb{C}$-Hilbert space. $A \in \mathcal{B}(H)$. A is normal if and only if the real and imaginary parts of A commute.*

Unitary

Proposition 5.11. *$A \in \mathcal{B}(H)$,the following statement are equivalent.*

(a) *A is an isometry*

(b) *$A^*A = I$.*

(c) *$\langle Ah_1, Ah_2 \rangle = \langle h_1, h_2 \rangle$ for all h_1, h_2 in H.*

Proposition 5.12. *$A \in \mathcal{B}(H)$,the following statement are equivalent.*

(a) *A is unitary. (That is, A is a surjective isometry.)*

(b) *$A^*A = AA^* = I$.*

(c) *A is a normal isometry.*

5.3 Projections and Idempotents; Invariant and Reducing Subspaces

Let M be any closed subspace of H. We have given the definiton of P_M : projection of H onto M in the preceding chapter. When we say a operator P is projection, we means that there exists $M \leqslant H$ such that $P = P_M$. Obviously, We have $\operatorname{ran}(P) = M$ and $\ker(P) = M^{\perp}$. If P is a projection, we konw $P^2 = P$. In some sence, this property characterizes "projection" which is not orthogonal but skewed.

Definition 5.5. An **idempotent** on H is a bounded linear operator E on H such that $E^2 = E$.

It is not difficult to construct an idempotent that is not a projection.

Example 5.10. Let H be the two-dimensional real Hilbert space $\mathbb{R}^2$, let

$$M = \left\{(x, 0) \in \mathbb{R}^2 : x \in \mathbb{R}\right\},\ N = \{(x, x\tan\theta) : x \in \mathbb{R}\},$$

where $0 < \theta < \frac{1}{2}\pi$. There is an idempotent E_θ with $\operatorname{ran} E_\theta = M$ and $\ker E_\theta = N$. We can compute that $\|E_\theta\| = (\sin\theta)^{-1} > 1$. So E_θ must not be a projection.

Proposition 5.13.

(a) *E is an idempotent if and only if $I - E$ is an idempotent.*

(b) $\operatorname{ran} E = \ker(I - E), \ker E = \operatorname{ran}(I - E)$ *and both ran E and ker E are closed subspaces of H.*

(c) $H = \operatorname{ran} E \oplus \ker(E)$.

Proof. Notice that $(I - E)^2 = I - 2E + E^2 = I - 2E + E = I - E$, thus $I - E$ is also an idempotent. E be any idempotent.

Since $E, I - E$ is continuous, $\ker(E)$ is a closed subspace of H. Also, $0 = (I - E)h = h - Eh$, if and only if $Eh = h$. So ran $E \supset \ker(I - E)$. On the other hand, if $h \in \operatorname{ran} E, h = Eg$ and so $Eh = E^2 g = Eg = h$, hence $\operatorname{ran} E = \ker(I - E)$. Similarly, $\operatorname{ran}(I - E) = \ker E$. $\square$

There is also a converse to 3. If $M, N \leqslant H, M \cap N = \{0\}$, and$M + N = H$, then there is an idempotent E such that $\operatorname{ran} E = M$ and $\ker E = N$. Moreover, E is unique. When proving this converse, we need to use *closed graph theorem* to show that E is bounded. The same fact is true in Banach spaces.

Now we turn our attention to projections, which are peculiar to Hilbert space. A natural question is what conditions are given can make an idempotent become a projeection?

Proposition 5.14. *If E is an idempotent on H and $E \neq 0$,the following statements are equivalent.*

(a) *E is a projection.*

(b) $\operatorname{ran} E \perp \ker E$.

(c) $\|E\| = 1$.

(d) *E is hermitian.*

(e) *E is normal.*

(f) $\langle Eh, h \rangle \geqslant 0$ *for all h in H.*

Proof. $3 \Rightarrow 1$. For any $x \in \operatorname{ran} E$ and $y \in \ker E$, we have $\|E(x + y)\| = \|x\| \leqslant \|x + y\|$, thus $Re\langle x, y \rangle \geqslant -\frac{1}{2}\|y\|^2$.Take any $y' \in H$ and $t > 0$, then $Re\langle x, y' \rangle \geqslant -\frac{t}{2}\|y'\|^2$. Let $t \to 0$, we have $Re\langle x, y' \rangle \geqslant 0$. From this we konw $Re\langle x, y' \rangle = 0$, then $\langle x, y' \rangle = 0$.

$5 \Rightarrow 1$. E is normal then $\|Eh\| = \|E^* h\|$ for every h. Hence $\ker E = \ker E^*$. But by Theorem 5.3, $(\operatorname{ran} E)^\perp = \ker E^*$, so E is a projection.

$6 \Rightarrow 1$. For any $x \in \operatorname{ran} E$ and $y \in \ker E$, we have $\langle E(x + y), x + y \rangle = \langle x, x + y \rangle \geqslant 0$, then $\langle x, y \rangle \geqslant -\|x\|^2$. Let $x = tx'$ where $t > 0$ and $x' \in H$. $\square$

投影算子的运算

Proposition 5.15. *Let* $\{M_i : i \in I\}$ *be a collection of closed subspaces of* H*. Then* $\cap \left\{M_i^{\perp} : i \in I\right\} = (\text{cspan}\,\{M_i : i \in I\})^{\perp}$ *and* $(\cap \{M_i : i \in I\})^{\perp} = \text{cspan}\left\{M_i^{\perp} : i \in I\right\}$*.*

Proposition 5.16. *Let* P *and* Q *be projections. Then* $P+Q$ *is a projection if and only if* $\operatorname{ran} P \perp \operatorname{ran} Q$*. If* $P+Q$ *is a projection, then* $\operatorname{ran}(P+Q) = \operatorname{ran} P \oplus \operatorname{ran} Q$ *and* $\ker(P+Q) = \ker P \cap \ker Q$*.*

Proposition 5.17. *Let* P_i *be a collection of projections such that* $P_iP_j = 0$ *if* $i \neq j$*. Then there exists a projection* P *such that* $P = \sum_{i\in I} P_i$ (s)*, i.e.,* $Ph = \sum_{i\in I} P_i h$ *for any* $h \in H$*. If we denote* $M_i = \operatorname{ran} P_i$*, and*

$$\bigoplus_{i\in I} M_i \coloneqq \left\{ \sum_{i\in I} x_i : x_i \in M_i,\ \sum_{i\in I} \|x_i\|^2 < \infty \right\}$$

then $\operatorname{ran} P = \bigoplus_{i\in I} M_i$*, or* $P = P_{\oplus_{i\in I} M_i}$*.*

Proposition 5.18. *Let* P *and* Q *be projections,* PQ *is a projection if and only if* $PQ = QP$*. If* PQ *is a projection, then* $\operatorname{ran} PQ = \operatorname{ran} P \cap \operatorname{ran} Q$ *and* $\ker PQ = \ker P + \ker Q$*.*

Definition 5.6. If H and N are two closed linear subspaces of H. Define

$$M \ominus N \equiv M \cap N^{\perp}$$

which is called the **orthogonal difference** of M and N.

Proposition 5.19. *If* P *and* Q *are projections, then the following statements are equivalent.*

(a) $P - Q$ *is a projection.*

(b) $\operatorname{ran} Q \subseteq \operatorname{ran} P$*.*

(c) $PQ = Q$.

(d) $QP = Q$.

If $P-Q$ *is a projection, then* $\operatorname{ran}(P-Q) = (\operatorname{ran} P) \ominus (\operatorname{ran} Q)$ *and* $\ker(P-Q) = \operatorname{ran} Q \oplus \ker P$.

Exercise 5.1. $\{P_n\}$ are projections and is monotonous, i.e. $P_n \leqslant P_{n+1}$ for all n or $P_n \geqslant P_{n+1}$ for all n, then there exists a projection P such that $P_n \to P$ (s).

Invariant and Reducing Subspaces

Definition 5.7. Let $A \in \mathcal{B}(H)$ and $M \leqslant H$, say that M is an **invariant subspace for** A if $Ah \in M$ whenever $h \in M$, i.e., $AM \subseteq M$. Say that M is a **reducing subspace for** A if $AM \subseteq M$ and $AM^{\perp} \subseteq M^{\perp}$

Proposition 5.20. *If* $A \in \mathcal{B}(H), M \leqslant H$, *then* M *is invariant for* A *if and only if* $P_M A P_M = A P_M$.

Proof. □

Proposition 5.21. *If* $A \in \mathcal{B}(H), M \leqslant H$, *then*

(a) M *is a reducing subspace for* A *.*

(b) $P_M A = A P_M$.

(c) M *is imvariant for both* A *and* A^*.

Proof. $3 \Rightarrow 1$. If $h \in M^{\perp}$ and $g \in M$, then $\langle g, Ah \rangle = \langle A^* g, h \rangle = 0$ since $A^* g \in M$ since g was an arbitrary vector in $M, Ah \in M^{\perp}$. That is, $AM^{\perp} \subseteq M^{\perp}$. □

Remark 5.3. If $M \leqslant H$, then $H = M \oplus M^{\perp}$. If $A \in \mathcal{B}(H)$, then A can be written as a 2×2 matrix with operator entries,

$$A = \begin{bmatrix} W & X \\ Y & Z \end{bmatrix}$$

where $W \in \mathcal{B}(M), X \in \mathcal{B}\left(M^{\perp}, M\right), Y \in \mathcal{B}\left(M, M^{\perp}\right)$, and $Z \in \mathcal{B}\left(M^{\perp}\right)$. Then M is invariant for A if and only if $Y = 0$, $W = A|_M$. M reduces A if and only if $Y = Z = 0$ and $W = A|_M$, $Z = A|_{M^{\perp}}$. This is the reason for the terminology.

Appendix A

A.1 Metric space

Compact Subset of a Metric Space

Definition A.1. Let A be a subset of a metric space (X, d). we say A is **bounded** if A is contained in a ball of finite radius, i.e. there exists some $x \in X$ and $r > 0$ such that $A \subset B(x, r)$.

Definition A.2. A is a subset of a metric space (X, d) and $\varepsilon > 0$. A subset $F_\varepsilon \subset X$ is called an ε-**net** for A if each $x \in A$ there is an element $y \in F_\varepsilon$ such that $d(x, y) < \varepsilon$.

Definition A.3. A subset A of a metric space (X, d) is **totally bounded** if for any $\varepsilon > 0$ there is a *finite* ε-net $F_\varepsilon \subset X$ for A. That is, there is a finite set $F_\varepsilon \subset X$ such that

$$A \subset \bigcup_{x \in F_\varepsilon} B(x, \varepsilon).$$

Remark A.1. A subset A of a metric space (X, d) is totally bounded if and only if for any $\varepsilon > 0$ there is a finite ε-net $F_\varepsilon \subset A$ for A.

Obviously, every toally bounded set of a metric space is bounded. The following examples shows that boundedness does not, in general, imply total boundedness.

Example A.1. Let $X = \ell_2$ and $B = B(0, 1) = \{x \in X : \|x\| \leqslant 1\}$. B is bounded but not totally bounded.

Theorem A.1. *A subset K of a metric space (X, d) is totally bounded if and only if every sequence in K has a Cauchy subsequence.*

Proof. Assume that K is totally bounded and let (x_n) be an infinite sequence in K. There is a finite set of points $\{y_{11}, y_{12}, \ldots, y_{1r}\}$ in K such that

$$K \subset \bigcup_{j=1}^{r} B\left(y_{1j}, \frac{1}{2}\right).$$

At least one of the balls $B\left(y_{1j}, \frac{1}{2}\right), j = 1, 2, \ldots, r$, contains an infinite subsequence (x_{n1}) of (x_n). Again, there is a finite set $\{y_{21}, y_{22}, \ldots, y_{2s}\}$ in K such that

$$K \subset \bigcup_{j=1}^{s} B\left(y_{2j}, \frac{1}{2^2}\right).$$

At least one of the balls $B\left(y_{2j}, \frac{1}{2^2}\right), j = 1, 2, \ldots, s$, contains an infinite subsequence (x_{n2}) of (x_{n1}) Continuing in this way, at the m-th step, we obtain a subsequence (x_{nm}) of $\left(x_{n(m-1)}\right)$ which is contained in a ball of the form $B\left(y_{mj}, \frac{1}{2^m}\right)$.

Claim: The diagonal subsequence (x_{nn}) of (x_n) is Cauchy.

Indeed, if $m > n$, then both x_{nn} and x_{mm} are in the ball of radius 2^{-n}. Hence, by the triangle inequality,

$$\|x_{nn} - x_{mm}\| < 2^{1-n} \to 0 \text{ as } n \to \infty.$$

Conversely, assume that every sequence in K has a Cauchy subsequence and that K is not totally bounded. Then, for some $\epsilon > 0$, no finite ϵ-net exists for K. Hence, if $x_1 \in K$, then there is an $x_2 \in K$ such that $\|x_1 - x_2\| \geq \epsilon$. Otherwise, $\|x_1 - y\| < \epsilon$ for all $y \in K$ and consequently $\{x_1\}$ is a finite ϵ-net for K, a contradiction.) Similarly, there is an $x_3 \in K$ such that

$$\|x_1 - x_3\| \geq \epsilon \text{ and } \|x_2 - x_3\| \geq \epsilon$$

Continuing in this way, we obtain a sequence (x_n) in K such that $\|x_n - x_m\| \geq \epsilon$ for all $m \neq n$. Therefore (x_n) cannot have a Cauchy subsequence, a contradiction. □

Definition A.4. A metric space (X, d) is called **sequentially compact** if every sequence in X has a convergent subsequence, a subset K of X is sequentially compact when (K, d) is sequentially compact.

Theorem A.2. *A subset of a metric space is sequentially compact if and only if it is totally bounded and complete.*

Proof. Let K be a sequentially compact subset of a normed linear space (X, d), and (x_n) be a sequence in K. By sequential compactress of K, (x_n) has a subsequence (x_{n_k}) which converges in K. since every convergent sequence is Cauchy, by Theorem A.1, K is totally bounded.

Conversely, assume that K is a totally bounded and complete subset of a normed linear space (X, d). Let (x_n) be a sequence in K. By Theorem A.1, (x_n) has a Cauchy subsequence (x_{n_k}) . since K is complete, (x_{n_k}) converges in K. Hence K is sequentially compact. □

Definition A.5. A metric space is called **compact** if each of its open covers has a finite subcover, a subset K of X is compact when (K, d) is compact.

Theorem A.3. *(X, d) is a metric space and $K \subset X$. Then K is compact if and only if K is sequentially compact.*

Proof. □

Baire's Category Theorem

Definition A.6. A subset S of a metric space (X, d) is called **nowhere dense** in X if the closure of S contains no interior points.

Theorem A.4. *Let (X, d) be a complete metric space.*

(a) If (G_n) is a sequence of nonempty, open and dense subsets of X then $G = \cap_{n\in\mathbb{N}} G_n$ is dense in X.

(b) If (F_n) is a sequence of closed, nowhere dense subsets of X, then $F = \cup_{n\in\mathbb{N}} F_n$ cotains no interior points.

Proof. We ony need wo show 1. Let $x \in X$ and $\epsilon > 0$. Since G_1 is dense in X, there is a point x_1 in the open set $G_1 \cap B(x, \epsilon)$. Let r_1 be a number such that $0 < r_1 < \frac{\epsilon}{2}$ and $B(x_1, r_1) \subset G_1 \cap B(x, \epsilon)$. By induction, we obtain a sequence (x_n) in X and a sequence (r_n) of radii such that for each n, $0 < r_n < \frac{\epsilon}{2^n}$, and

$$\overline{B(x_{n+1}, r_{n+1})} \subset G_{n+1} \cap B(x_n, r_n) \text{ and } \overline{B(x_1, r_1)} \subset G_1 \cap B(x, \epsilon)$$

Hence, (x_n) is a Cauchy sequence in X. since X is complete, there is a $y \in X$ such that $x_n \to y$ as $n \to \infty$. since x_k lies in the closed set $\overline{B(x_n, r_n)}$ if $k > n$, it follows that y lies in each $\overline{B(x_n, r_n)}$. Hence y lies in each G_n. That is, $G = \bigcap_{n \in \mathbb{N}} G_n \neq \varnothing$. It is also clear that $y \in B(x, \epsilon)$. $\square$

A subset S of a metric space (X, d) is said to be

(a) of **first category** or **meagre** in X if S can be written as a countable union of sets which are nowhere dense in X.

(b) of **second category** or **nonmeagre** in X if it is not of first category in X.

Theorem A.5 (Baire's Catagory Theorem)**.** *A complete metric space* (X, d) *is of second category in itself.*

A.2 不可数级数的和

Definition A.7. X 是一赋范线性空间, $\{\alpha_i\}_{i\in I}$ 是 X 中一族向量, 记 I 的有限子集全体为 $\mathcal{I}$. 容易知道 $(\mathcal{I},\subset)$ 是一定向集, 故 $\{\sum_{i\in S}\alpha_i : S\in\mathcal{I}\}$ 是 X 中的网, 若网 (依范数诱导的拓扑) 收敛 β , 就称级数 $\sum_{i\in I}\alpha_i$ 收敛 (有时也称为无条件收敛,unconditionally convergent) 于 β, 记为 $\beta=\sum_{i\in I}\alpha_i$.

Remark A.2. 显然, $\sum_{i\in I}\alpha_i$ 收敛于 β 即是: 对任意 $\epsilon>0$, 存在 I 的有限子集 S, 使得 I 任何包含 S 的有限子集 T , 有

$$\|\sum_{i\in T}\alpha_i-\beta\|\leqslant\epsilon$$

Proposition A.6. *X 是一赋范线性空间,$\alpha_i,\beta_i\in X(i\in I)$, 且 $\sum_{i\in I}\alpha_i$, $\sum_{i\in I}\beta_i$ 皆收敛. 则*

(a) $\sum_{i\in I}(\alpha_i+\beta_i)$ 收敛, 且 $\sum_{i\in I}(\alpha_i+\beta_i)=\sum_{i\in I}\alpha_i+\sum_{i\in I}\beta_i$.

(b) 任意 $\lambda\in\mathbb{F}$, $\sum_{i\in I}\lambda\alpha_i$ 收敛, 且 $\sum_{i\in I}\lambda\alpha_i=\lambda\sum_{i\in I}\alpha_i$.

(c) A 是 X 到一赋范线性空间 Y 的有界线性算子, 则 $\sum_{i\in I}A\alpha_i$ 是 Y 收敛的级数, 且 $\sum_{i\in I}A\alpha_i=A(\sum_{i\in I}\alpha_i)$.

Proof. 利用定义易证. □

Proposition A.7. *X 是赋范线性空间,$\alpha_i\in X,i\in I$, 且 $\sum_{i\in I}\alpha_i$ 收敛. 则对任意 $\epsilon>0$, 存在 I 的有限子集 S, 使得 I 任何与 S 不交的有限子集 J, 有*

$$\|\sum_{i\in J}\alpha_i\|\leqslant\epsilon$$

当 X 是 Banach 空间时, 逆命题也成立.

Proof. 逆命题证明时, 构造柯西列利用空间完备性. □

Corollary A.8. *$\sum_{i\in I}\alpha_i$ 收敛, 则只有至多可数项 α_i 非零.*

Proof. 利用上述命题, 对任何 n, 有 $S_n \in \mathcal{I}$ 对任何与 S_n 不交的 J, 有

$$\|\sum_{i \in J} \alpha_i\| \leqslant \frac{1}{n}$$

令 $S = \bigcup_{n=1}^{\infty} S_n$. 可见只有 $i \notin S$ 时必然有 $\alpha_i = 0$. $\square$

Proposition A.9. X 是一赋范线性空间,$\alpha_i \in X(i \in I)$, 且 I 是可数集. 则 $\sum_{i \in I} \alpha_i$ 收敛于 β 当且仅当对任意 $\mathbb{N}$ 到 I 的双射 f, 级数 $\sum_{k=1}^{\infty} \alpha_{f(k)}$ 皆收敛于 β .

Proof. 充分性用反证法构造矛盾, 必要性易证 $\square$

Proposition A.10. $x_i \in \mathbb{R}$, 且 $x_i \geqslant 0 \, i \in I$. 则 $\sum_{i \in I} x_i$ 收敛于 $y \in \mathbb{R}$ 当且仅当

$$y = \sup\{\sum_{i \in S} x_i : S \text{ 是} I \text{ 有限子集}\}.$$