Functional Analysis

泛函分析笔记

何憾[1]

July 9, 2020

[1]访问我的个人主页, 可查看并下载其他笔记.

Contents

0.1 Baisc Linear Algebra

Let X and Y be linear spaces over the same scalar field $\mathbb{F}$. A mapping T defined on a linear subspace $D = D(T)$ of X and taking values in Y is said to be *linear*, if

$$T(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 T x_1 + \alpha_2 T x_2, \text{ for all } x_1, x_2 \in D \text{ and all } \alpha, \beta \in \mathbb{F}.$$

T is also called a *linear operator* or *linear transformation* on $D(T) \subset X$ into Y. If T is taking values in the scalar field, i.e., $Y = \mathbb{F}$, we say T is a *linear functional* on $D(T)$. We say that $D(T)$ is the *domain* of T and sometimes is denoted by $\operatorname{dom}(T)$. Moreover, the *range* the of T, denoted by $R(T)$ or $\operatorname{ran}(T)$, is given by

$$R(T) \coloneqq T(X) = \{y \in Y : y = Tx \text{ for some } x \in D(T)\}.$$

The *null space* or the *kernel* of T, denoted by $N(T)$ or $\ker(T)$, is given by

$$N(T) \coloneqq T^{-1}\{0\} = \{x \in D(T) : Tx = 0\}.$$

Clearly, $N(T)$ is a linear subspace of X and $R(T)$ is a linear subspace of Y.

If a linear operator T gives a one-to-one map of $D(T)$ onto $R(T)$, then the inverse map T^{-1} gives a linear operator on $R(T)$ onto $D(T)$:

$$T^{-1} T x = x \quad \text{for } x \in D(T) \; ; \; T T^{-1} y = y \quad \text{for } y \in R(T).$$

T^{-1} is the *inverse operator* or, in short, the *inverse* of T. Clearly, T admits the inverse T^{-1} if and only if $N(T) = \{0\}$; i.e., $Tx = 0$ implies $x = 0$.

Let T_1 and T_2 be linear operators with domains $D(T_1)$ and $D(T_2)$ both contained in X, and ranges $R(T_1)$ and $R(T_2)$ both contained in Y. Then $T_1 = T_2$ if $D(T_1) = D(T_2)$ and $T_1 x = T_2 x$ for all $x \in D(T_1) = D(T_2)$. If $D(T_1) \subset D(T_2)$ and $T_1 x = T_2 x$ for all $x \in D(T_1)$, then T_2 is called an *extension* of T_1, and T_1 a *restriction* of T_2; we shall then write $T_2|_{D(T_1)} = T_1$.

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

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

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

Exercise 0.1. Let X be a vector space, and $S, T \in \mathcal{L}(X)$. Suppose that $ST = TS$ and that ST is bijective. Then also S and T are bijective.

Lemma 0.1. *Let X be a vector space over $\mathbb{F}$. Let $\ell, \ell_1, \cdots, \ell_n$ be linear functionals on X. Then $\ell \in \mathrm{span}\{\ell_1, \cdots, \ell_n\}$ if and only if*

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

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

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

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

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

Thus

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

Corollary 0.2. *Let X be a vector space over $\mathbb{F}$. Let f, g be linear functionals on X. Then there exists a scalar $c \in \mathbb{F}$ so that $g = cf$ if and only if $\ker(f) \subset \ker(g)$.*

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: $p(x + y) \leq p(x) + p(y)$ 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 $\|x\| := d(x, 0)$ 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., $x_n \to x$ (with respect to 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 $\lambda_n x_n \to \lambda x$.

Definition 1.2. 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 1.3. 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 1.1. *$(X, \|\cdot\|)$ 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 $(X, \|\cdot\|)$ 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. $\square$

1.1.1 Examples

We give some examples of normed linear spaces and Banach spaces. We always 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

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

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

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

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

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

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 $C_b(X,\mathbb{F})$ be all bounded continuous functions $f : X \to \mathbb{F}$. Define

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

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

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

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

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

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

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

Thus,

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

so that $f \in C_0(X,\mathbb{F})$.

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 $\{f \in C(X_\infty,\mathbb{F}) : f(\infty) = 0\}$, with the norm it inherits as a subspace of $C(X_\infty,\mathbb{F})$, is *isometrically isomorphic* to $C_0(X,\mathbb{F})$. Paricularly,

- $C_0(\mathbb{R},\mathbb{F}) =$ all of the $\mathbb{F}$-valued continuous functions f such that $|f(x)| \to 0$ as $|x| \to \infty$.

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

- Let $\mathbb{N}$ equipped with discrete topology, then $C_0(\mathbb{N}, \mathbb{F}) = c_0$.

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

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

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

Example 1.8. Let I be a set and $1 \leq p < \infty$. Define $\ell^p = \ell^p(I, \mathbb{F})$ 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 $\ell^p(I, \mathbb{F}) = L^p(I, 2^I, \#; \mathbb{F})$.

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

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

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

Example 1.10 (Sobolev space)**.** Let $1 \leq p < \infty$ and $n \in \mathbb{N}$ and let $W_p^n[0,1] = W_p^n([0,1], \mathbb{R})$ be all the functions $f : [0,1] \to \mathbb{R}$ 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.

1.1.2 Separability and 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 1.2. *X is separable if and only if it contains a countable set S such that $\overline{\operatorname{span}}(S) = X$.*

Proof. It suffices to show that if there is a countable subset S of X so that $\overline{\operatorname{span}}(S) = X$, then X is separable. 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 $\operatorname{span}(S)$. Hence A is a countable dense subset of X. $\square$

Example 1.11. $\mathbb{R}, \mathbb{C}$ 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 1.4. 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 $\overline{\operatorname{span}}\{e_n : n \in \mathbb{N}\} = X$ Uniqueness of the expansion clearly implies that the set $\{e_n : n \in \mathbb{N}\}$ is linear independent.

Example 1.12.

- 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 1.3. In 1937, Per Enflo constructed a separeble Banach space with no Schauder basis.

1.1.3 The External Direct Sum of Banach Spaces

Now for the product or (external) 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 vector 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.

Let $\|\cdot\|$ denote the norm on each X_i. For $1 \leqslant p < \infty$, define

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

and

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

If $\{X_1, X_2, \ldots\}$ is a sequence of normed spaces, define

$$\oplus_0 X_n \equiv \left\{ x \in \prod_{n=1}^\infty X_n : \|x(n)\| \to 0 \right\}$$

give $\oplus_0 X_n$ the norm it has as a subspace of $\oplus_\infty X_n$.

Proposition 1.3. *Let $\{X_i\}_{i\in I}$ be a collection of normed spaces and $p \in [1, \infty]$. Denote $X = \oplus_p X_i$, then*

(a) *X is a normed space, and X is a Banach space if and only if each X_i is a Banach space.*

(b) *The canonical projection $P_i : X \to X_i$ is a continuous open linear map with $\|P_i x\| \le \|x\|$ for each x in X.*

Proof. Observe that for $x = (x_i)$ and $y = (y_i)$ in X, by Minkovski inequality,

$$\|x+y\| = \left(\sum_i \|x_i + y_i\|^p\right)^{1/p} \le \left(\sum_i (\|x_i\| + \|y_i\|)^p\right)^{1/p}$$
$$\le \left(\sum_i \|x_i\|^p\right)^{1/p} + \left(\sum_i \|y_i\|^p\right)^{1/p} = \|x\| + \|y\|\,.$$

It's now easy to see that X is a normed vector space. Obviously, if X is complete then each X_i is. On the other hand, suppose each X_i is complete and $\{x^{(n)}\}$ a Cauchy sequence in X. Then for every $\epsilon > 0$, there exists a positive integer $N = N_\epsilon > 0$ so that

$$\|x^{(n)} - x^{(m)}\| = \left(\sum_i \|x_i^{(n)} - x_i^{(m)}\|^p\right)^{1/p} \le \epsilon\,.$$

Thus for each $i \in I$, $\{x_i^{(n)}\}$ is a Cauchy sequence in X_i. Suppose $x_i^{(n)} \to x_i$ in X_i. Given any fixed fininte subset S of I, for $n \ge N_\epsilon$, we have

$$\left(\sum_{i\in S} \|x_i^{(n)} - x_i\|^p\right)^{1/p} = \lim_{m\to\infty} \left(\sum_{i\in S} \|x_i^{(n)} - x_i^{(m)}\|^p\right)^{1/p} \le \epsilon\,.$$

Taking the limit, we get

$$\left(\sum_{i\in I} \|x_i^{(n)} - x_i\|^p\right)^{1/p} \le \epsilon\,, \qquad \text{for } n \ge N_\epsilon\,.$$

Thus $x = (x_i) \in X$ and $x^{(n)} \to x$. So X is a Banach space. Part (b) is trivial, so we omit the proof. □

Remark 1.4. A similar result holds for $\oplus_0 X_n$, but the formulation and proof of this is the same as before.

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 linear operator A on $D(A) \subset X$ into Y is said to be *bounded* if there exists a constant $M > 0$ such that

$$\|Ax\| \le M\|x\|,^{2} \text{ for all } x \in D(A).$$

Obviously, A is bounded if and only if A maps a bounded set into a bounded set. Moreover, boundedness of lienar operator in the sense of Definition 1.5 is equivalent to the continuity. The proof is easy so we omit it.

Proposition 1.4. *Let A be a linear operator on $D(A)$ into Y. Then the following statements are equivalent: (i) A is bounded. (ii) A is continuous on $D(A)$. (iii) A is continuous at some point x_0 in $D(A)$.*

Definition 1.6. Let A be a linear operator on $D(A)$ into Y. The *operator norm* (or simply norm) of A, denoted by $\|A\|$, is defined as

$$\|A\| := \inf\{M : \|Tx\| \le M\|x\|, \ \forall\, x \in D(A)\}\,.$$

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

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

We shall denote by $\mathcal{B}(X, Y)$ the set of all bounded linear operators defined on X taking values in Y. We shall write $\mathcal{B}(X)$ for $\mathcal{B}(X, X)$. 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?

Theorem 1.5. *$\mathcal{B}(X, Y)$ is a Banach space if and only if Y is 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 .

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 $x \leq 1$,

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

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 $\|T_n - T_m\| \leq \|\ell\|\,\|y_n - y_m\|$, 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 1.5. 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 1.4. Let X, Y be normed vector spaces. Let $T \in \mathcal{B}(X,Y), S \in \mathcal{B}(Y,Z)$. Then $ST \in \mathcal{B}(X,Z)$ and $\|ST\| \leq \|S\|\|T\|$.

1.2.1 Examples

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

Example 1.13. 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 $\|A\| = \sup\limits_{1 \le i \le n} \sum\limits_{j=1}^n |\alpha_{ij}|$.

Example 1.14. 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 1.15. 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 1.16 (Multiplication Operator)**.** Let $(X, \mathcal{F}, \mu)$ be a σ-finite measure space and $\phi \in L^\infty(\mu)$. Let $1 \le p \le \infty$. 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 $M_\phi \in \mathcal{B}\left(L^p(\mu)\right)$ 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

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

Thus, $M_\phi \in \mathcal{B}(L^p(\mu))$ and $\|M_\phi\| \le \|\phi\|_\infty$. 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

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

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

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

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

Remark 1.6. 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 1.17 (Integral Operator I)**.** Let $(X, \mathcal{F}, \mu)$ be a measure space. Suppose $k \in L^2(X \times X, \mathcal{F} \times \mathcal{F}, \mu \times \mu)$. Define $K : L^2(\mu) \to L^2(\mu)$ by

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

Then K is a bounded linear operator (in fact K is compact, see Example 7.6) with

$$\|K\| \le \|k\|_{L^2}\,.$$

K 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 $f \in L^2(\mu)$, using the Cauchy-Schwartz inequality, we have

$$\begin{aligned}\|Kf\|^2 &= \int \left| \int k(x,y) f(y) \mu(dy) \right|^2 \mu(dx) \\ &\leq \int \left[\int |k(x,y)|^2 \mu(dy) \int |f(y)|^2 \mu(dy) \right] \mu(dx) = \|k\|_{L^2}^2 \|f\|^2 \,.\end{aligned}$$

Example 1.18 (Integral Operator II)**.** 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

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

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

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

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

This operator is also 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 show this, for $f \in L^p(\mu)$,

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

Hence

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

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

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

Example 1.19. 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

In this section we will show two surprising results. Firstly, there norm on finite dimensional space is unique in the sense of equivalence. Secondly, the completeness of the unit ball implies the dimension of sapce is finite. Now let us make the "equivalence" clear.

Definition 1.7. 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.

Lemma 1.6. *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*

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

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

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

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 surprisingly result told us, this norm is the unique norm on X in the sense of equivalence.

Theorem 1.7. *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

$$C_1|x| \le \|x\| \le C_2|x| .$$

Hence we complete the proof. $\square$

By the theorem above, we can find that

Corollary 1.8. *Every finite-dimensional normed linear space $(X, \|\cdot\|)$ is Banach space.*

Corollary 1.9. *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 1.10. *Let X and Y be normed linear spaces and $\dim(X) < \infty$. Then $\mathcal{B}(X, Y) = \mathcal{L}(X, Y)$.*

Next, we shall 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 1.11 (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*

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

The element y is thus "nearly orthogonal" to M.

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 $u \in M$,

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

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

Second proof. By Corollary 1.34, one can find a non-trivial linear functional ℓ on X which vanishes on M with $\|\ell\| = 1$. By definition of the operator norm $\|\ell\|$ of ℓ, one can find a unit vector x such that $|\ell(x)| \geq 1 - \epsilon$. The claim follows from that $d(x, M) = \|\widetilde{x}\| \geq |\widetilde{\ell}\widetilde{x}| = |\ell(x)| \geq 1 - \epsilon$. $\square$

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

Theorem 1.12. *$(X, \|\cdot\|)$ is finite-dimensional if and only if 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. $\square$

To give an example of a relatively compact infinite subset of a Banach space of infinite dimension, we shall give the

Theorem 1.13 (Ascoli-Arzelá)**.** *Let (S, d) be a compact metric space, and $C(S)$ the Banach space of $\mathbb{F}$-valued continuous functions on S endowed with supremum norm. Then a subset $F \subset C(S)$ is relatively compact in $C(S)$ if the following two conditions are satisfied:*

(a) *F is (uniformly) bounded, i.e.,* $\sup_{x \in F} \|x\| < \infty$*;*

(b) *F is equi-continuous, i.e.,*

$$\lim_{\delta \downarrow 0} \sup_{x \in F} \{|x(s') - x(s'')| : d(s', s'') \leq \delta\} = 0\,.$$

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 $x \in X$,

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

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

Theorem 1.14. *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*

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

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,

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

So the desired result follows. □

Remark 1.7. 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 1.15. *M be a proper 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 coincides with the quotient topology on X/M. In other 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 1.11, $\|Q\| = 1$.

(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. □

Remark 1.8. 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 (with respect to f).

Corollary 1.16. *Let X be a normed vector space. Let M be a closed subspace and N be a finite dimensional subspace. Then $M + N$ is closed.*

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. □

Example 1.20. If $A : X \to Y$ is a linear operator between vector spaces X and Y, then it descends to a injective linear operator from the quotient space $X/N(A)$ to Y. That is,

$$\tilde{A} : X/N(A) \to Y\,;\, \tilde{x} \to Ax$$

It's easy to see that $\tilde{A}$ is well-defined and is an injective lienar operater. Moreover, if X and Y are normed vector sapce and A is continuous, then so is $\tilde{A}$. Indeed,

$$\|\tilde{A}\| = \|A\|. \tag{1.2}$$

On the one hand, $\|\tilde{A}\tilde{x}\| = \|Ax\| \leq \|A\|\|x\|$ for all $x \in \tilde{x}$, and hence $\|\tilde{A}\| \leq \|A\|$. On the other hand, observe that $A = \tilde{A}Q$, where $Q : X \to X/N(A)$ is the quotinent map with $\|Q\| \leq 1$, we get $\|A\| \leq \|\tilde{A}\|\|Q\|$. Then (1.2) follows.

Theorem 1.17. *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 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.

1.5.1 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: $p(x + y) \leq p(x) + p(y)$ for all x, y in X.

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

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

Theorem 1.18. *Let X be a vector space over $\mathbb{R}$. Let p be a positive homogeneous, subaddititive functional on X. Let Y be a subspace in X, and ℓ be a linear functional on Y. If ℓ satisfies*

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

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

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

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

$$L\Big(\sum_i \alpha_i e_i + \sum_j \beta_j \varepsilon_j\Big) = \sum_i \alpha_i \ell\,(e_i) = \ell\Big(\sum_i \alpha_i e_i\Big).$$

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

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

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 $Z = Y \oplus \operatorname{span}\{z\}$. 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 $L(y + tz) = t\alpha_0 + \ell(y) \le p(y + tz)$. Hence

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

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

$$\alpha_0 \le p(y + z) - \ell(y) \tag{1.3}$$

for all y in Y. On the other hand, if α_0 satisfies (1.3), then by reversing the preceding argument, it follows that $t\alpha_0 + \ell(y) \le p(y + tz)$ whenever $t \ge 0$.

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

$$\alpha_0 \ge \ell(y') - p(y' - z)\ . \tag{1.4}$$

Combining (1.3) and (1.4), we see that that α_0 can be chosen satisfying them both simultaneously. Such an α_0 exists iff for all pairs $y, y' \in Y$,

$$\ell(y') - p(y' - z) \le p(y + z) - \ell(y) \tag{1.5}$$

Using the linearity of ℓ and subadditivity of p we have

$$\ell(y + y') \le p(y + y') \le p(y + z) + p(y' - z)\ .$$

So pick α_0 satisfying $\sup_{y' \in Y} \ell(y') - p(-z + y') \le \alpha_0 \le \inf_{y \in Y} p(z + y) - \ell(y)$, 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

$$(Z, \ell) \le (Z', \ell')$$

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 $(Z_v, \ell_v) \le (Z, \widehat{\ell})$ 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 1.19. *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

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

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

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

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

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

Theorem 1.20 (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,*

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

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

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

Proof. There are two case for $\mathbb{F}$. When $\mathbb{F}$ is $\mathbb{R}$, then $\ell(y) \leq |\ell(y)| \leq p(y)$ for y in Y. By Theorem 1.18 , ℓ can be extended to X such that $\ell(x) \leq p(x)$ for all $x \in X$. Hence $-\ell(x) = \ell(-x) \leq p(-x) = p(x)$, That is, $|\ell(x)| \leq p(x)$.

When $\mathbb{F} = \mathbb{C}$. Let $\ell_1 = \operatorname{Re} \ell$. By Lemma 1.19, $\ell_1 \leq p$. From the proof above, ℓ_1 can be extended on X as $\mathbb{R}$-linear functional such that $|\ell_1| \leq p$. Let

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

for all x in X, then ℓ is a extension. By Lemma 1.19, $|\ell| \leq p$. □

1.5.2 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

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

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 $|\lambda| \leq 1$, we have $\lambda K \subset K$.

Example 1.21. Let p be real-valued functional on X.

(a) If p is a positive homogeneous, subadditive 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

$$p_K(x) = \inf\left\{t > 0 : \frac{x}{t} \in K\right\} , \text{ for } x \in X . \tag{1.6}$$

Lemma 1.21. *Let K be a balanced absorbing convex in X, then p_K is a seminorm.*

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

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

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

Lemma 1.22. *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 $|t| < \epsilon$,

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

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

Note that in (1.6), 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 preceding theorems, we have

Lemma 1.23. *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 other 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 1.24. *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 1.25 (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 other 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 $\operatorname{span}\{y\}$ by

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

We claim that for all such ay,

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

This is obvious for $a \leq 0$, for then $\ell(ay) = a \leq 0$ while $p_K \geq 0$. If $a > 0$, since $y \notin K$, $p_K(y) \geq 1$. So, $p_K(ay) \geq a = \ell(ay)$ 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 1.23, since K is absorbing at each of it's points, it follows from that $p_K(x) < 1$ for every x in K, thus

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

Corollary 1.26. *Let K denote a absorbing convex set. For any y not in K there is a nonzero linear functional ℓ that satisfies*

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

Theorem 1.27 (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*

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

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

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

1.6 Continuous 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 1.28. *Let X be a normed lienar space. Let M be a hyperplane in X. Then either M is closed or M is dense.*

Proof. Note that $\dim(X/\overline{M}) \le \dim(X/M)$, then the desired result follows. $\square$

Example 1.22. 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 $n \ge 1$, ker f is dense but clearly ker $f \neq c_0$.

Theorem 1.29. *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.*

First 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. □

Second proof. If there exises $\{x_n\}$ and $\epsilon > 0$, so that $\|x_n\| \to 0$ but $|f(x_n)| \geq \epsilon$ 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. Since $f(\frac{x_1}{f(x_1)}) = 1$. □

We will denote by X^* all the continuous linear functional on the norm lienar space X. In other words, $X^* := \mathcal{B}(X, \mathbb{F})$. Then it follows from Theorem 1.5 that X^* is a Banach space endowed with the operator norm.

Theorem 1.30 (Hahn-Banach Extension Theorem in Normed Space)**.** *Let X be a normed vector space. Let Y is a subspace in X. Let $f : Y \to \mathbb{F}$ is a bounded linear functional, then there is an $\bar{f}$ in X^* such that $\bar{f}|_Y = f$ and $\|\bar{f}\| = \|f\|$.*

Proof. Use Theorem 1.20 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 $\bar{f}$ on $\overline{Y}$ such that $\bar{f}|_Y = f$ and $\|\bar{f}\| = \|f\|$.

Corollary 1.31. *If X is a normed space, $\{x_1, x_2, \ldots, x_d\}$ 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 $1 \leq j \leq d$.*

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. $\square$

Corollary 1.32. *Let X be a normed linear space. For any $x \in X$,*

$$\|x\| = \sup_{x^* \in X^*, \|x^*\| \leq 1} |\langle x, x^* \rangle| \, . \tag{1.7}$$

Moreover, this supremum is attained.

Proof. Obviously,

$$\sup_{x^* \in X^*, \|x^*\| \leq 1} |\langle x, x^* \rangle| \leq \|x\| \, .$$

On the other hand, define $f : \mathrm{span}\{x\} \to \mathbb{F}$ by

$$f(\beta x) = \beta \|x\| \, , \text{ for all } \beta \in \mathbb{F} \, .$$

Then f is bounded and $\|f\| = 1$. By Hahn-Banach extension theorem, there is a x^* in X^* such that $\|x^*\| = 1$ and $x^*(x) = f(x) = \|x\|$. $\square$

Corollary 1.33. *X, Y are normed space. $T \in \mathcal{B}(X, Y)$, then*

$$\|T\| = \sup_{\|x\| \leq 1, \|y^*\| \leq 1} |\langle Tx, y^* \rangle| \, . \tag{1.8}$$

Corollary 1.34. *Let X be a normed linear space. Let M be a proper closed subspace of X. For any $x_0 \notin M$, there exists f in X^*, depending on x_0, such that*

(a) $\|f\| = 1$.

(b) $f(x_0) = d(x_0, M)$.

(c) $f(x) = 0$ *for all* x *in* M.

Proof. Let $Y = \operatorname{span}\{x_0\} \oplus M$. Define a functional $f : Y \to \mathbb{F}$ by

$$f(\alpha x_0 + m) = \alpha d\,(x_0, M) \qquad \text{for } \alpha \in \mathbb{F},\ m \in M\,.$$

Then it's easy to check that f is a continuous linear function on Y. We show that $\|f\| = 1$. To this end, observe that for $\alpha \in \mathbb{F}$ and $m \in M$,

$$|f(\alpha x_0 + m)| = |\alpha| d\,(x_0, M) = \inf_{m' \in M} \|\alpha x_0 + m'\|\,.$$

On the one hand, $|f(\alpha x_0 + m)| \leq \|\alpha x_0 + m\|$ implies $\|f\| \leq 1$. On the other hand, running m though M

$$\inf_{m' \in M} \|\alpha x_0 + m'\| \leq \|f\| \inf_{m \in M} \|\alpha x_0 + m\|$$

which implies $\|f\| \leq 1$. So by Hahn-Banach extension theorem, we can extend f to X as a continuous linear functional with the same norm. $\square$

Corollary 1.35. *Let* X *be a normed linear space. Let* S *be a subset of* X. *Then* $\overline{S} = X$ *if and only if for any* $f \in X^*$, $f(x) = 0\,, \forall x \in S$ *implies* $f = 0$.

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 $x_n \to x$, $y_n \to y$ with respect to 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. □

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 trivial 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. □

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.

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 $C[a, b]$.

The (External) Direct Sum of Hilbert Spaces Suppose H and K are Hilbert spaces. We want to define $H \oplus K$ so that it becomes a Hilbert space. This is not a difficult assignment.

Proposition 2.4. *If H and K are Hilbert spaces, the external direct sum $H \oplus K$ endowed with the inner product*

$$\langle (h_1, k_1), (h_2, k_2)(h_2, k_2)\rangle := \langle h_1, k_1\rangle + \langle h_2, k_2\rangle$$

for every $(h_1, k_1), (h_2, k_2) \in H \oplus K$ is a Hilbert space.

Now what happens if we want to define $H_1 \oplus H_2 \oplus \cdots$ for a sequence of Hilbert spaces $H_1, H_2, \ldots$? There is a problem about the completeness of this infinite (external) direct sum, but this can be overcome as follows.

Proposition 2.5. *Let $\{H_i\}_{i \in I}$ be a collection of Hilbert spaces. Let*

$$H = \left\{ (h_i) : h_i \in H_i \text{ and } \sum_i \|h_i\|^2 < \infty \right\} .$$

For $h = (h_i)$ and $g = (g_i)$ in H, define

$$\langle h, g \rangle = \sum_i \langle h_i, g_i \rangle \tag{2.4}$$

Then $\langle , \rangle$ is an inner product on H and the norm relative to this inner product is $\|h\| = \left[\sum_i \|h_i\|^2 \right]^{1/2}$. With this inner product, H is a Hilbert space.

Proof. If $h = (h_i)$ and $g = (g_i) \in H$, then the CBS inequality implies

$$\sum_i |\langle h_i, g_i \rangle| \leqslant \sum_i \|h_i\| \, \|g_i\| \leqslant \left(\sum_i \|h_i\|^2 \right)^{1/2} \left(\sum_i \|g_i\|^2 \right)^{1/2} < \infty.$$

Hence the series in (2.4) converges absolutely. Trivially, we check that $\langle , \rangle$ is an inner product. H is a Hilbert sapce follows form Proposition 1.3. $\square$

The space H is called the (external) direc sum of $\{H_i\}_{i \in I}$ and is denotd by $\oplus_i H_i$.

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 $M^{\perp} = (\operatorname{span} M)^{\perp} = (\overline{\operatorname{span}} M)^{\perp}$.

2.2.1 Best Approximation

Theorem 2.6 (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. □

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 2.6 holds.

Theorem 2.7 (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*

$$\operatorname{Re}\langle x - k_0, k - k_0\rangle \le 0, \ \textit{for all } k \in K\,. \tag{2.5}$$

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 (2.5) holds , we get $\|x-k\|^2 \ge \|x-k_0\|^2$. so k_0 is a best approximation. The uniqueness is guaranteed by Theorem 2.6.

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

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

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

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

So we let λ tends to zero , we get (2.5). □

2.2.2 Orthogonal Decomposition

Theorem 2.8. *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 2.7, $m \in M$ is the unique best approximation to x from M if and only if

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

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

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

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

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

the desired result follows. $\square$

Theorem 2.8 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

$$P_M : H \to M; x \mapsto P_M(x) \tag{2.6}$$

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

Theorem 2.9 (Orthogonal Decomposition)**.** *H is Hilbert space, M is a closed subspace. Then*

$$H = M \oplus M^{\perp}\,. \tag{2.7}$$

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

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

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

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 2.11 (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 $n \in \mathbb{N}$,*

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

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

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

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

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

Theorem 2.12 (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,

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

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 $\lambda_i\ne 0$. So $\sum_{i\in I}\lambda_i e_i$ converges, by the first step. □

Lemma 2.13. *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*

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

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

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

Theorem 2.14 (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 2.13 and Proposition A.11, the theorem follows. □

Theorem 2.15 (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) *$x \in \overline{\mathrm{span}}(\mathcal{E})$.*

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

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

Using Lemma 2.13 we have

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

So

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

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

Remark 2.4. There exists the projection of x on $\overline{\operatorname{span}}(\mathcal{E})$ 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 2.15 that an orthonomal set $\mathcal{E} = \{e_i\}_{i\in I}$ is complete is equivalent to the follows propositions. (a) $\overline{\operatorname{span}}(\mathcal{E}) = H$. (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 2.16. *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 2.17. *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

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

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

Theorem 2.18. *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 1.2 we have H is separable. 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. $\square$

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

Theorem 2.19. *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 2.12, T is a surjection. By Theorem 2.15, T isometry. $\square$

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 $\ell^2(\mathbb{Z},\mathbb{C})$.

2.5 F.Riesz's Representation Theorem

In a inner product space, we can introduce the notion of *orthogonality* of two vectors. Thanks to this fact, a Hilbert space may be identified with its dual space, i.e., the space of bounded linear functionals. This result is the representation theorem of F.Riesz, and the whole theory of Hilbert spaces is founded on this theorem.

Let $(X,\langle\cdot,\cdot\rangle)$ be an inner product space over a field $\mathbb{F}$. Choose and fix $y\in X$. Define a map $f_y : X\to\mathbb{F}$ by $f_y(x)=\langle x,y\rangle$. We claim that f_y is a bounded linear functional on X. In fact, for any $x\in X$, $|f_y(x)| = |\langle x,y\rangle| \le \|x\|\|y\|$ (by the CBS Inequality). That is, f_y is bounded and $\|f_y\|\le\|y\|$. Since

$$f_y(y)=\langle y,y\rangle=\|y\|^2 \Rightarrow \frac{|f_y(y)|}{\|y\|}=\|y\|$$

we have that $\|f_y\|=\|y\|$. The above observation simply says that each element y in an inner product space $(X,\langle\cdot,\cdot\rangle)$ determines a bounded linear functional on X. The following theorem asserts that if $X=H$ is a Hilbert space then the converse of this statement is true. That is, every bounded linear functional on a Hilbert space H is, in fact, determined by some element $y\in H$.

Theorem 2.20 (F.Riesz's Representation Theorem)**.** *Let H be a Hilbert space over $\mathbb{F}$. If $f : H\to\mathbb{F}$ is a bounded linear functional on H (i.e., $f\in H^*$), then there exists a uniquely determined vector $y=y_f\in H$ such that*

$$f(x)=\langle x,y\rangle \quad \textit{for all } x\in H \quad \textit{and } \|f\|=\|y\|\,.$$

Proof. The uniqueness of y is clear, since $\langle x, y - y' \rangle = 0$ for all $x \in X$ implies $y = y'$. To prove its existence, consider the null space $N = \ker(f)$. If $f = 0$ then take $y = 0$. Then assume that $f \neq 0$, and hence N is a closed proper subspace of H. By Theorem 2.9, there exists $z \in N^{\perp}$ and $z \neq 0$. Without loss of generality, let $\|z\| = 1$. Then define $f_z : H \to \mathbb{F}$ by $f_z(x) = \langle x, z \rangle$. We have shown that f_z is a bounded linear functional. Observe that

$$\ker(f) = N \subset \ker(f_z)\,,$$

then we have that there exists a constant $c \in \mathbb{F}$ so that $f_z = cf$. Since $f_z(z) = \langle z, z \rangle = \|z\|^2 = 1$, we get that $c = 1/f(z)$. Thus

$$f(x) = \frac{1}{f(z)} \langle x, z \rangle \qquad \text{for all } x \in H\,.$$

Taking $y = \dfrac{1}{\bar{f(z)}} z$, the desired result follows. □

Remark 2.5. The conclusion of Riesz's representation theorem may fail if $(X, \langle \cdot, \cdot \rangle)$ is an incomplete inner product space. It's easy to give a counterexample. In fact, let H be the completion of X. Take any $z \in H \backslash X$ and define $f : X \to \mathbb{F}$ by $f(x) = \langle x, z \rangle$ for all x. Obviously, there does not exist an element $y \in X$ so that $\langle x, y \rangle = \langle x, z \rangle$ for all $x \in H$. If exists, since X is dense in H, we have $y = z$, which contradicts to $z \notin X$.

Let H be a Hilbert space. Riesz's representation theorem gives a norm-preserving, one-to-one correspondence $f \leftrightarrow y_f$ between H^* and H. By this correspondence, H^* may be identified with H as an abstract set but it is not allowed to identify, by this correspondence, H^* with H as linear spaces, since the correspondence $f \leftrightarrow y_f$ is *conjugate linear* :

$$(\alpha_1 f_1 + \alpha_2 f_2) \leftrightarrow (\bar{\alpha}_1 y_{f_1} + \bar{\alpha}_2 y_{f_3})$$

where α_1, α_2 are scalars. Indeed, the dual space H^* of H constitutes also a Hilbert space. Firstly, the operator norm on H^* satisfies the *parallelogram identity.*

Secondly, in the case of complex field, for every $f, g \in H^*$,

$$
\begin{aligned}
\langle f, g\rangle &= \frac{\|f+g\|^2 - \|f-g\|^2}{4} + i\left(\frac{\|f+ig\|^2 - \|f-ig\|^2}{4}\right) \\
&= \frac{\|y_f+y_g\|^2 - \|y_f-y_g\|^2}{4} + i\left(\frac{\|y_f-iy_g\|^2 - \|y_f+iy_g\|^2}{4}\right) \\
&= \frac{\|y_g+y_f\|^2 - \|y_g-y_f\|^2}{4} + i\left(\frac{\|y_g+iy_f\|^2 - \|y_g-iy_f\|^2}{4}\right) \\
&= \langle y_g, y_f\rangle = f(y_g)\,.
\end{aligned}
$$

Moreover, any continuous linear functional T on the Hilbert space H^* is thus identified with a uniquely determined element y_T of H as follows:

$$
T(f) = f(y_T) \quad \text{for all} \quad f \in H^*\,.
$$

This fact will be referred to as the *reflexivity* of Hilbert spaces.

Chapter 3

Topological 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.

3.1.1 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. $\square$

As a consequence of this proposition, every vector topology τ is *translation-invariant* (or simply invariant, for brevity): A set $U \in \tau$ 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 $\lambda V \subset U$ whenever $|\lambda| \leq \epsilon$. If we let

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

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.*

3.1.2 Separation Properties

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

$$(K + V) \cap (C + V) = \emptyset\,.$$

Proof. For any $x \in K$, there is a balanced open neighborhood V_x of 0 such that $(x + V_x + V_x) \cap C = \emptyset$. 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 Remark 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.

Proposition 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.*

3.1.3 Types of Topological Vector Spaces

Let (X, τ) be a TVS. A subset B of X is called *bounded* if for any neighborhood U of 0, there exist some $\epsilon > 0$, depending on U, such that $|\lambda| B \subset U$ for all $|\lambda| \leq \epsilon$. For example, every compact subset of X is bounded. Note that any subset of a bounded set is also bounded.

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. Since that for any seminorm p on X and for $\epsilon > 0$,

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

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

Example 3.1. X be a vector space. Let $\mathcal{P}$ be a family of seminorms on X. Let τ be the weakest topology on X staisfying that p is continuous for each $p \in \mathcal{P}$. What do the basic open sets for this topology look like? So we need to find a base of τ. For any $x_0 \in X$, let

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

where Φ is a finite subset of $\mathcal{P}$ and $\epsilon > 0$. Clearly, $V(x_0,\Phi,\epsilon) \in \tau$. Moreover, τ is the weakest topology containing $\{V(x_0,\Phi,\epsilon)\}$. To see this, it suffices to show that every p is continuous relative to the topology generated by $\{V(x_0,\Phi,\epsilon)\}$. Take any $x_0 \in X$ and $\epsilon > 0$, then

$$\{x : |p(x) - p(x_0)| < \epsilon\} = \bigcup_{x:|p(x)-p(x_0)|<\epsilon} V(x,p,\epsilon - |p(x) - p(x_0)|)\,.$$

Thus τ is generated by $\{V(x_0,\Phi,\epsilon)\}$.

Recall that a collection $\mathfrak{B}$ of subsets of a set X is a base for a topology on X if and only if

- $X = \cup\{B : B \in \mathfrak{B}\}$; i.e., each $x \in X$ belongs to some $B \in \mathfrak{B}$, and
- if $x \in B_1 \cap B_2$ for some B_1 and B_2 in $\mathfrak{B}$, then there is a $B_3 \in \mathfrak{B}$ such that $x \in B_3 \subset B_1 \cap B_2$.

Then, it's obvious to show that indeed

$$\mathfrak{B} = \{V(x; \Phi; \epsilon) : x \in X,\ \Phi \subset \mathcal{P} \text{ is finite}, \epsilon > 0\}$$

is a base for the topology τ. We write $V(0, \Phi, \epsilon)$ as $V(\Phi, \epsilon)$ for brevity. Then one can see that

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

is a balanced convex open neighborhood base of 0. Besides, it's easy to find that,

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

which hints that (X, τ) is a LCS.

We now show that addition and scalars mutipulication are continous relative to τ. In fact, 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 $\alpha_0 x_0 + V(\Phi, \epsilon)$, let $|\alpha - \alpha_0| \leq \delta_1$ and $x \in x_0 + V(\Phi, \delta_2)$, then for all $p \in \Phi$,

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

Thus scalars mutipulication is continuous.

Now we try to make (X, τ) be a TVS, then since 0 has a (balanced) convex local base, it is a LCS. To this end, it suffices to make (X, τ) be a Hausdorff space. We need to assume that $\mathcal{P}$ is *separating*:

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

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. Conversely, it's easy to check that this condition is necessary.

Thus when $\mathcal{P}$ is separating, (X, τ) is a LCS. 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. We shall show that any LCS has the same structure as in Example 3.1.

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) $0 \in \operatorname{int} V(p, 1)$. (c) p is continuous at 0.*

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

To show (b) implies (c), since $0 \in \operatorname{int}\{x : p(x) < 1\}$, 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 obvious. □

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 $p \leq q$.*

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 $0 \leq r \leq 1$. 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 neighborhood 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 $0 \leq r \leq 1, |\beta| = 1$.

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

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. Firstly, we show that $\mathcal{P}$ is separating. If $p_U(x) = 0$ for all $U \in \mathcal{B}$, $x \in U$ 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 1.23, 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 $\tau \subset \tau_1$. □

3.2.1 Metrizable Locally Convex Spaces

We recall that a topology τ on a set X is said to be metrizable if there is a metric d on X which is compatible with τ. In that case, the balls with radius $1/n$ centered at x form a local base at x. This gives a necessary condition for metrizability which, for topological vector spaces, turns out to be also sufficient.

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.

If $\mathcal{P}$ is a family of seminorms on topological vector X, say that $\mathcal{P}$ determines the topology on X if the topology of X is the same as the topology induced by $\mathcal{P}$.

Lemma 3.10. *Let (X, τ) be a LCS , $\mathcal{P} = \{p_n : n \in \mathbb{N}\}$ is countable separating 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 on X, and the topology induced by d coincides with τ. Moreover, $\{B(0, \frac{1}{n}) : n \geq 1\}$ is a balanced (may NOT convex) local base τ.

Proof. It's easy to check that d is a translation invariant metric on X. We only show that the topology induced by the metric d, denoted by τ_d, coincides with τ. Firstly, in order that $\tau_d \subset \tau$, it suffices to show that for every fixed y, the (open) ball centered at y with radius $r > 0$ is in τ, i.e.,

$$B(y, r) \coloneqq \{x : d(x, y) < r\} \in \tau .$$

Observe that for fixed $y \in X$, the series

$$\sum_{n=1}^{\infty} \frac{\min\{p_n(x-y), 1\}}{2^n} \text{ converges to } d(x, y) \text{ uniformly in } x .$$

Since each p_n is continuous relative to τ, the map $x \mapsto d(x, y)$ is continuous relative to τ for fixed y. Thus $B(y, r) \in \tau$.

On the other hand, in order that $\tau \subset \tau_d$, it suffices to show that for each fixed n, p_n is continuous relative to τ_d. For any gien $\epsilon \in (0, 1/2^n)$, note that

$$p_n(x) < \epsilon \quad \text{for all} \quad d(x, 0) < \frac{1}{2^n} \frac{\epsilon}{1+\epsilon} .$$

Thus p_n is continuous at 0 relative to τ_d. By Proposition 3.6, p_n is continuous.

Finally, we show that $B(0, 1)$ is balanced and convex. For $x \in B(0, 1)$ and $|\lambda| \leq 1$, since $x \mapsto \frac{x}{1+x}$ is increasing on $[0, \infty)$, we have

$$d(\lambda x, 0) = \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{|\lambda| p_n(x)}{1 + |\lambda| p_n(x)} \leq \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{p_n(x)}{1 + p_n(x)} = d(x, 0) < 1 .$$

So $\lambda x \in B(0,1)$. Thus $B(0,1)$ is balanced. □

Remark 3.3. We should emphasize that the open ball relative to such a d may not be convex. Indeed, for every $x, y \in B(0,1)$ and $\lambda \in (0,1)$, we want

$$\begin{aligned} d(\lambda x + (1-\lambda)y, 0) &= \sum_{n=1}^{\infty} \frac{1}{2^n} \frac{p_n(\lambda x + (1-\lambda)y)}{1 + p_n(\lambda x + (1-\lambda)y)} \\ &\leq \sum_{n=1}^{\infty} \frac{1}{2^n} \left[\frac{\lambda p_n(x)}{1+p_n(x)} + \frac{(1-\lambda)p_n(y)}{1+p_n(y)} \right] \\ &= \lambda d(x,0) + (1-\lambda) d(y,0) \,. \end{aligned}$$

In order that the argument valid, we need the inequality

$$\frac{p_n(\lambda x + (1-\lambda)y)}{1 + p_n(\lambda x + (1-\lambda)y)} \leq \frac{\lambda p_n(x)}{1+p_n(x)} + \frac{(1-\lambda)p_n(y)}{1+p_n(y)} \,.$$

But take $y = 0$, the inequality is in the opposite direction.

However, there does exist a translation invariant metric ρ on X compatible with the topology τ so that all open balls relative to ρ centered at 0 are balanced and convex. See 1.24 Theorem in *Functional Analysis* by W.Rudin.

If $\mathcal{P}$ is a family of seminorms on LCS X, say that $\mathcal{P}$ determines the topology on X if the topology of X is the same as the topology induced by $\mathcal{P}$.

Theorem 3.11 (Metrizable LCS)**.** *Let (X, τ) be a locally convex space. Then the followings are equivalent.*

(a) (X, τ) is measurable.

(b) (X, τ) is first-countable. In other words, (X, τ) has a countable local base.

(c) τ is determined by a countable separating family of seminorms.

Proof. $(a) \Rightarrow (b)$ is obvious. $(c) \Rightarrow (a)$ was proved in Lemma 3.10. We only need to show that $(b) \Rightarrow (c)$. To this end, assume that $\{U_n\}$ is a local base of τ.

Given $n \geq 1$, because (X, τ) is locally convex, there are continuous seminorms $\Phi_n = \{q_{n,1}, \ldots, q_{n,k}\}$ and positive real number ϵ_n such that $V(\Phi_n, \epsilon_n) \subset U_n$. Let

$$p_n = \frac{1}{\epsilon_n}(q_{n,1} + \cdots + q_{n,k}) .$$

Then $x \in U_n$ whenever $p_n(x) < 1$, i.e.,

$$V(p_n, 1) \subset U_n . \tag{3.2}$$

We claim that $\{p_n\}$ determines the topology τ. Denote by τ' the vector topology induced by $\{p_n\}$. Since p_n is continuous with respect to τ for each n, there must be $\tau' \subset \tau$. To show the other direction, note that (3.2) implies that $\{V(p_n, 1)\}$ is a local base of τ. Therefore, $\tau \subset \tau'$. We now complete the proof. $\square$

Remark 3.4. In fact, every first-countable topological vector space is metrizable. Specifically, if (X, τ) is a topological vector space with a countable local base, then there exists a translation invariant mertic ρ on X compatible with the topology τ so that all open balls relative to ρ centered at 0 are balanced. See 1.24 Theorem in *Functional Analysis* by W.Rudin.

3.2.2 Normable Locally Convex Spaces

Recall that if X is a TVS and $B \subset X$, then B is *bounded* if for every open set U containing 0, there is an $\epsilon > 0$ such that $\lambda B \subset U$ for all $|\lambda| \leq \epsilon$. When X is metrizable, there is a possibility of misunderstanding, since another very familiar notion of boundedness exists in metric spaces:

If d is a metric on a set X, a set $E \subset X$ is said to be d-bounded if there is a number $M < \infty$ such that $d(x, y) \leq M$ for all x and y in E.

If X is a topological vector space with a compatible metric d, the bounded sets and the d-bounded ones need NOT be the same, even if d is invariant. For instance, if d is a metric such as the one constructed in Lemma 3.10, then X itself is d-bounded (with $M = 1$) but, as we shall see presently, X cannot be bounded, unless $X = \{0\}$. If X is a normed space and d is the metric induced by the norm, then the two notions of boundedness coincide: B is bounded lerative to the topology if and onley if $\sup_{b\in B} \|b\| < \infty$; but if d is replaced by $d_1 = d/(1+d)$ (an invariant metric which induces the same topology they do not.

Theorem 3.12 (Normable LCS)**.** *Let (X, τ) be a LCS. Then (X, τ) is normable if and only if X locally bounded, i.e., 0 has a bounded open neighborhood.*

Proof. It has already been shown that the open unit ball in a normed space is bounded. So assume that X is a LCS that 0 has a bounded open neighborhood U. It must be shown that there is norm on X that defines the same topology. By local convexity, there is a continuous seminorm p such that $V(p, 1) \subset U$. It will be shown that p is a norm and defines the topology on X.

To see that p is a norm, we need to show that $p(x) = 0$ implies $x = 0$. For every neighborhood W of 0, there exists $\epsilon > 0$ so that $\epsilon U \subset W$. Observing that

$$x \in V(p, \epsilon) = \epsilon V(p, 1) \subset \epsilon U \subset W\,, \tag{3.3}$$

since W is arbitary and X is a Hausdorff space, we have $x = 0$. Thus p is a norm. Denote by τ' the topology induced by the norm p. Because p is continuous on (X, τ), $\tau' \subset \tau$. On the orther hand, it follows from (3.3) that $\{V(p, \epsilon)\}$ is a local base of τ. Thus $\tau \subset \tau'$. □

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^* for short) 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 in the following sense: there exists $M > 0$ and a neighborhood V of 0 such that $|f(x)| \le M$ 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

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

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. □

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 $p_1, \ldots, p_n$ in $\mathcal{P}$ and positive scalars $\alpha_1, \ldots, \alpha_n$ such that*

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

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

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

Obviously, q is a (continuous) seminorm, and

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

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

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

□

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.,

$$V \cap f(S) = \emptyset\,.$$

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 $\tau_1 = \tau_2$. □

Next, we shall give a topological characterization of the algebraic concept of finite dimensionality, as a generalization of Theorem 1.12.

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 $x_1, \dots, x_m$. Then $\dim Y \le m$. 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

$$\tau_M = \{U \in X/M : Q^{-1}(U) \in \tau\}\,. \tag{3.4}$$

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

Theorem 3.21. *Let M be a closed subspace of a TVS (X, τ). Let τ_M is quotient topology on X/M. Then the following propositions hold.*

(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). $\square$

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. $\square$

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. $\square$

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 $\tilde{\mathcal{P}} \coloneqq \{\tilde{p} : p \in \mathcal{P}\}$ is a separating 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

3.6.1 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*

$$\operatorname{Re} \ell x < c \leq \operatorname{Re} \ell y \tag{3.5}$$

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 $x_0 \notin C$, $p(x_0) \geq 1$.

Just like in Theorem 1.25, 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 Theorem 1.18, ℓ can be extened as a linear functional on X domainted by p. Note that $p(x) \leq 1$ 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 $\mathbb{R}$, 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 right end point of $\ell(A)$ to get the conclusion of (3.5). $\square$

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

Figure 3.1: ℓ is open mapping

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*

$$\mathrm{Re}\, \ell x < c < \mathrm{Re}\, \ell y \tag{3.6}$$

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

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

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 3.5. 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 $X = \mathbb{R}^2$, $A = \{(x, y) \in \mathbb{R}^2 : y \leq 0\}$, and $B = \left\{(x, y) \in \mathbb{R}^2 : y \geq x^{-1} > 0\right\}$, 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 F-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. Hence, $\ell(M)$ must be a *proper subspace* of the scalar field. This forces

$$\ell(M) = \{0\} \text{ and } \ell(x_0) \neq 0\,.$$

The desired functional is obtained by dividing ℓ by $\ell(x_0)$. ☐

There is an another useful corollary of the separation theorem.

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

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

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. □

3.6.2 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.*

3.6.3 An 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

$$N = \{y \in Y : \ell(y) = 0\}$$

and pick $y_{\mathrm{o}} \in Y$ such that $\ell(y_{\mathrm{o}}) = 1$. Since ℓ is continuous on Y, $y_{\mathrm{o}} \notin N = \mathrm{cl}_Y(N) = \mathrm{cl}_X(N) \cap Y$. So $y_{\mathrm{o}} \notin \mathrm{cl}_X(N)$. Then there exists a $\Lambda \in X^*$ such that $\Lambda y_{\mathrm{o}} = 1$ and $\Lambda = 0$ on $\mathrm{cl}_X(N)$. For any $y \in Y$, note that $y - \ell(y)y_{\mathrm{o}} \in N$, since $\ell(y_{\mathrm{o}}) = 1$. Hence

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

Thus $\Lambda = \ell$ on Y. □

3.7 Weak topologies on LCS

Let X be a LCS. Denote by X^* 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 3.2. (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

$$p_{x^*}(x) = |\langle x, x^* \rangle| \ , \text{ for all } x \in X \, . \tag{3.7}$$

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 $\{p_x : x \in X\}$, where

$$p_x(x^*) = |\langle x, x^* \rangle| \ , \text{ for all } x^* \in X^* \, . \tag{3.8}$$

In fact, the weak topology on X is the weakest one with respect to which each element in X^* is continuous. Also, regarding each $x \in X$ as a linear functional on X^*, denoted as Jx and given by

$$Jx : x^* \mapsto \langle x, x^* \rangle \ , \text{ for all } x^* \in X^* \, , \tag{3.9}$$

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,

$$\left\{ V(\Phi, \epsilon) \coloneqq \{x : |\langle x, x^* \rangle| < \epsilon, x^* \in \Phi\} : \Phi \subset X^* \text{ is finite, } \epsilon > 0 \right\}, \tag{3.10}$$

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

$$\left\{ V(\Phi, \epsilon) \coloneqq \{x^* : |\langle x, x^* \rangle| < \epsilon, x \in \Phi\} : \Phi \subset X \text{ is finite, } \epsilon > 0 \right\}. \tag{3.11}$$

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

$$\langle x_i, x^* \rangle \to \langle x, x^* \rangle \ , \text{ for all } x^* \in X^* \, . \tag{3.12}$$

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

$$\langle x, x_i^* \rangle \to \langle x, x^* \rangle \ , \text{ for all } x \in X \tag{3.13}$$

(c) Let $(Y, \mathcal{T})$ be a topological space. Then $f : (Y, \mathcal{T}) \to (X, \tau_w)$ is continuous if and only if

$$y \mapsto \langle f(y), x^* \rangle ; Y \to \mathbb{R}$$

is continuous for every $x^* \in x^*$. $g : (Y, \mathcal{T}) \to (X^*, \tau_{w^*})$ is continuous if and only if

$$y \mapsto \langle x, f(y) \rangle ; Y \to \mathbb{R}$$

is continuous for every $x \in X$.

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

$$\sup_{x \in M} |\langle x, x^* \rangle| < \infty\,, \text{ for all } x^* \in X^*\,. \tag{3.14}$$

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

$$\sup_{x^* \in N} |\langle x, x^* \rangle| < \infty\,, \text{ for all } x \in X\,. \tag{3.15}$$

3.7.1 Duality

Theorem 3.31. *Let (X, τ) be a 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, since $\tau_w \subset \tau$, thus $(X, \tau_w)^* \subset X^*$. □

Theorem 3.32. *Let (X, τ) be a 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 $x_1, \ldots, x_n$ in X and positive scalars $\alpha_1, \ldots, \alpha_n$, by Theorem 3.15, such that

$$|f(x^*)| \leq \sum_{k=1}^{n} \alpha_k |\langle x_k, x^* \rangle|\,, \quad \text{for all } x^* \in X^*.$$

Thus

$$\ker f \subset \bigcap_{k=1}^{n} \ker(Jx_k)\,.$$

By Lemma 0.1, $f \in \operatorname{span}\{Jx_1, \cdots, Jx_n\}$, of course $f \in J(X)$. $\square$

Remark 3.6. 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). Then one can show that this topology is exactly τ_w !

3.7.2 Weak closure

Theorem 3.33. *Let (X, τ) be a LCS. Let K be a nonempty convex subset of X. Then the weak closure of K is equal to its original closure, i.e.,*

$$\overline{K} = \overline{K}^{\tau_w}\,. \tag{3.16}$$

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 Hahn-Banach 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 $|\langle x - x_0, x^*\rangle| \geq \epsilon$ for all $k \in K$, i.e.,

$$K \cap V(x_0, x^*, \epsilon) = \emptyset\,.$$

Thus $x_0 \notin \overline{K}^{\tau_w}$, this is a contradiction. $\square$

Corollary 3.34. *Let (X, τ) be a LCS. Let K be a nonempty convex subset of X. Then K is closed if and only if K is weakly closed.*

Corollary 3.35. *Let (X, τ) be 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 $\{x_n\}$, let K be the weak closure of H. Clearly, $x \in K$. 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 in Corollary 3.35, consider the following example.

Example 3.3. Let X be a compact Hausdorff space, for example, the unit interval on the real line. Assume that f and f_n are continuous scalar (real or complex) functions on X such that

(a) $f_n(x) \to f(x)$ for every $x \in X$;

(b) $|f_n(x)| \leq 1$ 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 $C(X)$ be the Banach space of all continuous scalar functions on X, normed by the supremum. Then norm convergence is the same as uniform convergence on X. If μ is any scalar Borel measure on X, with finite total variation, 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 $C(X)^*$ with the space of all regular scalar Borel measures on X.

3.7.3 Annihilators

Annihilator is in an analogue of the orthogonal complement in Hilbert space.

Definition 3.3. Let (X, τ) be a LCS. Let X^* be the dual space of X.

- For a subspace of X, namely M, the *annihilator of* M is defined by

$$M^\perp = \{x^* \in X^* : \langle x, x^* \rangle = 0 \text{ for all } x \in M\} \ . \tag{3.17}$$

- For a subspace of X^*, namely N, the *annihilator of* N is defined by

$${}^\perp N = \{x \in X : \langle x, x^* \rangle = 0 \text{ for all } x^* \in N\} \ . \tag{3.18}$$

In other words, $M^\perp$ 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 a weak-closed subspace of X, and ${}^\perp N$ is a weak*-closed subspace of X^*. The following theorem describes the duality between these two types of annihilators.

Theorem 3.36. *Under the preceding hypotheses,*

$${}^\perp \left(M^\perp\right) = \overline{M}^{\tau_w} = \overline{M} \, , \ \left({}^\perp N\right)^\perp = \overline{N}^{\tau_{w^*}} \ . \tag{3.19}$$

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. □

3.8 The Krein-Milman Theorem

3.8.1 Compact Convex Sets

Let X be a vector space and $E \subset X$, the *convex hull* of E will be denoted by $\operatorname{co}(E)$. Recall that $\operatorname{co}(E)$ is the intersection of all convex subsets of X which contain E. Equivalently, $\operatorname{co}(E)$ is the set of all finite convex combinations of members of E. If X is a topological vector space and $E \subset X$, the *closed convex hull* of E, written $\overline{\operatorname{co}}(E)$, is the closure of $\operatorname{co}(E)$.

We now turn to the question: What can one say about the convex hull $\operatorname{co}(K)$ of a compact set K? Even in a Hilbert space, $\operatorname{co}(K)$ need not be closed, and there are situations in which $\overline{\operatorname{co}}(K)$, is not compact. In Fréchet spaces the latter pathology does not occur (Theorem 3.38). The proof of this will depend on the fact that a subset of a complete metric space is compact if and only if it is closed and totally bounded.

Lemma 3.37. *If $A_1, \ldots, A_n$ are compact convex sets in a topological vector space X, then $\operatorname{co}(A_1 \cup \cdots \cup A_n)$ is compact.*

Proof. Let S be the simplex in $\mathbb{R}^n$ consisting of all $s = (s_1, \ldots, s_n)$ with $s_i \geq 0$, $s_1 + \cdots + s_n = 1$. Put $A = A_1 \times \cdots \times A_n$. Define $f : S \times A \to X$ by

$$f(s, a) = s_1 a_1 + \cdots + s_n a_n$$

and put $K = f(S \times A)$. Clearly, $S \times A$ is compact and f is continuous, so K is compact with $K \subset \operatorname{co}(A_1 \cup \cdots \cup A_n)$. We will see that this inclusion is actually an equality because each A_i is convex. To this end, we show that K is convex, then since $A_i \subset K$, we have $K \supset \operatorname{co}(A_1 \cup \cdots \cup A_n)$.

If (s,a) and (t,b) are in $S\times A$ and if $\alpha\geq 0,\beta\geq 0,\alpha+\beta=1$ then

$$\alpha f(s,a)+\beta f(t,b)=f(u,c)$$

where $u=\alpha s+\beta t\in S$ and $c\in A$, because

$$c_i=\frac{\alpha s_i a_i+\beta t_i b_i}{\alpha s_i+\beta t_i}\in A_i \quad (1\leq i\leq n)\,.$$

We complete the proof. $\square$

Theorem 3.38. *Let X be a Frechet space and $K\subset X$ is compact, then $\overline{\mathrm{co}}(K)$ is compact.*

Proof. It suffices to show that $\mathrm{co}(K)$ is totally bounded. For any given $\epsilon>0$, there exists $\{k_1,\cdots,k_n\}\subset K$ so that

$$K\subset\{k_1,\cdots,k_n\}+B(0,\epsilon)\,.$$

Since X is locally convex, we suppose that $B(0,\epsilon)$ is convex. Thus

$$\mathrm{co}(K)\subset\mathrm{co}\{k_1,\cdots,k_n\}+B(0,\epsilon)\,.$$

By Lemma 3.37, $\mathrm{co}\{k_1,\cdots,k_n\}$ is compact, so there exists $\{k_1',\cdots,k_m'\}\subset\mathrm{co}(K)$ with

$$\mathrm{co}\{k_1,\cdots,k_n\}\subset\{k_1',\cdots,k_m'\}+B(0,\epsilon)\,.$$

Therefore,

$$\mathrm{co}(K)\subset\{k_1',\cdots,k_m'\}+B(0,2\epsilon)\,,$$

and the desired result follows. $\square$

We give an example that there exists a compact set in a Hilbert space whose convex hull is not closed.

Example 3.4. Consider

$$u_n=(\underbrace{0,\ldots,0}_{n-1},1/n,0,\ldots)\text{ for }n\geq 1\,,$$

and $K = \{u_n\} \cup \{0\}$ a compact subset of $\ell^2(\mathbb{N})$ since $u_n \to 0$. The convex hull of K is given by

$$\mathrm{co}(K) = \left\{ \sum_{n=1}^{k} a_n u_n : a_n \geq 0\,, \sum_{n=1}^{k} a_n \leq 1 \right\}.$$

So also $\sum_{n=1}^{k} 2^{-n} u_n$ lies in $\mathrm{co}(K)$ for each k. But this sequence converges to $\sum_{n=1}^{\infty} 2^{-n} u_n$ which does not lie in it. Thus $\mathrm{co}(K)$ is not closed.

Proposition 3.39. *If $E \subset \mathbb{R}^n$ and $x \in \mathrm{co}(E)$, then x lies in the convex hull of some subset of E which contains at most $n+1$ points (depending on x).*

Proof. It is enough to show that if $k > n$ and $x = \sum_{i=1}^{k+1} t_i x_i$ is convex combination of some $k+1$ vectors $x_i \in \mathbb{R}^n$, then x is actually a convex combination of some k of these vectors. Assume, with no loss of generality, that $t_i > 0$ for $1 \leq i \leq k+1$ The null space of the linear map (8)

$$(a_1, \ldots, a_{k+1}) \to \left(\sum_{1}^{k+1} a_i x_i, \sum_{1}^{k+1} a_i \right)$$

which sends R^{k+1} into $R^n \times R$, has positive dimension, since $k > n$ Hence there exists $(a_1, \ldots, a_{k+1})$, with some $a_i \neq 0$, so that $\sum a_i x_i = 0$ and $\sum a_i = 0$. since $t_i > 0$ for all i, there is a constant λ such that $|\lambda a_i| \leq t_i$ for all i and $\lambda a_j = t_j$ for at least one j. Setting $c_i = t_i - \lambda a_i$ we conclude that $x = \sum c_i x_i$ and that at least one c_j is 0; note also that $\sum c_i = \sum t_i = 1$ and that $c_i \geq 0$ for all i. $\square$

Corollary 3.40. *If K is a compact set in $\mathbb{R}^n$, then $\mathrm{co}(K)$ is compact.*

Proof. Let S be the simplex in $\mathbb{R}^{n+1}$ consisting of all $t = (t_1, \ldots, t_{n+1})$ with $t_i \geq 0$ and $\sum t_i = 1$. Let K be compact, $K \subset \mathbb{R}^n$. By the proposition that follows, $x \in \mathrm{co}(K)$ if and only if

$$x = t_1 x_1 + \cdots + t_{n+1} x_{n+1}$$

for some $t \in S$ and $x_i \in K$. In other words, $\mathrm{co}(K)$ is the image of $S \times K^{n+1}$ under the continuous mapping

$$(t, x_1, \ldots, x_{n+1}) \to t_1 x_1 + \cdots + t_{n+1} x_{n+1}\,.$$

Hence co(K) is compact. $\square$

3.8.2 The Krein-Milman Theorem

Let K be a convex subset of a linear space X. A nonempty set $S \subset K$ is called an *extreme set* of K if no point of S is an internal point of any line interval whose end points are in K, except when both end points are in S. Analytically, the condition can be expressed as follows: For any $z \in S$, if there exists $x, y \in K$ and $t \in [0,1]$ with

$$z = (1-t)x + ty\,,$$

then it must be $x, y \in S$. Clearly, if $\{S_i : i \in I\}$ is a family of extreme sets of K, then the intersection $\cap_{i \in I} S_i$ is also an extreme set of K. The *extreme points* of K are the extreme sets that consist of just one point. The set of all extreme points will be denoted by $E(K)$.

Example 3.5. K is the interval $0 \le x \le 1$; the two endpoints are extreme points.

Example 3.6. K is the closed disk

$$x^2 + y^2 \le 1\,.$$

Every point on the circle $x^2 + y^2 = 1$ is an extreme point. The open disk

$$x^2 + y^2 < 1$$

has no extreme points.

Example 3.7. K a polyhedron, including faces. Its extreme subsets are its faces, edges, vertices, and of course K itself.

Lemma 3.41. *Let K be a convex set, E an extreme subset of K, and F an extreme subset of E. Then F is an extreme subset of K.*

Proof. If $x, y \in K$, $t \in (0,1)$ so that $tx + (1-t)y \in F \subset E$, since E is an extreme subset of K, we have x, y in E. Since since F is an extreme subset of E, $tx + (1-t)y \in F$, we have x, y in F. Thus F is an extreme subset of K. $\square$

Lemma 3.42. *Let X, Y be two linear spaces. Let $T : X \to Y$ be a REAL linear. Let K be a convex subset of Y, E an extreme subset of K. Then $T^{-1}E$ is either empty or an extreme subset of $T^{-1}K$.*

Proof. If $T^{-1}E$ is not empty, then so is $T^{-1}K$. As we all konw, $T^{-1}K$ is convex. If $x_1 = T^{-1}y_1, x_2 = T^{-1}y_2$ in $T^{-1}E$, $t \in (0,1)$ satisfy that $tx_1+(1-t)x_2 \in T^{-1}K$, then $ty_1 + (1-t)y_2 \in E$. Since E is an extreme subset of K, y_1, y_2 are in K and hence x_1, x_2 are in $T^{-1}K$. □

Taking Y to be one dimensional, we get

Corollary 3.43. *Denote by H a convex subset of a linear space X, ℓ a REAL linear map of X into $\mathbb{R}$, $H_{\min}$ and $H_{\max}$ the subsets of H, where ℓ achieves its minimum and maximum, respectively.*

Assertion: When nonempty, $H_{\min}$ and $H_{\max}$ are extreme subsets of H.

The following two theorems show that under certain conditions $E(K)$ is quite a large set.

Theorem 3.44 (The Krein-Milman Theorem)**.** *Let X be a locally convex space. Let K be a nonempty compact convex subset of X. Then*

(i) K has at least one extreme point, i.e., $E(K) \neq \emptyset$; and

(ii) K the closed convex hull of the set of its extreme points, i.e.,

$$K = \overline{\text{co}}(E(K))\,.$$

Proof of part (i). Consider the collection

$$\{E : E \text{ is a nonempty closed extreme subsets of } K\}\ .$$

This collection is nonempty, for it contains K itself. Partially order this collection by inclusion. We claim that every totally ordered subcollection $\{E_j\}$ has a lower bound. That lower bound is the intersection $\cap_j E_j$. To see this, we have to show that $\cap_j E_j$ is nonempty, closed, and extreme.

We claim that every finite subset of the totally ordered collection $\{E_j\}$ has a nonempty intersection. This is because in being totally ordered by inclusion, the intersection of a finite subset of the collection $\{E_j\}$ is the smallest member of that subset, and hence $\cap_j E_j$ is nonempty as a consequence of the compactness of K. Being the intersection of closed sets, $\cap_j E_j$ is closed. $\cap_j E_j$ is an extreme subset of K, since the nonempty intersection of extreme subsets of a convex set K is itself an extreme subset of K.

We conclude from Zorn's lemma that K has a closed extreme subset E that is minimal with respect to inclusion. We claim that such an E consists of a single point. To see this, suppose, on the contrary, that E contains two distinct points. According to Theorem 3.25, there exists a continuous linear functional ℓ that separates these two points. Since E is compact, and $\mathrm{Re}\ell$ continuous and not constant on E, $\mathrm{Re}\ell$ achieves its maximum on some *proper* subset M of E.

Since ℓ is continuous and E is closed, M is closed. By Corollary 3.43 Since M is a real linear functional $\mathrm{Re}\ell$ assumes its maximum on a convex set E is an extreme subset of E. It is easy to show further that if E is an extreme subset of K, and M an extreme subset of E, then M is an extreme subset of K. Since E is a minimal extreme subset of K, and M an extreme subset smaller than E, we have a contradiction, into which we got by assuming that E contains more than one point. We conclude therefore that a minimal E consists of a single point. This single point is an extreme point of K. This completes the proof of part (i), and gives a little more:

(i') Every closed, extreme subset of K contains an extreme point. □

Proof of part (ii). To show that every point of K belongs to the closure of $\mathrm{co}(E(K))$ is the same as showing that a point z that does not belong to $\overline{\mathrm{co}}(E(K))$ does not belong to K. Clearly, the closure of $\mathrm{co}(E(K))$ is compact and convex. So, if z does not belong to the closure, then according to Theorem 3.25 there is a continuous linear functional ℓ such that

$$\mathrm{Re}\,\ell(y) < c < \mathrm{Re}\,\ell(z) \quad \text{for all } y \text{ in } \mathrm{co}(E(K)).$$

Since K is compact and ℓ continuous, $\mathrm{Re}\ell$ achieves its maximum over K on some closed subset M of K, and M is an extreme subset of K. According to part (i') noted above, M contains some extreme point p of K. Since p belongs to $E(K)$, and so to $\operatorname{co}(E(K))$, it follows that $\ell(p) < c$. since by construction $\operatorname{Re}\ell(p) = \max_{x\in K} \operatorname{Re}\ell(x)$, $\operatorname{Re}\ell(x) \leq \operatorname{Re}\ell(p) < c$ for all x in K. Since by $\operatorname{Re}\ell(z) > c$, this proves that z does not belong to K. $\square$

Theorem 3.45 (Stone-Weierstrass Theorem)**.** *Let S be a compact Hausdorff space, $C(S)$ the set of all real-valued continuous functions on S, equipped with the supremum norm. Then $C(S)$ ia a Banach algebra. Let E be a subalgebra of $C(S)$, that is,*

(i) E is a linear subspace of $C(S)$.

(ii) The product of two functions in E belongs to E.

In addition we impose the following conditions on E :

(iii) E separates points of S, that is, given any pair of points p and $q, p \neq q$, there is a function f in E such that $f(p) \neq f(q)$.

(iv) All constant functions belong to E.

Conclusion: E is dense in $C(S)$.

The classical Weierstrass theorem is a special case of this proposition, with S an interval of the x axis, and E the set of all *polynomials* or *trigonometric polynomials* in x. We present Louis de Branges's elegant proof, based on the Krein-Milman theorem, of Stone's generalization of the Weierstrass theorem.

Proof. According to the spanning criterion, E is dense in $C(S)$ if the only bounded linear functional ℓ on $C(S)$ that is zero on E is the zero functional. According to the Riesz-Kakutani representation theorem (see Example 5.6), the bounded linear functionals on $C(S)$ are of the form

$$\ell(f) = \int_S f dv$$

v a signed measure of finite total variation $\|v\| = \int d|v|$. So what we have to show is that if $\int_S f dv = 0$ for all f in $E, v = 0$.

Suppose not; denote by U the set of signed measures of finite total mass is ≤ 1 that annihilate all functions in E. This is a convex set, and according to Theorem 5.14, compact in the weak-star topology. So according to the Krein-Milman theorem, if U contained a nonzero measure, it would contain a nonzero extreme point; call it μ. Since μ is extreme, $\|\mu\| = 1$.

Since E is an algebra, if f and g belong to E, so does gf. Since μ annihilates every function in E,

$$\int (fg) d\mu = 0\,.$$

It follows that the measure $gd\mu$ also annihilates every function in E. Let g be a function in E whose values lie between 0 and 1, i.e., $0 < g(p) < 1$ for all p in S. Denote

$$a = \|g\mu\| = \int gd|\mu|\ ,\ b = \|(1-g)\mu\| = \int (1-g)d|\mu|\,.$$

Clearly a and b are positive. Add them:

$$a + b = \int |d\mu| = 1$$

The identity

$$\mu = a\frac{g\mu}{a} + b\frac{(1-g)\mu}{b}$$

represents μ as a nontrivial convex combination of $g\mu/a$ and $(1-g)\mu/b$, both points in U. Since μ is an extreme point, μ must be equal to $g\mu/a$.

Define the support of the measure μ to be the set of points p that have the property that $\int_N |d\mu| > 0$ for any open set N containing p. If $\mu = g\mu/a$, it follows that g has the same value at all points of the support of μ

We claim that the support of μ consists of a single point. To see this, suppose that both p and $q, p \neq q$, belong to the support μ. since the functions in E separate points of S, there is a function h in $E, h(p) \neq h(q)$. Adding a large enough constant to h and dividing it by another large constant, we obtain a function g whose values lie between 0 and 1, and $g(p) \neq g(q)$. This contradicts our previous conclusion.

A measure μ whose support consists of a single point p, and $\|\mu\| = 1$, is a unit point mass at p. Therefore

$$\int f d\mu = f(p) \text{ or } -f(p)$$

since, by hypothesis, the constant 1 belong to E, $\int f d\mu \neq 0$ for $f \equiv 1$ in E, a contradiction.

$\square$

3.9 Vector-Valued Integration

Sometimes it is desirable to be able to integrate functions f that are defined on some measure space $(\Omega, \mathcal{F}, \mu)$ and whose values lie in some topological vector space X. The first problem is to associate with these data a vector in X that deserves to be called

$$\int_\Omega f d\mu$$

i.e., which has at least some of the properties that integrals usually have. For instance, the equation

$$\Lambda\left(\int_\Omega f d\mu\right) = \int_\Omega (\Lambda f) d\mu$$

ought to hold for every $\Lambda \in X^*$, because it does hold for sums, and because integrals are (or ought to be) limits of sums in some sense or other. In fact, our definition will be based on this single requirement.

Many other approaches to vector-valued integration have been studied in great detail; in some of these, the integrals are defined more directly as limits of sums (see Exercise 3.1).

Definition 3.4. Let $(\Omega, \mathcal{F}, \mu)$ be a measure space, let X be a topological vector space on which X^* separates points. Let f be a function from Ω into X such that the scalar functions Λf are integrable with respect to μ, for every $\Lambda \in X^*$. If there exists a vector $y \in X$ such that

$$\Lambda y = \int_\Omega (\Lambda f) d\mu \quad \text{for every } \Lambda \in X^*,$$

then we define

$$\int_\Omega f d\mu := y\,.$$

Remark 3.7. It is clear that there is at most one such y, because X^* separates points on X. Thus there is no uniqueness problem.

Existence will be proved only in the rather special case (sufficient for many applications) in which Ω is compact, f is continuous and X is a Fréchet space. In that case, $f(\Omega)$ is compact, and the closed convex hull of $f(\Omega)$ should be compact by Theorem 3.38. Recall that a Borel measure on a compact (or locally compact) Hausdorff space Ω is a measure defined on the σ-algebra of all Borel sets in Ω this is the smallest σ-algebra that contains all open subsets of Ω. A probability measure is a positive measure of total mass 1.

Theorem 3.46 (Existence of Vector-Valued Integrations)**.** *Let X be a Fréchet space, and μ is a Borel probability measure on a compact Hausdorff space Ω. If $f : \Omega \to X$ is continuous, then the integral*

$$y = \int_\Omega f d\mu$$

exists in the sense of Definition 3.4. Moreover, $y \in \overline{co}(f(\Omega))$.

Remark 3.8. If v is any positive Borel measure on Ω, then some scalar multiple of v is a probability measure. The theorem therefore holds (except for its last sentence) with v in place of μ. It can then be extended to real-valued Borel measures (by the Jordan decomposition theorem and (if the scalar field of X is $\mathbb{C}$) to complex ones.

Proof. We have to prove that there exists $y \in \overline{\mathrm{co}}(f(\Omega))$ such that

$$\Lambda y = \int_\Omega (\Lambda f) d\mu \quad \text{for every } \Lambda \in X^*\,. \tag{3.20}$$

Let $L = \{\Lambda_1, \dots, \Lambda_n\}$ be a finite subset of X^*. Let E_L be the set of all $y \in \overline{\mathrm{co}}(f(\Omega))$ that satisfy (3.20) for every $\Lambda \in L$. Each E_L is closed (by the continuity of Λ) and is therefore compact, since $\overline{\text{co}}(f(\Omega))$ is compact by Theorem 3.38. If no E_L is empty, the collection of all E_L has the finite intersection property. The intersection of all E_L is therefore not empty, and any y in it satisfies (3.20) for every $\Lambda \in X^*$. It is therefore enough to prove $E_L \neq \emptyset$. In other words, there exists $y \in \overline{\text{co}}(f(\Omega))$ with

$$\Lambda_i y = \int_\Omega \Lambda_i f d\mu \quad (1 \leq i \leq n) .$$

Without loss of generality, regard X as a real vector space. Regard $L = (\Lambda_1, \dots, \Lambda_n)$ as a (continuous) mapping from X into $\mathbb{R}^n$. Define

$$m_i = \int_\Omega (\Lambda_i f)\, d\mu \quad (1 \leq i \leq n) . \tag{3.21}$$

It suffices to show that $m = (m_1, \dots, m_n) \in L(\overline{\text{co}}(f(\Omega)))$. In fact,

$$m \in L(\text{co}(f(\Omega))) = \text{co}(L(f(\Omega))) .$$

Put $K = L(f(\Omega))$, then K is a compact subset of $\mathbb{R}^n$. If $m \notin \text{co}(K)$, by Theorem 3.40, Hahn-Banach theorem on $\mathbb{R}^n$ and the known form of the linear functionals on $\mathbb{R}^n$, there are real numbers $c_1, \dots, c_n$ such that

$$\sum_{i=1}^n c_i u_i < \sum_{i=1}^n c_i m_i$$

if $u = (u_1, \dots, u_n) \in K$. Hence

$$\sum_{i=1}^n c_i \Lambda_i f(\omega) < \sum_{i=1}^n c_i m_i \quad \omega \in \Omega .$$

Since μ is a probability measure, integration of the left side gives $\sum c_i m_i < \sum c_i m_i$, that is a contradiction! This completes the proof. $\square$

The following proposition is intuitive and easy.

Proposition 3.47. *Suppose Ω is a compact Hausdorff space, X is a Banach space, $f : \Omega \to X$ is continuous, and μ is a positive Borel measure on Ω. Then*

$$\left\| \int_\Omega f d\mu \right\| \leq \int_\Omega \|f\| d\mu$$

Proof. Let $y = \int_\Omega f d\mu$, then by Corollary 1.32,

$$\|y\| = \sup_{\Lambda \in X^*, \|\Lambda\| \le 1} |\Lambda y| = \sup_{\Lambda \in X^*, \|\Lambda\| \le 1} \left| \int_\Omega \Lambda f d\mu \right| .$$
$$\le \sup_{\Lambda \in X^*, \|\Lambda\| \le 1} \int_\Omega |\Lambda f| d\mu \le \sup_{\Lambda \in X^*, \|\Lambda\| \le 1} \|\Lambda\| \int_\Omega \|f\| d\mu = \int_\Omega \|f\| d\mu .$$

We complete the proof. □

Theorem 3.48. *Let X be a Fréchet space. Let Ω be a compact subset of X. Let y be a vector in X. Then $y \in \overline{\mathrm{co}}(\Omega)$ if and only if there is a regular Borel probability measure μ on Ω such that*

$$y = \int_\Omega x \, \mu(dx) .$$

Remark 3.9. The integral is to be understood as in Definition 3.4 with $f(x) = x$. The integral represents every $y \in \overline{\mathrm{co}}(\Omega)$ as a "weighted average" of Ω, or as the "center of mass" of a certain unit mass distributed over Ω.

Recall that a positive Borel measure on Q is said to be *regular* if

$$\mu(E) = \sup\{\mu(K) : K \subset E\} = \inf\{\mu(G) : E \subset G\}$$

for every Borel set $E \subset Q$, where K ranges over the compact subsets of E and G ranges over the open supersets of E.

Proof. Regard X again as a real vector space. Let $C(\Omega)$ be the Banach space of all real continuous functions on Ω, with the supremum norm. The Riesz representation theorem identifies the dual space $C(\Omega)^*$ with the space of all real Borel measures on Ω that are differences of regular positive ones. With this identification in mind, we define a mapping

$$\phi : C(\Omega)^* \to X \; ; \; \phi(\mu) = \int_\Omega x d\mu(x) .$$

Let P be the set of all regular Borel probability measures on Ω. The theorem asserts that $\phi(P) = \overline{\mathrm{co}}(\Omega)$. For each $x \in \Omega$, the unit mass δ_x concentrated at x

belongs to P, since $\phi(\delta_x) = x$, we see that $\Omega \subset \phi(P)$. Since ϕ is linear and P is convex, it follows that $\mathrm{co}(\Omega) \subset \phi(P)$, where $\mathrm{co}(\Omega)$ is the convex hull of Ω. By Theorem 3.46, $\phi(P) \subset \overline{\mathrm{co}}(\Omega)$. Therefore all that remains to be done is to show that $\phi(P)$ is closed in X. This is a consequence of the following two facts:

(i) P is weak*-compact in $C(Q)^*$. (It's easy to show that P is weak*-closed, then use Theorem 5.14.)

(ii) The mapping ϕ is continuous if $C(Q)^*$ is given its weak*-topology and if X is given its weak topology.

Once we have (i) and (ii), it follows that $\phi(P)$ is weakly compact, hence weakly closed, and since weakly closed sets are strongly closed, we have the desired conclusion. $\square$

Exercise 3.1. Suppose μ is a Borel probability measure on a compact Hausdorff space Ω, X is a Fréchet space, and $f : \Omega \to X$ is continuous. A partition of Ω is, by definition, a finite collection of disjoint Borel subsets of Ω whose union is Ω. Prove that to every neighborhood V of 0 in X there corresponds a partition $\{E_i\}$ such that the difference

$$z = \int_{\Omega} f d\mu - \sum_i \mu(E_i) f(s_i)$$

lies in V for every choice of $s_i \in E_i$. (This exhibits the integral as a strong limit of " Riemann sums.") Suggestion: Take V convex and balanced. If $\Lambda \in X^*$ and if $|\Lambda x| \leq 1$ for every $x \in V$, then $|\Lambda z| \leq 1$, provided that the sets E_i are chosen so that $f(s) - f(t) \in V$ whenever s and t lie in the same E_i

3.10 Holomorphic Functions

In the study of Banach algebras, as well as in some other contexts, it is useful to enlarge the concept of holomorphic function from complex-valued ones to vector-valued ones. (Of course, one can also generalize the domains, by going from $\mathbb{C}$ to $\mathbb{C}^n$ and even beyond. But this is another story.) The most important examples in spectral theory are operator valued holomorphic functions. There are at least two very natural definitions of "holomorphic" available in this general setting, a "weak" one and a "strong" one. They turn out to define the same class of functions if the values are assumed to lie in a Fréchet space.

Let Ω be an open set in $\mathbb{C}$ and let X be a complex topological vector space.

(a) A function $f : \Omega \to X$ is said to be *weakly holomorphic* in Ω if Λf is holomorphic in the ordinary sense for every $\Lambda \in X^*$.

(b) A function $f : \Omega \to X$ is said to be *strongly holomorphic* in Ω if

$$\lim_{w \to z} \frac{f(w) - f(z)}{w - z}$$

exists (in the topology of X) for every $z \in \Omega$. Note that the above quotient is the product of the scalar $(w - z)^{-1}$ and the vector $f(w) - f(z)$ in X.

The continuity of the functionals Λ that occur in (a) makes it obvious that every strongly holomorphic function is weakly holomorphic. The converse is true when X is a Fréchet space, but it is far from obvious. (Recall that weakly convergent sequences may very well fail to converge originally.) The Cauchy theorem will play an important role in this proof, as will Theorem 3.18.

The index of a point $z \in \mathbb{C}$ with respect to a closed path Γ that does not pass through z will be denoted by $\mathrm{Ind}_\Gamma(z)$. We recall that

$$\mathrm{Ind}_\Gamma(z) = \frac{1}{2\pi i} \int_\Gamma \frac{d\zeta}{\zeta - z}\,.$$

All paths considered here and later are assumed to be piecewise continuously differentiable, or at least rectifiable.

Theorem 3.49. *Let Ω be open in $\mathbb{C}$, let X be a complex Fréchet space, and assume that $f : \Omega \to X$ is weakly holomorphic. The following conclusions hold:*

(a) f is strongly continuous in Ω.

(b) The Cauchy theorem and the Cauchy formula hold: If Γ is a closed path in Ω such that $\mathrm{Ind}_\Gamma(w) = 0$ for every $w \notin \Omega$, then

$$\int_\Gamma f(\zeta)d\zeta = 0$$

and

$$f(z) = \frac{1}{2\pi i} \int_\Gamma \frac{f(\zeta)}{\zeta - z} d\zeta$$

if $z \in \Omega$ *and* $\mathrm{Ind}_\Gamma(z) = 1$. *If* Γ_1 *and* Γ_2 *are closed paths in* Ω *such that*

$$\mathrm{Ind}_{\Gamma_1}(w) = \mathrm{Ind}_{\Gamma_2}(w)$$

for every $w \notin \Omega$, *then*

$$\int_{\Gamma_1} f(\zeta) d\zeta = \int_{\Gamma_2} f(\zeta) d\zeta$$

(c) f *is strongly holomorphic in* Ω.

The integrals in (b) are to be understood in the sense of Theorem 3.46. Either one can regard $d\zeta$ as a complex measure on the range of Γ (a compact subset of $\mathbb{C}$), or one can parametrize Γ and integrate with respect to Lebesgue measure on a compact interval in $\mathbb{R}$.

The following extension of Liouville's theorem concerning bounded entire functions does not even depend on Theorem 3.31. It can be used in the study of spectra in Banach algebras.

Theorem 3.50. *Let* X *be a complex Fréchet space. Suppose* $f : \mathbb{C} \to X$ *is weakly holomorphic and* $f(\mathbb{C})$ *is a weakly bounded subset of* X. *Then* f *is constant.*

Proof. For every $\Lambda \in X^*, \Lambda f$ is a bounded (complex-valued) entire function. If $z \in \mathbb{C}$, it follows from Liouville's theorem that

$$\Lambda f(z) = \Lambda f(0) \,.$$

Since X^* separates points on X, this implies $f(z) = f(0)$, for every $z \in \mathbb{C}$. ☐

Chapter 4

Fundermental Principles of Bounded Operators

4.1 Convergence of Operaters

Let X and Y be normed linear spaces over the same scalar field $\mathbb{F}$. We shall to define several types of convergence on the continuous operators from X to Y.

Definition 4.1. Let $A, A_n (n \geq 1)$ be a sequence of operators in $\mathcal{B}(X, Y)$.

(a) We sya that $\{A_n\}$ *converges uniformly* to A if

$$\|A_n - A\| \to 0\,. \tag{4.1}$$

In this case, A is called the *uniform limit* of the sequence $\{A_n\}$.

(b) We say that $\{A_n\}$ *converges strongly* to A if $A_n x$ converges strongly to Ax in Y for each $x \in X$, i.e.,

$$\|A_n x - Ax\| \to 0\,. \tag{4.2}$$

In this case, A is called the *strong limit* of the sequence $\{A_n\}$, denoted by $A = s\text{-}\lim A_n$ or $A_n \xrightarrow{s} A$.

(c) We say that $\{A_n\}$ *converges weakly* to A if $A_n x$ converges weakly to Ax in Y for each $x \in X$, i.e.,

$$\langle A_n x, y^* \rangle \to \langle Ax, y^* \rangle \quad \text{for each } y^* \in Y^* . \tag{4.3}$$

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. However, 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 on $\mathcal{B}(X,Y)$. In fact, we can induce some topologies on $\mathcal{B}(X,Y)$ to describe this serval convergence.

(a) the operator norm $\|\cdot\|$ determines a natural norm topology $\mathcal{T}$ (or uniform topology) on $\mathcal{B}(X,Y)$.

(b) 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 operator topology, denoted by $\mathcal{T}_s$. clearly, $\mathcal{T}_s$ is the weakest topology with respect to which $A \mapsto Ax$ is continuous on $\mathcal{B}(X,Y)$ for all $x \in X$, and $A_n \xrightarrow{s} A$ iff $A_n \to A$ in $\mathcal{T}_s$.

(c) 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 weak operator topology, denoted by $\mathcal{T}_w$. Clearly, $\mathcal{T}_w$ 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 $y^* \in Y^*$, and $A_n \xrightarrow{w} A$ iff $A_n \to A$ in $\mathcal{T}_w$.

Evidently, $\mathcal{T}_w \subset \mathcal{T}_s \subset \mathcal{T}$. Moreover, when $Y = \mathbb{F}$, the scalars field, $\mathcal{B}(X,Y)$ is exactly X^*, then it's easy to find the $\mathcal{T}_s, \mathcal{T}_w$ coincides with $\sigma(X^*, X)$, the weak* topology on X.

4.2 Principle of Uniform Boundedness

Let $\mathcal{A}$ be a subset of $\mathcal{B}(X,Y)$. That is $\mathcal{A}$ is a family of continuous linear operators from X to Y.

(a) We say that $\mathcal{A}$ is *uniformly bounded* or *norm bounded* , if

$$\sup_{A\in\mathcal{A}} \|A\| < \infty\,.$$

(b) We say that $\mathcal{A}$ is *pointwise bounded* or *bounded for the strong operator topology*, if for each $x \in X$,

$$\sup_{A\in\mathcal{A}} \|Ax\| < \infty\,,$$

i.e., $\mathcal{A}$ is bounded with respect to the strong topology $\mathcal{T}_s$.

(c) We say that $\mathcal{A}$ is *bounded for the weak operator topology*, if for each $x \in X$ and $y^* \in Y^*$,

$$\sup_{A\in\mathcal{A}} |\langle Ax, y^*\rangle| < \infty\,,$$

i.e., $\mathcal{A}$ is bounded with respect to the weak topology $\mathcal{T}_w$.

Obviously, $\mathcal{A}$ is uniformly bounded implies that it is pointwise bounded; and $\mathcal{A}$ is pointwise bounded implies that it is bounded for the weak operator topology. Surprisingly, in the case that X is a Banach space, these three types of boundedness is equivalent. Let's see the famous principle of uniform boundedness first, which asserts that pointwise boundedness is equivalent to uniform boundedness.

Theorem 4.1 (Principle of Uniform Boundedness)**.** *Let X be a Banach space and let Y a normed space. If $\mathcal{A} \subseteq \mathcal{B}(X, Y)$ is pointwise bounded, then $\mathcal{A}$ is uniformly bounded.*

Proof. For each $k \geq 1$, 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 $\operatorname{int}(E_m) \neq \emptyset$.

That is, there is an $x_0 \in E_m$ and an $\epsilon > 0$ such that $B(x_0, 2\epsilon) \subset E_m$. Then for any $\|x\| \leq 1$, $x_0 + \epsilon x \in B(x_0, 2\epsilon)$, so

$$\|A(x_0 + \epsilon x)\| \leq m\,, \quad \text{for all } A \in \mathcal{A}\,.$$

Thus

$$\|Ax\| = \frac{\|Ax_0\| + m\|}{\epsilon} \leq \frac{2m}{\epsilon}\,, \quad \text{for all } A \in \mathcal{A}\,.$$

So $\mathcal{A}$ is uniformly bounded. $\square$

Taking $Y = \mathbb{F}$, we get the following corollary:

Corollary 4.2. *X is a Banach space and $N \subset X^*$, 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.

To see the relationship between weak boundedness and uniform boundedness in $\mathcal{B}(X, Y)$, we shall to use PUB carefully. It should be emphasized again that for any $x \in X$, we can regard x as a linear functional on X^*:

$$x^* \mapsto \langle x, x^* \rangle \ ; \ X^* \to \mathbb{F} .$$

In order to ensure rigor, denote by Jx this functional. It is easy to verify that $Jx \in X^{**}$, and $|Jx| = \|x\|$. So using PUB to X^*, we get :

Lemma 4.3. *Let X be a normed vector space. $N \subset X^*$, 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.

Using PUB twice, we get the equivalence between boundedness for weak operator topology and for norm topology.

Theorem 4.4. *Let X be a Banach space and let Y be a normed space. If $\mathcal{A} \subset \mathcal{B}(X, Y)$ is weakly bounded, i.e., 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.

Example 4.3. This example shows that the hypothesis that X is complete cannot be removed in PUB. Consider the space $X = \ell_0$ with the supremum norm. We have show that it is not complete, but is a linear subspace of ℓ^∞ whose closure $\bar{\ell_0} = c_0$. Define the linear operators $A_n : \ell_0 \to \ell_0$ and $A : \ell_0 \to \ell_0$ by

$$A_n x := (x_1, 2x_2, \ldots, nx_n, 0, 0, \ldots) , \quad Ax := (nx_n)_{n \geq 1}$$

for $n \geq 1$ and $x = (x_n)_{n\geq 1} \in \ell_0$. Then $Ax = \lim_{n\to\infty} A_n x$ for every $x \in X$ and $\|A_n\| = n$ for every $n \geq 1$. Thus the sequence $\{A_n x\}_{n\geq 1}$ is bounded for every $x \in X$, the linear operator A is not bounded, and the sequence A_n converges strongly to A.

Example 4.4 (Bilinear Map)**.** Let X, Y and Z be vector spaces. Let $B : X \times Y \to Z$. Associate to each $x \in X$ and to each $y \in Y$ the mappings

$$B_x : Y \to Z \quad \text{and} \quad B^y : X \to Z$$

by defining

$$B_x(y) = B(x, y) = B^y(x)\,.$$

B is said to be *bilinear* if every B_x and every B^y are linear. If X, Y, Z are normed vector spaces and if every B_x and every B^y is continuous, then B is said to be *separately continuous.* If B is continuous (relative to the product topology of $X \times Y$) then B is obviously separately continuous. In certain situations, the converse can be proved with the aid of the PUB.

Let X be a Banach space and Y, Z be normed linear space. Then the followsing propositions are equivalent.

(a) B is *bounded*, i.e. there is a constant $C \geq 0$ such that

$$\|B(x, y)\| \leq C\|x\|\|y\|\,, \quad \text{for all } x \in X \text{ and all } y \in Y\,.$$

(b) B is continuous.

(c) B is obviously separately continuous.

Appications. We will use the PUB to deduce some results about strong convergence.

Proposition 4.5. *Let X be a Banach space and let Y be a normed space. Let $A, A_n \in \mathcal{B}(X, Y)$. If A_n converges to A strongly, then $\{A_n\}$ is uniformly bounded and,*

$$\|A\| \leq \liminf_{n\to\infty} \|A_n\|\,. \tag{4.6}$$

Proof. It follows from PUB that $\{A_n\}$ is uniformly bounded. Notice that for each $x \in X$,

$$\|Ax\| = \lim_{n\to\infty} \|A_n x\| \le \liminf_{n\to\infty} \|A_n\| \|x\| \,.$$

Thus $\|A\| \le \liminf_n \|A_n\|$. $\square$

The following lemma is quite useful:

Lemma 4.6. *Let X be a Banach space and let Y be a normed space. Let $\{A_n\}$ be a sequence in $\mathcal{B}(X, Y)$. Then $\{A_n\}$ is strongly convergent iff*

(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

$$Ax \coloneqq \lim_{n\to\infty} A_n x \,.$$

A is linear, clearly. A is continuous by the proof of Proposition 4.5. $\square$

Theorem 4.7 (Banach-Steinhaus)**.** *Let X, Y be two Banach spaces. Let $\{A_n\}$ be a sequence in $\mathcal{B}(X, Y)$. Then $\{A_n\}$ is strongly convergent iff*

(a) $\{A_n\}$ *is uniformly bounded and,*

(b) there exists a dense subset D of X so that $\{A_n x\}$ converges for each $x \in D$.

Proof. It suffices 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 $\|x - x'\| \le \epsilon$. Then

$$\begin{aligned}\|A_{n+p}x - A_n x\| &\le \|A_{n+p}x - A_{n+p}x'\| + \|A_{n+p}x' - A_n x'\| + \|A_n x' - A_n x\| \\ &\le 2\sup_n \|A_n\|\epsilon + \|A_{n+p}x' - A_n x'\|\end{aligned}$$

Since $\{A_n x'\}$ is a Cauchy sequence, it's easy to find that $\{A_n x\}$ is a Cauchy sequence. Then the desired result follows. $\square$

4.3 Open Mappings and Closed Graphs

4.3.1 Open Mapping Theorem

Let X and Y be two normed linear spaces. Let $A : X \to Y$ be a linear operator. We call A an *open mapping* if $A(U)$ is open in Y whenever U is open in X.

Theorem 4.8 (Open Mapping Theorem)**.** *Let X, Y be Banach spaces. Let $A \in \mathcal{B}(X,Y)$. If A is surjective, i.e., $R(A) = Y$, 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 have only to prove that $AB_X(0,1)$ is open. To this end, it must be shown that there is constant $r > 0$ such that $B_Y(0,r) \subset AB_X(0,1)$. However, this is also sufficient, in deed, for any $y = Ax \in AB_X(0,1)$, take $\epsilon > 0$ such that $\|x\| + \epsilon < 1$, then

$$B_Y(y,\epsilon) = Ax + B_Y(0,\epsilon) \subset AB_X(0,1)\,,$$

which implies that $AB_X(0,1)$ is open.

Step 2. We shall show that 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,

$$Y = AX = A\left(\bigcup_{n=1}^{\infty} nB_X(0,1)\right) = \bigcup_{n=1}^{\infty} nAB_X(0,1) = \bigcup_{n=1}^{\infty} n\overline{AB_X(0,1)}\,.$$

By Baire's category theorem, there is $m \geq 1$ such that $\left(m\overline{AB_X(0,1)}\right)^{\circ} \neq \emptyset$. This implies that $\left(\overline{AB_X(0,1)}\right)^{\circ} \neq \emptyset$. Hence, there is a constant $r > 0$ and an element $y_0 \in Y$ such that $B_Y(y_0, 2r) \subset \overline{AB_X(0,1)}$. Since $y_0 \in \overline{AB_X(0,1)}$, it follows, by symmetry, that $-y_0 \in \overline{AB_X(0,1)}$. Therefore

$$B_Y(0,2r) = B_Y(y_0,2r) - y_0 \subset \overline{AB_X(0,1)} + \overline{AB_X(0,1)}\,.$$

Since $\overline{TB_x(0,1)}$ is convex, $\overline{AB_X(0,1)}+\overline{AB_X(0,1)} = 2\overline{AB_X(0,1)}$. Hence, $B_Y(0,2r) \subset 2\overline{AB_X(0,1)}$ and, consequently, $B_Y(0,r) \subset \overline{AB_X(0,1)}$.

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

$$B_Y(0,r/2^n) \subset \overline{AB_X(0,1/2^n)}\,,\ n \geq 1.$$

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 $B_Y(0,r/2) \subset AB_X(0,1)$. $\square$

Remark 4.1. Conversely, it's evident to see that if $A : X \to Y$ is an open mapping, then A is surjective.

From the open mapping theorem, we get the following theorem directly.

Theorem 4.9 (Inverse Mapping Theorem)**.** *Let X, Y be Banach spaces. Let $A \in \mathcal{B}(X,Y)$. If A is bijective, then $A^{-1} \in \mathcal{B}(Y,X)$. In other words, A is a linear homomorphism.*

In fact, the inverse mapping theorem and the open mapping theorem are equivalent. To see this, suppose $A \in \mathcal{B}(X,Y)$ is a surjection. Observe from Example 1.20 and that

$$\tilde{A} : X/N(A) \to Y\,;\, \tilde{x} \mapsto Ax$$

is bijective continous linear operator from $X/N(A)$ onto Y. By Theorem 1.17, $X/N(A)$ is Banach space. So the inverse mapping theorem implies that

$$\tilde{A}^{-1} : Y \to X/N(A)\,;\, Ax \mapsto \tilde{x} \quad \text{is continuous}\ .$$

Then we show that there exists $r > 0$ so that $B_Y(0,r) \subset AB_X(0,1)$. Notice that for any $y = Ax \in B_Y(0,r)$

$$\|\tilde{x}\| = \|\tilde{A}^{-1}\tilde{A}\tilde{x}\| \leq \|\tilde{A}^{-1}\|\|y\| \leq \|\tilde{A}^{-1}\|r\,.$$

Let $\|\tilde{A}^{-1}\|r < 1$, then $\|\tilde{x}\| \leq 1$, so there exist $m \in N(A)$ so that $\|x+m\| < 1$ and $Ax = A(x+m) \in AB_X(0,1)$.

As a consequence of the inverse mapping theorem we will see that when X is a Banach space with respect to two different norms then determining whether those two norms are equivalent is simpler.

Theorem 4.10 (Equivalence of Norms on Banach Spaces)**.** *Let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two norms defined on X and let X be a Banach space with respect to both of these norms. Suppose that there exists a constant $C > 0$ such that for all $x \in X$ we have that $\|x\|_2 \leq C\|x\|_1$. Then $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent.*

Proof. Consider the identity mapping $I : (X, \|\cdot\|_1) \to (X, \|\cdot\|_2)$. Trivially, I is a continuous linear bijection from X onto itself. By the inverse mapping theorem, $I^{-1} = I : (X, \|\cdot\|_2) \to (X, \|\cdot\|_1)$ is bounded, which implies the desired result. □

Example 4.5. Let $(X, \|\cdot\|)$ be a Banach space and let $M, N \subset X$ be two closed linear subspaces such that

$$X = M \oplus N$$

i.e., $M \cap N = \{0\}$ and every vector $x \in X$ can be written as $x = m + n$ with $m \in M$ and $n \in N$. Then it follows from Corollary 4.10 that there exists a constant $C \geq 0$ such that

$$\|m\| + \|n\| \leq C\,\|m+n\|$$

for all $m \in M$ and all $n \in N$. In other words, the canonical mapping from $(M \oplus N, \|\cdot\|)$ onto $M \oplus_1 N$, $m + n \mapsto (m, n)$ is a linear homeomorphism.

Example 4.6. This example shows that the hypothesis that X and Y are complete cannot be removed in the open mapping theorem and the inverse mapping theorem. Let $X = \ell_0$, equipped with the supremum norm. Thus X is a normed vector space but is not a Banach space. Define the operator $A : X \to X$ by $Ax := \left(k^{-1}x_k\right)_{k\in\mathbb{N}}$ for $x = (x_k)_{k\in\mathbb{N}} \in X$. Then A is a bijective bounded linear operator but its inverse is unbounded.

Example 4.7. Here is an example where X is complete and Y is not. Let $X = Y = C([0,1])$ be the space of continuous functions $f : [0,1] \to \mathbb{R}$ equipped with the norms

$$\|f\|_X := \sup_{0 \le t \le 1} |f(t)|, \quad \|f\|_Y := \sqrt{\int_0^1 |f(t)|^2 dt}$$

Then X is a Banach space, Y is a normed vector space, and the identity

$$I : X \to Y$$

is a bijective bounded linear operator with an unbounded inverse.

4.3.2 Closed Graph Theorem

Let X and Y be two linear spaces. Let A be a linear operator on $D(A) \subset X$ into Y. The *graph* of A, denoted by $G(A)$, is the subset of $X \times Y$ given by

$$G(A) = \{(x, Ax) : x \in X\}.$$

Obviously, $G(A)$ is a linear subspace of $X \times Y$.

Let X and Y be normed linear spaces on the same scalar field. It follows from Section **??** that we can equip the product space $X \times Y$ a norm, so that the projections $(x, y) \mapsto x$ and $(x, y) \to y$ are continuous. Specifically, the norm defined by

$$\|(x, y)\| := \|x\| + \|y\| \text{ for all } x \in X, y \in Y$$

satisfies the desired properties. We write $X \oplus_1 Y$ for $X \times Y$ endowed with this norm. Clearly, if X, Y are Banach space, then so is $X \oplus_1 Y$.

The graph norm of A on the linear subspace $D(A) \subset X$ is the norm function $D(A) \to [0, \infty) : x \mapsto \|x\|_A$ defined by

$$\|x\|_A := \|x\| + \|Ax\|$$

for $x \in D(A)$. Note that a linear operator $A : X \supset D(A) \to Y$ is always a continuous linear operator with respect to the graph norm.

Definition 4.2. Let X and Y be normed linear spaces, and A a linear operator from $D(A) \subset X$ into Y. Then A is called a *closed linear operator* if it's graph $G(A)$ is a closed subspace of $X \oplus_1 Y$.

Evidently, a linear operator $A : X \supset D(A) \to Y$ is closed if and only if for any sequence $\{x_n\}$ in $D(A)$ so that $x_n \to x \in X$ and $Ax_n \to y \in Y$, we have $x \in D(A)$ and $y \in R(A)$. Therefore, if A is a continuous lienar operator and $D(A)$ is a closed subspace of X, then A is closed. When Y is a Banach space, every continuous lienar operator can be regarded as a closed operator in the followsing sense.

Lemma 4.11. *Let X be a normed vector space and let Y be a Banach space. Let $A : X \supset D(A) \to Y$ be a continuous linear operator. Then there exists a unique continuous linear operator $\bar{A} : \overline{D(A)} \to Y$ so that*

$$\bar{A}|_{D(A)} = A \quad \text{and} \quad \|\bar{A}\| = \|A\| . \tag{4.7}$$

Proof. For any $x \in \overline{D(A)}$, there exists a sequence $\{x_n\}$ in $D(A)$ so that $x_n \to x$. Since A is a bounded linear operator on $D(A)$, $\{Ax_n\}$ is a Cauchy sequence in Y. By the completeness of Y, $\{Ax_n\}$ is convergent in Y. Define

$$\bar{A}x \coloneqq \lim_{n\to\infty} Ax_n .$$

Clearly, $\bar{A}$ is well-defined linear operator on $\overline{D}(A)$ with $\bar{A}|_{D(A)} = A$ and $\|\bar{A}\| = \|A\|$. We noe complete the proof. ☐

We will give an example of a discontinuous but closed operator.

Example 4.8. Let $X = C[0,1]$ endowed with sup-norm. Let $D = C^1[0,1] \subset C[0,1]$ and let T be the linear operator on D into X defined by

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

This T is not continuous, since, for $x_n(t) = t^n$, $\|x_n\| = 1$, but $\|Tx_n\| = n$ for $n \geq 1$. However, T is closed. In fact, let $\{x_n\} \subset D(T)$, $x_n \to x$ and $Tx_n \to y$. a Then $x_n'(t)$ converges uniformly to $y(t)$, and $x_n(t)$ converges uniformly to $x(t)$.

Hence $x(t)$ must be differentiable with continuous derivative $y(t)$. This proves that $x \in D(T)$ and $Tx = y$.

Besides, the graph norm of T on $D(T) = C^1[0,1]$ agrees with the usual C^1-norm

$$\|f\|_{C^1} = \sup_{0\le t\le 1} |f(t)| + \sup_{0\le t\le 1} |f'(t)| \qquad \text{for } f \in C^1[0,1]$$

and $C^1[0,1]$ is a Banach space with this norm.

Theorem 4.12 (Closed Graph Theorem)**.** *Let X and Y be Banach spaces. Let A be a linear operator from $D(A) \subset X$ into Y. If $D(A)$ is a closed subspace of X and A is closed operator, then A is acontinuous lienar operator.*

Proof. Define the projection

$$P : G(A) \to D(A) \subset X\,;\ (x, Ax) \mapsto x\,.$$

It is easy to check that P is a continuous bijection between $G(A)$ and $D(A)$. Since $G(A)$ is the closed subspace of Banach space $X \oplus_1 Y$, and $D(A)$ is a closed subspace of Banach space X, both $G(A)$ and $D(A)$ are Banach space. By the inverse mapping theorem, $P^{-1} : D(A) \to G(A)$ is continuous. Thus $A : D(A) \to Y$ is the composition of the continuous map $P^{-1} : D(A) \to G(A)$ and the continuous map of $G(A) \to Y$ defined by $(x, Ax) \mapsto Ax$. Therefore, A is continuous too. □

In fact, the closed graph theorem is equivalent to the inverse mapping theorem (or the open mappting theorem). The proof is trivial. Let $A \in \mathcal{B}(X,Y)$ be a bijection. It suffices to show that A^{-1} is a closed operator. Let $y_n = Ax_n \to y$ and $x_n = A^{-1}y_n \to x$, then it follows from the continuity of A that $Ax = \lim_n Ax_n = \lim_n y_n = y$. So the desired result follows.

Example 4.9 (Hellinger-Toeplitz Theorem)**.** Let H be a Hilbert space over $\mathbb{F}$. Let $A : H \to H$ be a selfadjoint linear operator, i.e.,

$$\langle x, Ay\rangle = \langle Ax, y\rangle \qquad \text{for all } x, y \in H\,.$$

Then A is continuous. To see this, it suffices to prove that A has a closed graph. Thus assume that $\{x_n\}$ is a sequence in H and $x, y \in H$ are vectors such that

$$\lim_{n\to\infty} x_n = x, \quad \lim_{n\to\infty} Ax_n = y\,.$$

In order that $Ax = y$, notice that for any $z \in H$,

$$\langle y, z\rangle = \lim_{n\to\infty} \langle Ax_n, z\rangle = \lim_{n\to\infty} \langle x_n, Az\rangle = \langle x, Az\rangle = \langle Ax, z\rangle\,.$$

Then the desired result follows.

Example 4.10 (Douglas Factorization)**.** Let X, Y, Z be Banach spaces and let $A : X \to Y$ and $B : Z \to Y$ be bounded linear operators. Assume A is injective. Then the following are equivalent.

(a) $R(B) \subset R(A)$.

(b) There exists a continuous linear operator $T : Z \to X$ such that $AT = B$.

We now use the closed graph theorem to prove this. If (b) holds, then $R(B) = R(AT) \subset R(A)$. Conversely, suppose that $R(B) \subset R(A)$ and define

$$T := A^{-1} \circ B : Z \to X$$

Then T is a linear operator and $AT = B$. We prove that T has a closed graph. To see this, let $(z_n)\, d$ be a sequence in Z such that the limits

$$z := \lim_{n\to\infty} z_n, \quad x := \lim_{n\to\infty} Tz_n$$

exist. Then

$$Ax = \lim_{n\to\infty} ATz_n = \lim_{n\to\infty} Bz_n = Bz$$

and hence $x = Tz$. Thus T has a closed graph and $D(T) = Z$, by the closed graph theorem, hence T is continuous.

Closeable operators. For a linear operator that is defined on a proper linear subspace it is an interesting question whether it can be extended to a linear operator with a closed graph. Such linear operators are called *closeable.* So, any continuous lienar operator A form $D(A) \subset X$ into Y is closeable, if Y is complete.

Lemma 4.13 (Characterization of Closeable Operators)**.** *Let X and Y be Banach spaces, let $D(A) \subset X$ be a linear subspace, and let $A : D(A) \to Y$ be a linear operator. Then the following are equivalent.*

(i) A is closeable.

(ii) If $\{x_n\}$ is a sequence in $D(A)$ and $y \in Y$ is a vector such that $x_n \to 0$ and $Ax_n \to y$, then $y = 0$.

(iii) The projection $\pi_X : G(A) \to X$, $(x, y) \mapsto x$ onto the first factor is injective.

Proof. (i) implies (ii) is trivial. We prove that (ii) implies (iii). The closure of any linear subspace of a normed vector space is again a linear subspace. Hence graph (A) is a linear subspace of $X \times Y$ and the projection $\pi_X : G(A) \to X$ is a linear map by definition. By (ii) the kernel of this linear map is the zero subspace and hence it is injective.

We prove that (iii) implies (i). Define

$$D\left(\bar{A}\right) := \pi_X(\overline{G(A)}) \subset X$$

This is a linear subspace and the map $\pi_X : \overline{G(A)} \to D\left(\bar{A}\right)$ is bijective by (iii). Denote its inverse by $\pi_X^{-1} : D\left(\bar{A}\right) \to \overline{G(A)}$ and denote by

$$\pi_Y : \overline{G(A)} \to Y$$

the projection onto the second factor. Then

$$\bar{A} := \pi_Y \circ \pi_X^{-1} : D\left(\bar{A}\right) \to Y\,;\, (x, y) \mapsto y$$

is a continuous linear operator, since π_X and π_Y are continuous. Its graph is the linear subspace

$$G\left(\bar{A}\right) = \overline{G(A)} \subset X \times Y$$

and $\bar{A}|_{D(A)} = A$ holds because $G(A) \subset G\left(\bar{A}\right)$. $\square$

Example 4.11 (Symmetric Operators)**.** Let H be a Hilbert space over $\mathbb{F}$ and let $A : D(A) \to H$ be a linear operator, defined on a *dense* linear subspace $D(A) \subset H$. Suppose A is selfadjoint, i.e.

$$\langle x, Ay \rangle = \langle Ax, y \rangle \quad \text{for all } x, y \in D(A) .$$

Then A is closeable. To see this, choose a sequence $\{x_n\}$ in $D(A)$ such that $x_n \to 0$ and the sequence $Ax_n \to y \in H$. Then for all $z \in D(A)$,

$$\langle y, z \rangle = \lim_{n\to\infty} \langle Ax_n, z \rangle = \lim_{n\to\infty} \langle x_n, Az \rangle = 0 .$$

Since $D(A)$ is a dense subspace of H, we deduce that $y = 0$. Thus A is closeable.

Example 4.12 (Differential Operators)**.** This example shows that *evry differential operator is closeable.* Let $\Omega \subset \mathbb{R}^n$ be a nonempty open set, fix a constant $1 < p < \infty$, and consider the Banach space $X := L^p(\Omega)$. Then the space

$$D(A) := C_0^\infty(\Omega)$$

of smooth functions $u : \Omega \to \mathbb{R}$ with compact support is a dense linear subspace of $L^p(\Omega)$. Let $m \geq 1$ and, for every multiindex $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$ with $|\alpha| = \alpha_1 + \cdots + \alpha_n \leq m$, let $\phi_\alpha : \Omega \to \mathbb{R}$ be a smooth function. Define the operator $A : C_0^\infty(\Omega) \to L^p(\Omega)$ by

$$Au := \sum_{|\alpha| \leq m} \phi_\alpha \partial^\alpha u \tag{4.8}$$

Here the sum runs over all multi-indices $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n$ with $|\alpha| \leq m$ and

$$\partial^\alpha = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}.$$

We prove that A is closeable. To this end, define the constant $1 < q < \infty$ by $1/p + 1/q = 1$ and define the *formal adjoint* of A as the operator $B : C_0^\infty(\Omega) \to L^q(\Omega)$, given by

$$Bv := \sum_{|\alpha| \leq m} (-1)^{|\alpha|} \partial^\alpha \left(\phi_\alpha v\right)$$

for $v \in C_0^\infty(\Omega)$. Then integration by parts shows that

$$\int_\Omega v(Au) = \int_\Omega (Bv)u \tag{4.9}$$

for all $u, v \in C_0^\infty(\Omega)$. Now let $u_k \in C_0^\infty(\Omega)$ be a sequence of smooth functions with compact support and let $v \in L^p(\Omega)$ such that

$$\lim_{k\to\infty} \|u_k\|_{L^p} = 0, \quad \lim_{k\to\infty} \|v - Au_k\|_{L^p} = 0$$

Then, for every test function $\phi \in C_0^\infty(\Omega)$, we have

$$\int_\Omega \phi v = \lim_{k\to\infty} \int_\Omega \phi\,(Au_k) = \lim_{k\to\infty} \int_\Omega (B\phi)u_k = 0$$

since $C_0^\infty(\Omega)$ is dense in $L^q(\Omega)$, this implies that

$$\int_\Omega \phi v = 0 \quad \text{for all } \phi \in L^q(\Omega)$$

Now take $\phi := \mathrm{sign}(v)|v|^{p-1} \in L^q(\Omega)$ to obtain $\int_\Omega |v|^p = 0$ and hence v vanishes almost everywhere. Hence the linear operator A is closeable, as claimed.

Example 4.13. Let $H = L^2(\mathbb{R})$. Define $\Lambda : C_c(\mathbb{R}) \to \mathbb{R}$ by

$$\Lambda(f) := f(0)\,.$$

This linear functional is NOT closeable because there exists a sequence of continuous functions $f_n : \mathbb{R} \to \mathbb{R}$ with compact support such that $f_n(0) = 1$ and $\|f_n\|_{L^2} \le \frac{1}{n}$ for all $n \in \mathbb{N}$.

EXERCISE

Exercise 4.1. Let X, Y be Banach spaces and let $A : X \supset D(A) \to Y$ be a linear operator. Then, A is a closed linear operator if and only if $D(A)$ is a Banach space with respect to the graph norm.

Exercise 4.2. Let $A \in \mathcal{B}(X, Y)$. Show that $A^{-1} : R(A) \to X$ exists and is continuous, i.e., A is an linear homeomorphism between X and $R(A)$ if and noly if there is a constant $m > 0$ such that

$$\|Ax\| \ge m\|x\| \quad \text{for each } x \in X\,.$$

In this case, if X is a Banach space then $R(A)$ is closed.

Exercise 4.3 (Open Mapping Theorm for Closed Operators). Let X and Y be Banach spaces. Let $A : X \supset D(A) \to Y$ be a closed lienar operator. Show that if $R(A)$ is of the second category in Y, then A is an open mapping form $D(A)$ onto Y. In particular, $R(A) = Y$.

4.4 Projections, Direct Sum Decomposition

Let X be a normed vector space. We say $P \in \mathcal{B}(X)$ is called a *projection* if

$$P^2 = P\,.$$

Proposition 4.14.

(a) *P is an projection if and only if $I - P$ is an projection.*

(b) *$N(P) = R(I-P)$, $R(P) = N(I-P)$ and both $R(P)$ and $N(P)$ are closed subspaces of X.*

(c) *$X = R(P) \oplus N(P)$.*

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, $N(P)$ is a closed subspace of X. Also,

$$x \in R(P) \Leftrightarrow Px = x \Leftrightarrow (I-P)x = 0 \Leftrightarrow x \in N(I-P).$$

Similarly, $R(I-P) = NP$.

To show (c), note that $x = Px + (I-P)x$. □

From this proposition, we say P is the projection from X to $R(P)$. Moreover, M is a closed subspace of X, we say there exists a projection P from X to M, if there exists projection P such that $R(P) = M$. Clearly, if X is a Hilbert space such a projection exists but is not determined uniquely since the "orthogonality" is not asked. When X is a Banach space, does such projections exist?

Theorem 4.15. *Let X be a Banach space. If M, N are two closed subspace of X, such that $X = M \oplus N$. Then there is a projection P satisfying*

$$R(P) = M \text{ and } N(P) = N .$$

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

$$Px \coloneqq m .$$

P is linear, since the composition is unique. Clearly, $P^2 = P$. It follows from Example 4.5 that P is bounded. So the desired result follows.

Also, we can sue the closed graph theorem. It suffices to show P has a closed graph. Assume $x_n \to x$ and $Px_n \to m$, then

$$x_n - Px_n \to x - m .$$

Since M, N are closed, $m \in M$ and $x - m \in N$. Thus we have $Px = m$. So P is bounded. □

Definition 4.3. 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 4.16. *M is a closed subspace of normed space X. If*

$$\dim M < \infty \quad \text{or} \quad \operatorname{codim} M \coloneqq \dim(X/M) < \infty ,$$

there is a projection P form X to M, i.e., M is complemented.

Proof. *Case 1.* Assume 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, $c_j(\cdot)$ is continuous linear functional on M. By Hahn-Banach extension theorem, it can be extended on X as a continuous linear functional. Then define $P : X \to X$ by

$$Px = \sum_{i=1}^{n} c_j(x)e_j \, ,$$

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 $e_j \in \widetilde{e}_j$, 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. □

Schauder Bases. Let X be a separable Banach space and let $(e_j)_{j\geq 1}$ be a *Schauder basis* of X. Recall that this means that, for each element $x \in X$, there exists a *unique* sequence $(\alpha_j(x))$ of scalars such that

$$\lim_{n\to\infty} \left\| x - \sum_{i=1}^{n} \alpha_j(x)e_j \right\| = 0 \tag{4.10}$$

Let $n \geq 1$ and define the map $\Pi_n : X \to X$ by

$$\Pi_n(x) := \sum_{i=1}^{n} \alpha_j(x) e_j$$

for $x \in X$, where $(\alpha_j(x))$ is the unique sequence that satisfies (4.10). Clearly, the operators $\Pi_n : X \to X$ are linear and satisfy

$$\Pi_n \Pi_m = \Pi_m \Pi_n = \Pi_{n \wedge m}$$

for all integers $n, m \geq 1$. In particular, $\Pi_n^2 = \Pi_n$ for all $n \geq 1$. If $\{\Pi_n\}$ were projections, since $\Pi_n \to I$ in the strong operator topology, it follows from the PUB that $\{\Pi_n\}$ is uniformly bounded. Let's check if this is true.

Theorem 4.17. *Define a map $X \to [0, \infty) : x \mapsto \|x\|'$ by the formula*

$$\|x\|' := \sup_{n \geq 1} \|\Pi_n(x)\| \qquad \textit{for all } x \in X\,.$$

Then $(X, \|\cdot\|')$ is a Banach space with $\|x\|' \leq \|x\|$ for all $x \in X$. Hence by Theorem 4.10, there exists a constant $C > 0$ such that

$$\sup_{n \in \mathbb{N}} \|\Pi_n(x)\| \leq C\|x\| \qquad \textit{for all } x \in X\,.$$

In particular, $\{\Pi_n\}$ are projections.

Proof. It's easy to check that $\|\cdot\|'$ is a norm on X and

$$\|x\| = \lim_{n \to \infty} \|\Pi_n(x)\| \leq \sup_{n \geq 1} \|\Pi_n(x)\| = \|x\|'\,. \tag{4.11}$$

It suffices to show that $(X, \|\cdot\|')$ is complete. Let (x_k) be a Cauchy sequence in $(X, \|\cdot\|')$. Then (x_k) is a Cauchy sequence in $(X.\|\cdot\|)$ by (4.11). Suppose

$$\lim_{k \to \infty} \|x_k - x\| = 0\,. \tag{4.12}$$

We have only to show that

$$\lim_{k \to \infty} \|x_k - x\|' = \lim_{k \to \infty} \sup_{n \geq 1} \|\Pi_n(x_k) - \Pi_n(x)\| = 0\,.$$

Firstly, since (x_k) be a Cauchy sequence in $(X, \|\cdot\|')$, for any fixed n, $\{\Pi_n(x_k)\}$ is a Cauchy sequence in $(X, \|\cdot\|)$. Thus there exists a sequence $\{\xi_n\}$ in X with $\|\Pi_n(x_k) - \xi_n\| \to 0$ as $k \to \infty$ uniformly in n; i.e.,

$$\lim_{k\to\infty} \sup_{n\geq 1} \|\Pi_n(x_k) - \xi_n\| = 0\,. \tag{4.13}$$

So it's enough to show that $\xi_n = \Pi_n(x)$ for all n. Clearly $\xi_n \in \operatorname{span}\{e_1, \cdots, e_n\}$. Besides, since Π_m is continuous on $\operatorname{span}\{e_1, \cdots, e_n\}$ for $m \leq n$,

$$\Pi_m \xi_n = \lim_{k\to\infty} \Pi_m \Pi_n(x_k) = \lim_{k\to\infty} \Pi_m(x_k) = \xi_m\,.$$

It's sufficient to show that $\|\xi_n - x\| \to 0$ as $n \to \infty$. To this end, take any $\epsilon > 0$. Notice that for each $n \geq 1$ and each $k \geq 1$,

$$\|\xi_n - x\| \leq \|\xi_n - \Pi_n(x_k)\| + \|\Pi_n(x_k) - x_k\| + \|x_k - x\|\;. \tag{4.14}$$

Combine (4.13) and (4.12), there exists $k_0 = k_0(\epsilon)$ so that

$$\sup_{n\geq 1} \|\xi_n - \Pi_n(x_{k_0})\| \leq \epsilon \text{ and } \|x_{k_0} - x\| \leq \epsilon\,.$$

Recall that $\|\Pi_n(x_{k_0}) - x_{k_0}\| \to 0$ as $n \to \infty$. Taking $k = k_0(\epsilon)$ and letting $n \to \infty$, we get

$$\limsup_{n\to\infty} \|\xi_n - x\| \leq 2\epsilon\,.$$

Since $\epsilon > 0$ is arbitrary, the desired result follows. $\square$

Corollary 4.18. *For each $j \geq 1$, $\alpha_j(\cdot) \in X^*$, where $(\alpha_j(x))$ is the unique sequence that satisfies (4.10).*

Chapter 5

Duality in Banach Spaces

5.1 Dual Spaces

Let X be a norm linear space. Recall that we denote by X^* the set of all continuous linear functionals on X. As a corollary of Proposition 1.5, X^* is Banach space. We call X^* the *dual* of X.

Remark 5.1. It should be emphasized that we did NOT assume X is complete. In fact, if $\bar{X}$ is its completion, then X^* and $\bar{X}^*$ are isometrically isomorphic.

Proposition 5.1. *X^* is isometrically 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 embedding. $\square$

Theorem 5.2. *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

$$\|x_n\| = 1 \quad \text{and} \quad |\langle x_n, x_n^* \rangle| > \frac{1}{2}\,.$$

We claim that $\overline{\operatorname{span}}\{x_n\} = X$. By Corollary 1.35, it suffices to show that $\ell \in X^*$ vanishing at $\{x_n\}$ implies $\ell = 0$. If not, without loss of generality we assume that $\|\ell\| = 1$. Then there exists x_k^* with $\|x_k^* - \ell\| < 1/2$. Hence $|\ell(x_k) - x_k^*(x_k)| < 1/2$. But $\ell(x_k) = 0$, so $|x_k^*(x_k)| < 1/2$, which is a contradiction! $\square$

The converse of Theorem 5.2 is not true.

Example 5.1. ℓ^1 is separable but $(\ell^1)^* = \ell^\infty$(see Example 5.5) 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.

5.1.1 Representation of Dual Spaces

Let X and X' be two metric spaces. Recall that a map $M : X \to X'$ is called an *isometry* if one has

$$d(Mx, My) = d(x,y) \qquad \text{for any } x, y \in X\,.$$

Particularly, if X, X' are normed linear space, and the distances are induced by norm, then M is *isometry* if and only if

$$\|Mx\| = \|x\| \qquad \text{for any } x \in X\,.$$

Obviously, an isometry is automatically injective, and it can not be surjective.

Example 5.2. 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.* Clearly, T is an isometry isomorphism if and only if T is bijective and

$$\|T\| = \|T^{-1}\| = 1\,.$$

In this case, X and Y are said to be *isometrically isomorphic* and we write $X \cong Y$. By abuse of the notation, sometime we write $X = Y$ for short.

Example 5.3. The dual space of c_0 is ℓ_1, i.e., $c_0^* \cong \ell_1$.

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 (e_n), the Schauder basis for c_0, 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

$$f(x) = \sum_{n=1}^{\infty} x_n f\left(e_n\right)\,.$$

Take any $k \in \mathbb{N}$, let

$$x_n^{(k)} = \begin{cases} \left|f\left(e_n\right)\right| / f\left(e_n\right), & n \le k \text{ and } f(e_n) \neq 0\,. \\ 0\,, & \text{otherwise}\,. \end{cases}$$

Then $x^{(k)} = (x_n^{(k)}) \in c_0$ and $\|x^{(k)}\| = 1$. So $f(x^{(k)}) = \sum_{n=1}^{k} |f(e_n)| \le \|f\|$. 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 5.4. $L^p(\mu)^* \cong L^q(\mu)$. 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

$$\ell_g(f) = \int fgd\mu$$

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 5.5. $L^1(\mu)^* \cong L^\infty(\mu)$, where μ is a σ-finite measure. Specifically, (X, Ω, μ) is a σ-finite measure space. For any $g \in L^\infty(X, \Omega, \mu)$, define ℓ_g : $L^1(\mu) \to \mathbb{F}$ by

$$\ell_g(f) = \int fgd\mu \ .$$

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 5.6 (Riesz - Kakutani Representation Theorem)**.** $C_0(X)^* \cong \mathcal{M}(X)$, where X is a locally compact Hausdorff (LCH) space. Specifically, $\mathcal{M}(X)$ denotes the space of all $\mathbb{F}$-valued Radon 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

$$\ell_\mu(f) = \int fd\mu \ .$$

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 $l^1 = M(\mathbb{N})$.

5.1.2 Dual Spaces of Subspaces and Quotients

Let X be a normed linear space. Let M be a closed subspace in X. If $f \in X^*$, then $f|_M$, the restriction of f to M, belongs to M^* and $\|f|_M\| \le \|f\|$. According to the Hahn-Banach theorem, each bounded linear functional on M is obtainable as the restriction of a functional from X^*. Indeed, more can be said.

Let now M be a subset of X. The *annihilator* of M, denoted by $M^\perp$, is given by

$$M^\perp := \{x^* \in X^* : \langle x, x^* \rangle = 0 \text{ for all } x \in M\} .$$

Let N be any subset of X^*, then the *annihilator* of N, denoted by ${}^\perp N$, is given by

$${}^\perp N := \{x \in X : \langle x, x^* \rangle = 0 \text{ for all } x^* \in N\} .$$

It is easy to observe that $M^\perp$ and ${}^\perp N$ are closed subspace of X^* and X.

Lemma 5.3. *M is a subspace of X, N is a subspace of X^*, then*

$${}^\perp(M^\perp) = \overline{M}\,, \ ({}^\perp N)^\perp = \overline{N}^{\sigma(X^*,X)} . \tag{5.1}$$

Proof. Use Hahn-Banach theorem, as in the proof of Theorem 3.36. □

Theorem 5.4. *Let M be a closed linear subspace of X. Then $M^* \cong X^*/M^\perp$. Moreover, the map $\rho : X^*/M^\perp \to M^*$ defined by*

$$\rho : \tilde{f} \mapsto f|_M$$

is an isometric isomorphism.

Proof. Clear ρ is a well-defined linear operator. It follows from the Hahn-Banach theorem that ρ is surjective. So we have only to show

$$\|\tilde{f}\| = \|f|_M\| \quad \text{ for all } \tilde{f} \in X^*/M^\perp .$$

To this end, notice that $\|f|_M\| = \|g|_M\| \le \|g\|$ for all $g \in \tilde{f}$. Besides, by Hahn-Banach theorem, there exists $g \in \tilde{f}$ so that the equality holds. Then the desired result follows. □

Theorem 5.5. *Let M be a closed linear subspace of X. Let $Q : X \to X/M$ be the natural map. Then the map $\rho : (X/M)^* \to M^\perp$ defined by*

$$\rho : \tilde{f} \mapsto \tilde{f} \circ Q$$

is an isometric isomorphism.

Proof. Clear ρ is a well-defined linear operator. We show that ρ is surjective. For any $\ell \in M^\perp$, define $\tilde{\ell} : X/M \to \mathbb{F}$ by

$$\tilde{\ell}\tilde{x} := \ell x \qquad \text{for } \tilde{x} \in X/M\,.$$

By Example 1.20, $\ell \in (X/M)^*$ and $\ell = \tilde{\ell} \circ Q$. Hence, ρ is surjective. Now, it suffices to show that

$$\|\tilde{f}\| = \|\tilde{f} \circ Q\|\,.$$

Since $\|Q\| \le 1$, we have $\|\tilde{f} \circ Q\| \le \|f\|$. On the other hand, take any $\lambda \in (0,1)$. Then we can find $\tilde{x} \in X/M$ with $\|\tilde{x}\| = 1$, satisfying $|\tilde{f}\tilde{x}| \ge \lambda\|\tilde{f}\|$. Take $m \in M$ so that $\|x + m\| \le 1/\lambda$, then

$$\lambda\|\tilde{f}\| \le |\tilde{f}\tilde{x}| = |(\tilde{f} \circ Q)(x+m)| \le \|\tilde{f} \circ Q\|\frac{1}{\lambda}\,.$$

Letting $\lambda \uparrow 1$, the desired result follows. $\square$

5.2 Bidual Space and Reflexivity

Definition 5.1. The dual space of $(X^*, \|\cdot\|)$ is called the *second dual space* or *bidual space* of X, and denote as X^{**}.

Let $(X, \|\cdot\|)$ be a normed linear space over $\mathbb{F}$. For any fixed $x \in X$, define a functional $Jx : X^* \to \mathbb{F}$ by

$$\langle x^*, Jx\rangle = \langle x, x^*\rangle \qquad \text{for all } x^* \in X^*\,.$$

It is easy to verify that $Jx \in X^{**}$, and $|Jx| = \|x\|$. It now follows that we can define a map

$$J : X \to X^{**},\ x \mapsto Jx. \tag{5.2}$$

Obviously, J_X is linear, and hence 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 5.2. $(X, \|\cdot\|)$ is said to be *reflexive* if the canonical embedding is surjective, i.e., $JX = X^{**}$.

Remark 5.2. (a) X is reflexive 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 5.7. Every finite-dimensional normed linear space is reflexive.

Example 5.8. For $1 < p < \infty$, $L^p(\mu)$ is reflexive. See Example 5.4.

Example 5.9. 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

$$|\ell_s - \ell_t| = 2 \quad \text{for } s \ne t$$

since the $\{\ell_s\}$ form a nondenumerable collection, $C[0,1]^*$ cannot contain a dense denumerable subset. It follows now that $C[0,1]^{**} \ne C[0,1]$.

Lemma 5.6. *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 5.7. *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 5.8. *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 $x \in X$,

$$|\langle x, x^* \rangle| = |\langle J_X x, \xi \rangle| \leq \|\xi\| \|x\| \, ,$$

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 1.34 that there is a functional $J_{X^*} x^* \in X^{***}$ such that $\|x^*\| = 1$, and $\langle x^{**}, J_{X^*} x^* \rangle > 0$ and

$$\langle J_X x, J_{X^*} x^* \rangle = 0 \quad \text{for all } x \in X \, .$$

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. □

5.3 Weak and Weak-star 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-star topology on X^*.

Proposition 5.9. *$(X, \|\cdot\|)$ 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 have 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,

$$\bigcap_{x^* \in \Phi} \ker(x^*) \subset B(0,1).$$

Since X is infinite-dimensional, $\cap_{x^* \in \Phi} \ker(x^*) \neq \{0\}$ is a subspace of X, this is a contradiction. □

Note that there are three topologies on X^*: the norm topology, the weak topology, denoted by $\sigma(X^*, X^{**})$, and the weak-star topology, denoted by $\sigma(X^*, X)$.

Proposition 5.10. *Let τ^* denote the norm topology on X^*. Then*

(a) $\sigma(X^*, X) \subset \sigma(X^*, X^{**}) \subset \tau^*$

(b) $\sigma(X^*, X^{**}) = \tau^*$ *if and only if* 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. $\square$

5.3.1 Convergence

In Example 3.1, we have pointed that

- a net $\{x_i\}$ in X is convergent to some $x \in X$ in $\sigma(X, X^*)$ 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 $\sigma(X^*, X)$ 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, $x_n \to x$ in $\sigma(X, X^*)$, or $x_n^* \to x^*$ in $\sigma(X, X^*)$. In order to be totally clear we sometimes emphasize norm convergence by saying, "$x_n \to x$ in norm," meaning that $\|x_n - x\| \to 0$.

Example 5.10. 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 5.11. If a sequence in l^1 converges weakly, it converges in norm (Proposition 5.18). Note that this result demonstrates in a dramatic way that in discussions concerning the weak topology it is essential to consider nets and not just sequences.

Example 5.12. 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, so x_n doesn't converge to 0 in norm.

Generally, H is a Hilbert space and $\{e_n\}$ is a orthonormal system, by Bessel's inequality we know $e_n \xrightarrow{w} 0$ but $\|e_n\| = 1$ for all n.

Proposition 5.11. *Let $\{x_n\}$ be a sequence in the normed linear space X.*

(a) If $x_n \to x$ weakly in $\sigma(X, X^*)$, then $\{x_n\}$ is norm bounded and,

$$\|x\| \le \liminf_{n\to\infty} \|x_n\| . \tag{5.3}$$

(b) If $x_n \to x$ weakly in $\sigma(X, X^*)$ and if $x_n^* \to x^*$ in norm, then

$$\langle x_n, x_n^* \rangle \to \langle x, x^* \rangle . \tag{5.4}$$

Proof. To show (a), by PUB, we have that $\{x_n\}$ is bounded, and note that

$$|\langle x, x^* \rangle| = \lim_{n\to\infty} |\langle x_n, x^* \rangle| \le \liminf_{n\to\infty} \|x_n\| \|x^*\|$$

for any $x^* \in X^*$, Thus

$$\|x\| \le \liminf_{n\to\infty} \|x_n\| .$$

To show (b), note that

$$\begin{aligned} |\langle x_n, x_n^* \rangle - \langle x, x^* \rangle| &\le |\langle x_n, x_n^* \rangle - \langle x_n, x^* \rangle| + |\langle x_n, x^* \rangle - \langle x, x^* \rangle| \\ &\le \|x_n^* - x^*\| \sup_n \|x_n\| + |\langle x_n, x^* \rangle - \langle x, x^* \rangle| . \end{aligned}$$

We now complete the proof. □

By using Banach-Steinhaus theorem, we get the following result, which is useful when we discussing sequential compactness.

Theorem 5.12. *Let X be a normed linear space. A sequence $\{x_n\}$ in X is weakly convergent if and only if*

(a) $\{x_n\}$ is norm bounded and,

(b) there exists a dense subset N of X^* so that $\{\langle x_n, x^*\rangle\}$ converges for each $x^* \in N$.

Theorem 5.13. *Let X be a Banach space. A sequence $\{x_n^*\}$ in X^* is convergent relative to the weak-star topology if and only if*

(a) $\{x_n^*\}$ is uniformly (norm) bounded and,

(b) there exists a dense subset M of X so that $\{\langle x, x_n^*\rangle\}$ converges for each $x \in M$.

5.3.2 Compact and Sequentially Compact Sets in Weak-star Topology

Observe that $X^* \subset \mathbb{F}^X = \prod_X \mathbb{F}$ and that the weak-star topology $\sigma(X^*, X)$ on X^* is the relative topology on X^* induced by the product topology on $\prod_X \mathbb{F}$.

Theorem 5.14 (Banach-Alaoglu Theorem)**.** *Let $(X, \|\cdot\|)$ be a normed vector space. Then the closed unit ball in X^* is weak-star compact, i.e.,*

$$B_{X^*} = \{x^* \in X^* : \|x^*\| \leq 1\}$$

is compact with respect to the topology $\sigma(X^*, X)$.

Proof. Step 1. For each $x \in X$, let

$$D_x = \{\lambda \in \mathbb{F} : |\lambda| \leq \|x\|\}.$$

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 $|f(x)| \leq \|x\|$ 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-star 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

$$|\langle x, x^* \rangle| = \lim_i |\langle x, x_i^* \rangle| \leq \|x\|$$

for all $x \in X$, x^* is continuous, and $\|x^*\| \leq 1$. That is,

$$x^* \in B_{X^*}.$$

Therefore B_{X^*} is closed in D and hence compact. □

Theorem 5.15 (Sequential Compactness)**.** *Let X be a separable Banach space. Then B_{X^*} is weak-star sequentially compact.*

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

$$\langle x_k, y_n \rangle \text{ converges for each } x_k\,.$$

By Theorem 5.13, $\{y_n^*\}$ is converges in weak-star topology. □

Separability and Metrizability. In fact, for a separable Banach space X, the closed unit ball B_{X^*} is metrizable in the weak-star topology. So the weak-star compactness of B_{X^*} implies it's sequential compactness.

We should emphasize that the weak and weak-star topologies on an infinite dimensional Banach space are never metrizable. It is possible, however, to show that under certain conditions these topologies are metrizable when restricted to bounded sets. In applications this is often sufficient.

The next lemma is rather easy, but it will be used so often that it should be explicitly stated and proved.

Lemma 5.16. *Let X, Y be two Hausdorff topological spaces and X is compact. If $f : X \to Y$ is bijective and continuous, then f is a homeomorphism.*

Proof. If F is a closed subset of X, then F is compact. Thus $f(F)$ is compact in Y and hence closed. Since f maps closed sets to closed sets, f^{-1} is continuous. So f is a homeomorphism. $\square$

Theorem 5.17. *If X is a Banach space, then closed unit ball B_{X^*} is weak-star metrizable if and only if X is separable.*

Proof. Assume that X is separable and let $\{x_n\}$ be a countable dense subset of B_{X^*}. For each $n \geq 1$, let $D_n = \{\alpha \in \mathbb{F} : |\alpha| \leq 1\}$. Put $Y = \prod_{n=1}^{\infty} D_n$. Then Y is a compact metric space. So if $(B_{X^*}, \sigma(X^*, X))$ is homeomorphic to a subset of Y, B_{X^*} is weak-star metrizable.

Define $\varphi : B_{X^*} \to X$ by $\varphi(x^*) = \{\langle x_n, x^* \rangle\}$. We show that φ is a homeomorphism. If $\{x_i^*\}$ is a net in B_{X^*} and $x_i^* \to x^*$ in $\sigma(X^*, X)$, then $\langle x_n, x_i^* \rangle \to \langle x_n, x^* \rangle$ for each $n \geq 1$, hence $\varphi(x_i^*) \to \varphi(x^*)$ and φ is continuous. If $\varphi(x^*) = \varphi(y^*)$, $\langle x_n, x^* - y^* \rangle = 0$ for all n, since $\{x_n\}$ is dense, $x^* - y^* = 0$. Thus φ is injective. Since x^* is wk * compact, φ is a homeomorphism onto its image (Lemma 5.16) and B_{X^*} is weak-star metrizable.

Now assume that $(B_{X^*}, \sigma(X^*, X))$ is metrizable. Thus there are open sets $\{U_n\}$ in $(B_{X^*}, \sigma(X^*, X))$ such that $0 \in U_n$ and $\cap_n U_n = \{0\}$. By the definition of the relative weak-star topology on B_{X^*}, for each n there is a finite set F_n contained in X such that

$$\{x^* \in B_{X^*} : |\langle x, x^* \rangle| < 1 \text{ for all } x \text{ in } F_n\} \subset U_n.$$

Let $F = \bigcup_{n=1}^{\infty} F_n$; so F is countable. We claim that $\overline{\operatorname{span}}\{F\} = X$, and hence X is separable. To show this, by Lemma 1.35 it suffices to show that

$$F^{\perp} = \{0\}\,.$$

If $x^* \in F^{\perp}$, then for each $n \geq 1$ and for each x in F_n, $|\langle x, x^*/\left\|x^*\right\|\rangle| = 0 < 1$. Hence $x^*/\left\|x^*\right\| \in U_n$ for all $n \geq 1$; thus $x^* = 0$. We now complete the proof. □

Is there a corresponding result for the weak topology? If X^* is separable, then the weak topology on ball X is metrizable. In fact, this follows from Theorem 5.17 if the embedding of X into X^{**} is considered. This result is not very useful since there are few examples of Banach spaces X such that X^* is separable. Of course if X is separable and reflexive, then X^* is separable, but in this case the weak topology on X^* is the same as its weak-star topology when X is identified with X^{**}. Thus Theorem 5.17 is adequate for a discussion of the weak topology on the unit ball of a separable reflexive space. If $X = c_0$, then $X = l^1$ and this is separable but not reflexive. This is one of the few nonreflexive spaces with a separable dual space.

If X is separable, is $(B_X, \sigma(X, X^*))$ metrizable? The answer is no, as the following result of Schur demonstrates. So if $(B(\ell^1), \sigma(\ell^1.\ell^{\infty}))$ were metrizable, the following proposition would say that the weak and norm topologies on l^1 agree. But this is not the case.

Proposition 5.18. *If a sequence in ℓ^1 converges weakly, it converges in norm.*

Proof. Recall that $\left(l^1\right)^* = l^{\infty}$. Since l^1 is separable, Theorem 5.17 implies that $B(l^{\infty})$ is weak-star metrizable. Indeed, one can check an equivalent metric on $(B(l^{\infty}), \sigma(\ell^{\infty}, \ell^1))$ is given by

$$d(\phi, \psi) = \sum_{j=1}^{\infty} \frac{|\phi(j) - \psi(j)|}{2^j} \qquad \text{for } \phi, \psi \in \ell^{\infty}\,.$$

By Alaoglu's theorem, $B(l^\infty)$ is weak-star compact. Hence $(B(l^\infty), \sigma(\ell^\infty, \ell^1))$ is a complete metric space and the Baire category theorem is applicable.

Let $\{f_n\}$ be a sequence of elements in l^1 such that $f_n \to 0$ weakly. Let $\epsilon > 0$, and for each positive integer m let

$$F_m = \{\phi \in B(l^\infty) : |\langle f_n, \phi\rangle| \le \epsilon \text{ for } n \ge m\} \,.$$

It is easy to see that F_m is weak-star closed in $B(l^\infty)$ and, because $f_n \to 0$ weakly,

$$\bigcup_{m=1}^{\infty} F_m = B(l^\infty)\,.$$

By the theorem of Baire, there exists $m = m_\epsilon$ so that F_m has non-empty weak-star interior. Thus there is a ϕ in F_m and a $\delta = \delta_\epsilon > 0$ such that $\{\psi \in B(l^\infty) : d(\phi, \psi) < \delta\} \subset F_m$. Given any fixed $n \ge m$, define ψ in l^∞ by

$$\psi(j) = \begin{cases} \phi(j)\,, & \text{for } 1 \le j \le J\,; \\ \operatorname{sign}(f_n(j))\,, & \text{for } j > J\,. \end{cases}$$

where J is a large integer that we determine it later. Thus $\psi(j)f_n(j) = |f_n(j)|$ for $j > J$. It is easy to see that $\psi \in B(l^\infty)$. Also,

$$d(\phi, \psi) = \sum_{j>J} 2^{-j}|\phi(j) - \psi(j)| \le 2 \cdot 2^{-J} < \delta\,.$$

for $J = J_\epsilon$. So $\psi \in F_m$ and hence $|\langle \psi, f_n\rangle| \le \epsilon$, i.e.,

$$\left|\sum_{j\le J} \phi(j) f_n(j) + \sum_{j>J} |f_n(j)|\right| \le \epsilon\,. \tag{5.5}$$

Since $f_n \to 0$ weakly, there is an $m' = m'_\epsilon$ such that for $n \ge m'$,

$$\sum_{j\le J} |f_n(j)| < \epsilon\,.$$

Combining this with (5.5) gives that for $n \geq \max\{m_\epsilon, m'_\epsilon\}$,

$$\begin{aligned}
\|f_n\| &= \sum_{j \leq J} |f_n(j)| + \sum_{j > J} |f_n(j)| \\
&< \epsilon + \left| \sum_{j > J} |f_n(j)| + \sum_{j \leq J} \phi(j) f_n(j) - \sum_{j \leq J} \phi(j) f_n(j) \right| \\
&< \epsilon + \left| \sum_{j > J} |f_n(j)| + \sum_{j \leq J} \phi(j) f_n(j) \right| + \sum_{j \leq J} |\phi(j) f_n(j)| \leq 3\epsilon \,.
\end{aligned}$$

So $\|f_n\| \to 0$. The desired result follows. $\square$

5.3.3 Compact and Sequentially Compact Sets in Weak Topology : Reflexivity Revisited

Theorem 5.19 (Goldstine)**.** *Let $(X, \|\cdot\|)$ be a normed linear space. Let J be the natural embedding 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-star 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 the weak-star topology is Hausdorff, so $B_{X^{**}}$ is closed in $\sigma(X^{**}, X^*)$ and

$$\overline{JB_X}^{\sigma(X^{**}, X^*)} \subset B_{X^{**}} \,.$$

On the other hand, if there exists $x_0^{**} \notin \overline{JB_X}^{\sigma(X^{**}, X^*)}$, since $\overline{JB_X}^{\sigma(X^{**}, X^*)}$ is balanced closed convex set in $(X^{**}, \sigma(X^{**}, X^*))$, by Theorem 3.28, there exists $x^* \in X^*$ (use Theorem 3.32 to deduce that $(X^{**}, \sigma(X^{**}, X^*))^* = X^*$) so that

$$|\langle x, x^* \rangle| \leq 1 < |\langle x^*, x_0^{**} \rangle| \text{ , for all } x \in B_X \,.$$

Thus $\|x^*\| \leq 1$, and $\|x_0^{**}\| > 1$. So $x_0^{**} \notin B_{X^{**}}$. Therefore,

$$B_{X^{**}} \subset \overline{JB_X}^{\sigma(X^{**}, X^*)} \,,$$

and the desired result follows. $\square$

Corollary 5.20. *Let X be a normed linear space. Let J be the canonical embedding of X into X^{**}. Then $J(X)$ is dense in X^{**} relative to the weak-star topology $\sigma(X^{**}, X^*)$ on X^{**}. That is,*

$$\overline{JX}^{\sigma(X^{**},X^*)} = X^{**}.$$

Proof. Note that

$$\begin{aligned}\overline{JX}^{\sigma(X^{**},X^*)} &= \overline{\bigcup_{n=1}^{\infty} nJB_X}^{\sigma(X^{**},X^*)} = \bigcup_{n=1}^{\infty} n\overline{JB_X}^{\sigma(X^{**},X^*)} \\ &= \bigcup_{n=1}^{\infty} nB_{X^{**}} = X^{**}.\end{aligned}$$

$\square$

Weak compactness and reflexivity. If X is a reflexive Banach space, then the weak and weak-star topologies agree on its dual space X^*, hence the closed unit ball in X^* is weakly compact by the Banach-Alaoglu theorem and so the closed unit ball in X is also weakly compact. The Eberlein-Šmulyan theorem asserts that this property characterizes reflexivity. It also asserts that weak compactness of the closed unit ball is equivalent to sequential weak compactness.

Theorem 5.21 (Eberlein-Šmulyan Theorem I)**.** *Let X be a normed lienar space. Let B_X be the closed unit ball in X. 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 $\sigma(X, X^*)$. □

Corollary 5.22. *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. □

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 5.23. *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.,

$$M := \overline{\text{span}}\{x_n\}\,.$$

Then M is separable and by Theorem 5.7 (or Exercise 5.1), M is reflexive. Since $M^{**} = M$ is separable, by Theorem 5.2, M^* is separable. Let $\{y_n^*\}$ dense in M^*. By the *diagonal process*, we can select a subspace $\{z_n\}$ of $\{x_n\}$ so that

$$\lim_{n\to\infty} \langle z_n, y_k \rangle \;\text{ exists for each }\; k \geq 1\,.$$

Then it follows from Theorem 5.13 that $\{z_n\}$ is weakly convergent. □

Remarkably, the converse of Theorem 5.23 is also true, namely the following

Theorem 5.24 (Eberlein-Šmulian Theorem II)**.** *Let be X a Banach space and let B_X be the closed unit ball in X. Then B_X is weakly sequentially compact if and only if X is reflexive.*

Proof. The proof is rather delicate and is omitted, see, e.g., *Functional Analysis* by K. Yosida. □

Remark 5.3. In order to clarify the connection between Theorem 5.23 and Theorem 5.24, 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 trivial.

(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 5.25 (Best Approximation)**.** *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 5.4. 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

$$\|k\| \leq \liminf_{n\to\infty} \|k_n\| = d$$

So we get the desired point. □

Exercise 5.1. Use Banach–Alaoglu theorem to 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.

5.4 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

$$\langle Tx, y^* \rangle = \langle x, T^* y^* \rangle \tag{5.6}$$

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* (or *dual*) of T. Moreover, 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 5.1: The adjoint of T

Example 5.13 (Adjoint of Multiplication Operator)**.** Let $(X, \mathcal{F}, \mu)$ and the multiplication operator M_ϕ on L^p, where $1 \leq p < \infty$, be as in Example 1.16. 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 other words, the adjoint of multiplication operator with symbol ϕ on L^p is the multiplication operator with symbol ϕ on L^q. To show this, notice that

$$\langle M_\phi f, g\rangle = \int (M_\phi f) g d\mu = \int \phi f g d\mu = \int f(M_\phi^* g) d\mu = \langle f, M_\phi^* g\rangle \, .$$

Example 5.14 (Adjoint of Integral Operator)**.** Let the integral operator K on L^p, where $1 \leq p < \infty$, and kernel k be as in 1.18 or as in Example 1.17 (p=2). Let $\frac{1}{p} + \frac{1}{q} = 1$, then $K^* : L^q \to L^q$ is the integer operator with kernel $k^*(x, y) \coloneqq k(y, x)$. To see this, using Fubini theorem, we have

$$\begin{aligned} \langle Kf, g\rangle &= \int \left[\int k(x,y) f(y) \mu(dy) \right] g(x) \mu(dx) \\ &= \int \left[\int K(x,y) g(x) \mu(dx) \right] f(y) \mu(dy) = \langle f, K^* g\rangle \end{aligned}$$

for all $f \in L^p$ and $g \in L^q$.

Example 5.15. Let X be a locally compact Hausdorff space and let $\phi : X \to X$ be a homeomorphism. Let $T : C_0(X) \to C_0(X)$ be the operator defined by

$Tf := f \circ \phi$ for $f \in C_0(X)$ (the *pullback* of f under ϕ). Then, under the identification $C_0(X)^* \cong \mathcal{M}(X)$ of the dual space of $C_0(X)$ with the space of $\mathbb{F}$-valued Radon measures with finite total variations on X, the dual operator of T is the operator $T^* : \mathcal{M}(X) \to \mathcal{M}(X)$, which assigns to every measure $\mu : \mathcal{B}(X) \to \mathbb{F}$ its *pushforward* $T^*\mu = \phi_*\mu$ under ϕ. This pushforward is given by $(\phi_*\mu)(B) := \mu\left(\phi^{-1}(B)\right)$ for every Borel set B in X. To see this, note that

$$\langle Tf, \mu\rangle = \int f \circ \phi d\mu = \int f d\phi_*\mu = \langle f, T^*\mu\rangle$$

for all $f \in C_0(X)$ and $\mu \in \mathcal{M}(X)$.

Proposition 5.26. *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

$$\|T\| = \sup_{x \in B_X, y^* \in B_{Y^*}} |\langle Tx, y^*\rangle| = \sup_{x \in B_X, y^* \in B_{Y^*}} |\langle x, T^*y^*\rangle| = \|T^*\|\,,$$

where B_X and B_{Y^*} is the closed unit ball of X and Y^*, respectively. □

Remark 5.5. From this property, we have that $T \mapsto T^{**}$ is an linear isometric embedding form $\mathcal{B}(X,Y)$ to $\mathcal{B}(X^{**}, Y^{**})$. Besides, we can regard T^{**} as a continuous extension of T on X^{**}; indeed,

$$T^{**} \circ J_X = J_Y \circ T\,. \tag{5.7}$$

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 5.27 (Duality)**.** *Let X and Y be normed linear spaces. Let $T \in \mathcal{B}(X,Y)$. Then the following hold.*

(a) $R(T)^{\perp} = N(T^*)$ *and* ${}^{\perp}R(T^*) = N(T)$.

(b) $\overline{R(T)} = {}^{\perp}N(T^*)$ *and* $\overline{R(T^*)}^{\sigma(X^*,X)} = N(T)^{\perp}$.

(c) *T is injective iff $R(T^*)$ is weak*-dense in X^*; $R(T)$ is dense in Y iff T^* is injective.*

Proof. Note that

$$\begin{aligned} y^* \in R(T)^{\perp} &\Leftrightarrow \langle Tx, y^* \rangle = 0 \text{ for all } x \\ &\Leftrightarrow \langle x, T^* y^* \rangle = 0 \text{ for all } x \Leftrightarrow T^* y^* = 0\,; \end{aligned}$$

and

$$\begin{aligned} x \in {}^{\perp}R(T^*) &\Leftrightarrow \langle x, T^* y^* \rangle = 0 \text{ for all } y^* \\ &\Leftrightarrow \langle Tx, y^* \rangle = 0 \text{ for all } y^* \Leftrightarrow Tx = 0\,, \end{aligned}$$

then (a) follows. Trivially, (b), (c) follows from (a). $\square$

Example 5.16. Define the operator $A : \ell^2 \to \ell^2$ by $Ax := \left(\frac{x_k}{k}\right)$ for $x = (x_k) \in \ell^2$. This operator is self-adjoint, i.e., $A = A^*$, injective, and has a dense image, but is not surjective. Thus $R(A) \subsetneqq \ell^2 =^{\perp} N(A^*)$.

Example 5.17. Let X be a normed vector space, let $M \subset X$ be a closed linear subspace, and let $Q : X \to X/M$ be the canonical projection. Then the dual operator $Q^* : (X/M)^* \to X^*$ is the isometric embedding of Theorem 5.5 whose image is $M^{\perp}$. The dual operator of the inclusion $i : M \to X$ is a surjective operator $i^* : X^* \to M^*$ with kernel $M^{\perp}$. It descends to the isometric isomorphism $X^*/M^{\perp} \to M^*$ in Theorem 5.4.

Many of the nontrivial properties of adjoints depend on the completeness 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 5.28. *X and Y are Banach spaces, and $T \in \mathcal{B}(X, Y)$. Then T is bijective if and only if T^* is bijective, and in this case*

$$(T^{-1})^* = (T^*)^{-1} . \tag{5.8}$$

Particularly, T is an isometry isomorphism if and only if T^* is an isometry isomorphism.

Proof. If T is a bijection, 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 bijective, and $(T^*)^{-1} = (T^{-1})^*$.

If T^* is bijective, by the preceding proof then T^{**} is, too. By the inverse mapping theorem, T^{**} is a linear homeomorphism, thus T is a 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 Theorem 5.27

$$R(T) = {}^{\perp}N(T^*) = Y . \qquad \square$$

An example of a Banach space isometry isomorphism is the pullback under a homeomorphism $\phi : X \to X$ of a LCH space, acting on the space of continuous functions on X vanishing at infity, equipped with the supremum norm. Its dual operator is the pushforward under ϕ, acting on the space of $\mathbb{F}$-valued Radon measures on X (see Example 5.15).

5.4.1 Closed Range Theorem

The main theorem of this subsection asserts that a continuous linear operator between two Banach spaces has a closed image if and only if its dual operator has a closed image. A key tool in the proof is that a bounded linear operator

$T : X \to Y$ between Banach spaces is surjective if and only if it's an open mapping.

Lemma 5.29. *Let X, Y be Banach spaces. Let $T \in \mathcal{B}(X, Y)$. Then T is surjective if and only if T^* is a linear homeomorphism from Y^* onto $R(T^*)$.*

Proof. Necessity. By Theorem 5.27 we have $N(T^*) = R(T)^\perp = \{0\}$. So T^* is a injection. To show $(T^*)^{-1} : R(T^*) \to Y^*$ is continuous, by Exercise 4.2, it suffices to show that there exists some $M > 0$ so that

$$M\|T^*y^*\| \geq \|y^*\|\,, \text{ for all } y^* \in Y^*\,.$$

That is

$$M \sup_{x \in T(B_X)} |\langle x, y^*\rangle| \geq \sup_{y \in B_Y} |\langle y, y^*\rangle|\,, \text{ for all } y^* \in Y^*\,, \tag{5.9}$$

where B_X and B_Y are the closed unit ball in X and Y, respectively. By the open mapping theorem, $T : X \to Y$ is an open mapping, so there exists $r > 0$, such that

$$rB_Y \subset T(B_X)\,.$$

Therefore,

$$\sup_{y \in B_Y} |\langle y, y^*\rangle| = \frac{1}{r} \sup_{y \in rB_Y} |\langle y, y^*\rangle| \leq \frac{1}{r} \sup_{x \in T(B_X)} |\langle x, y^*\rangle|$$

for all $y^* \in Y^*$, and hence (5.9) follows.

Sufficiency. We show that T is an open mapping, which implies that 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 balanced and convex. By Theorem 3.28, for any $y \notin \overline{T(B_X(0,1))}$, there exists some $y^* \in Y^*$ so that

$$|\langle Tx, y^*\rangle| \leq 1 < \langle y, y^*\rangle\,, \text{ for any } x \in B_X(0,1)\,.$$

Therefore,

$$\|T^*y^*\| < \langle y, y^*\rangle \leq \|y\|\|y^*\|$$

Since $(T^*)^{-1}$ is continuous, there exises some $M > 0$ so that

$$M\|T^*y^*\| \geq \|y^*\|\,, \text{ for all } y^* \in Y^*\,.$$

Thus

$$\|y\| \geq \frac{1}{M}\,,$$

and we get

$$B_Y(0, 1/M) \subset \overline{T(B_X(0,1))}\,.$$

We now complete the proof. □

The following consequence is useful in applications.

Corollary 5.30 (Surjection I)**.** *Let X, Y be Banach spaces. Let $T \in \mathcal{B}(X, Y)$. Then the following are equivalent.*

(a) $R(T) = Y$.

(b) There exists some $M > 0$ so that $M\|T^*y^*\| \geq \|y^*\|$ for all $y^* \in Y^*$.

(c) $N(T^*) = \{0\}$ *and* $R(T^*)$ *is norm-closed.*

Proof. Observe that T^* is a linear homeomorphism from Y^* onto $R(T^*)$ is equivalent to (b) and (c), the desired result follows. □

Observe that $N(T^*) = \{0\}$ implies $\overline{R(T)} = Y$. In fact, $R(T^*)$ is norm-closed implies that $R(T)$ is closed, and hence $R(T) = Y$. This is the so called closed range theorem.

Theorem 5.31 (Closed Range Theorem)**.** *Let X, Y be Banach spaces. Let $T \in \mathcal{B}(X, Y)$. Then each of the following three conditions implies the other two:*

(a) $R(T)$ *is closed in* Y ;

(b) $R(T^*)$ *is weak***-closed in* X^* ;

(c) $R(T^*)$ *is norm-closed in* X^* .

Proof. (a) $\Rightarrow$ (b). Notice that

$$\overline{R\left(T^{*}\right)}^{\sigma(X^{*},X)}=N(T)^{\perp}\,,$$

so it suffices to show that $R\left(T^{*}\right)=N(T)^{\perp}$; i.e., for any $x^{*}\in N(T)^{\perp}$, 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 $R(T)$ by

$$\langle Tx,y^{*}\rangle=\langle x,x^{*}\rangle\quad\text{for all } x\in X\,.$$

One can show that y^{*} is well-defined since $x^{*}\in N(T)^{\perp}$, and y^{*} is bounded linear functional on $R(T)$. The desired result follows from Hahn-Banach extension theorem.

(b) $\Rightarrow$ (c). This follows trivially from that the weak-star topology is weaker than norm topology.

(c) $\Rightarrow$ (a). Without loss of generality, let $\overline{R\left(T\right)}=Y$, or we let $Z=\overline{R\left(T\right)}$ and let $S\in\mathcal{B}(X,Z)$ with $Sx=Tx$ for each x, then $R\left(S\right)=R\left(T\right)$. By Hahn-Banach extension theorem, one can see that $R\left(S^{*}\right)=R\left(T^{*}\right)$.

Since $R(T)^{\perp}=N(T^{*})$, ${}^{\perp}N(T^{*})=\overline{R\left(T\right)}=Y$, hence

$$\overline{N\left(T^{*}\right)}^{\sigma(Y^{*},Y)}=N(T^{*})=\{0\}\,.$$

Thus by Theorem 5.30, $R\left(T\right)=Y$. ☐

Exercise 5.2. Let X and Y be real normed vector spaces and let $T:X\to Y$ be a bounded linear operator. Let $x^{*}\in X^{*}$. Show that the following are equivalent.

(a) $x^{*}\in R\left(T^{*}\right)$.

(b) There is a constant $c\geq 0$ such that $|\langle x^{*},x\rangle|\leq c\|Tx\|$ for all $x\in X$.

Exercise 5.3 (Surjection II). Let X, Y be Banach spaces. Let $T\in\mathcal{B}(X,Y)$. Show that the following are equivalent.

(a) $R(T^*) = X^*$.

(b) T is a linear homeomorphism form X onto $R(T)$.

(c) There exists some $M > 0$ so that $M\|Tx\| \geq \|x\|$ for all $x \in X$.

(d) $N(T) = \{0\}$ and $R(T)$ is norm-closed.

5.5 Uniformly Convex Space

A norm is called *strictly subadditive* if

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

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 $\psi(r)$ defined for positive r in $[0, 2]$,

$$\psi(r) > \psi(0) = 0, \text{ for any } r > 0\,, \quad \text{and} \quad \lim_{r \downarrow 0} \psi(r) = 0 \tag{5.10}$$

such that for all x, y such that $\|x\| \leq 1$, $\|y\| \leq 1$, the inequality

$$\left\| \frac{x+y}{2} \right\| \leq 1 - \psi(\|x - y\|) \tag{5.11}$$

holds.

Definition 5.3. A normed vector space $(X, \|\cdot\|)$ is called *uniformly convex*, if the norm satisfies (5.11) for all vectors x, y in B_X, where $\psi(r)$ is some function satisfying (5.10).

Some other textbooks defined uniformly convex space as follows : For 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 $\|x-y\| \geq \epsilon$.

The uniform convexity is a geometric property of the unit ball: if we slide a rule of length $\epsilon > 0$ 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 can not include any line segment.

Example 5.18. 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 5.2.

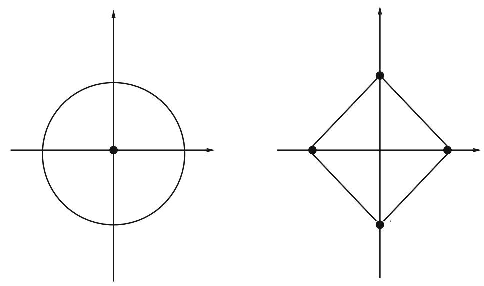

Figure 5.2: Unit ball B_X for $\|\cdot\|_2$ and $\|\cdot\|_1$

Example 5.19. L^p space is uniformly convex for $1 < p < \infty$. In fact, Clarkson's first inequality yields that for $2 \leq p < \infty$,

$$\left\|\frac{f+g}{2}\right\|_p^p + \left\|\frac{f-g}{2}\right\|_p^p \leq \frac{1}{2}\left(\|f\|_p^p + \|g\|_p^p\right) \quad \forall f, g \in L^p .$$

Besides, Clarkson's second inequality yields that for $1 < p \leq 2$:

$$\left\|\frac{f+g}{2}\right\|_p^q + \left\|\frac{f-g}{2}\right\|_p^q \leq \left(\frac{1}{2}\|f\|_p^p + \frac{1}{2}\|g\|_p^p\right)^{q/p} \forall f, g \in L^p .$$

Strong convergence and weak convergence. We conclude with a useful property of uniformly convex spaces.

Theorem 5.32. *Let X be a uniformly convex space. Let $x, x_n (n = 1, 2, \cdots)$ in X. Then $\{x_n\}$ converges strongly to x if and only if $\|x_n\| \to \|x\|$ and $\{x_n\}$ converges weakly to x.*

Proof. Without loss of generality, we assume $\|x\| > 0$, and $\|x_n\| > 0$ for all n. Note that

$$\epsilon\left(\left\|\frac{x_n}{\|x_n\|} - \frac{x}{\|x\|}\right\|\right) \le 1 - \frac{1}{2}\left\|\frac{x_n}{\|x_n\|} + \frac{x}{\|x\|}\right\| .$$

To show $\|x_n - x\| \to 0$, it suffices to show that

$$\epsilon\left(\left\|\frac{x_n}{\|x_n\|} - \frac{x}{\|x\|}\right\|\right) \to 0\,, \tag{5.12}$$

since $\epsilon(\cdot)$ is incresing, continuous at 0, $\epsilon(r) > 0$ for any positive r, and $\|x_n\| \to \|x\|$. To show (5.12), 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

$$\|x\| \le \liminf_{n\to\infty}\left\|\frac{x_n + x}{2}\right\|$$

On the other hand,

$$\limsup_{n\to\infty}\left\|\frac{x_n + x}{2}\right\| \le \limsup_{n\to\infty}\frac{\|x_n\| + \|x\|}{2} = \|x\| .$$

So we get the desired result. □

Best approximation. As in Theorem 2.6, the best approximation for a complete and convex subset of a uniformly convex space exists and, is unique.

Theorem 5.33 (Existence of the Unique Best Approximation)**.** *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*

$$\|x - k\| = dist(x, K) \coloneqq \inf_{k\in K}\|x - k\| . \tag{5.13}$$

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

$$\epsilon\left(\|x_n - x_m\|\right) \leq 1 - \left\|\frac{x_n + x_m}{2}\right\|\,.$$

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 $ck_n + (1-c)k_m \in K$.

$$\frac{x_n + x_m}{2} \geq \frac{d}{2}\left(\frac{1}{\|k_n\|} + \frac{1}{\|k_m\|}\right)$$

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

$$\epsilon\left(\left\|\frac{k - k'}{d}\right\|\right) \leq 1 - \left\|\frac{k + k'}{2d}\right\| \leq 0\,.$$

since $k + k'/2 \in K$, so $k = k'$. ☐

Remark 5.6. By Theorem 5.34, and Theorem 5.25, the existence of k is obvious, the uniform convexity guarantee the uniqueness.

Uniform convexity and reftexity. As everyone knows, every Hilbert space is reflexive. Surprisingly, every uniformly convex Banach space is, too.

Theorem 5.34 (Milman-Pettis)**.** *Every uniformly convex Banach space is reflexive.*

Proof. Suppose that X is a non-reflexive, uniformly convex Banach space. Then for some $\epsilon > 0$ there exists x^{**} in $B_{X^{**}}$ such that the distance between x^{**} and JB_X is 2ϵ. Let $\delta = \psi(\epsilon)$, so if x and y are in X with $\|x\| \leq 1, \|y\| \leq 1$ and $1 - \delta \leq \|\frac{x+y}{2}\|$, then $\|x - y\| \leq \epsilon$. Take x^* in X^*, with $\|x^*\| = 1$ such that $\langle x^*, x^{**}\rangle > 1 - \delta/2$. Let V be the weak-star neighborhood of x^{**} given by

$$V = V(x^{**}, x^*, \delta/2) = \{u^{**} \in X^{**} : |\langle x^*, u^{**} - x^{**}\rangle| < \delta/2\} \ .$$

If Jx and Jy are in JB_X belonging to V, then

$$\left\|\frac{x+y}{2}\right\| \geq \left|\left\langle \frac{x+y}{2}, x^*\right\rangle\right| > 1 - \delta$$

Hence $\|x - y\| \leq \epsilon$. Fixing $Jx \in V \cap JB_X$, we conclude that $V \cap JB_X \subset Jx + \epsilon B_{X^{**}}$. By Goldstein's theorem 5.19, we know that $V \cap JB_X$ is weak-star dense in $V \cap B_{X^{**}}$ which, since $Jx + \epsilon B\left(X^{**}\right)$ is weak-star closed, yields x^{**} belongs to $Jx + \epsilon B\left(X^{**}\right)$. But this means that the distance between x^{**} and $B(X)$ is less than or equal to ϵ, contradicting our choice of x^{**}. □

Remark 5.7. 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!

Chapter 6

Spectral Theory

6.1 The Resolvent and Spectrum

Let T be a linear operator whose domain $D(T)$ and range $R(T)$ both lie in the same complex linear topological space X. We consider the linear operator

$$T_\lambda = \lambda I - T$$

where λ is a complex number and I the identity operator. The distribution of the values of λ for which T_λ has an inverse and the properties of the inverse when it exists, are called the *spectral theory* for the operator T. We shall thus discuss the general theory of the inverse of T_λ.

Definition 6.1. If λ_0 is such that the range $R\left(T_{\lambda_0}\right)$ is dense in X and T_{λ_0} has a continuous inverse $\left(\lambda_0 I - T\right)^{-1}$, we say that λ_0 is in the *resolvent set* $\varrho(T)$ of T, and we denote this inverse $\left(\lambda_0 I - T\right)^{-1}$ by $R\left(\lambda_0; T\right)$ and call it the resolvent (at λ_0) of T.

Definition 6.2. All complex numbers λ not in $\varrho(T)$ form a set $\sigma(T)$ called the *spectrum* of T. The spectrum $\sigma(T)$ is decomposed into disjoint sets $\sigma_p(T), \sigma_c(T)$ and $\sigma_r(T)$ with the following properties:

- $\sigma_p(T)$ is the totality of complex numbers λ for which T_λ does not have an inverse $(N(\lambda I - T) \neq \{0\})$; $\sigma_p(T)$ is called the *point spectrum* of T.
- $\sigma_c(T)$ is the totality of complex numbers λ for which T_λ has a discontinuous inverse with domain dense in X $(\overline{R(\lambda I - T)} = X)$; $\sigma_c(T)$ is called the *continuous spectrum* of T.
- $\sigma_r(T)$ is the totality of complex numbers λ for which T_λ has an inverse whose domain is not dense in X $(\overline{R(\lambda I - T)} \neq X)$; $\sigma_r(T)$ is called the *residual spectrum* of T.

Remark 6.1. Why we need the scalar field be $\mathbb{C}$? The answer can be found in Theorem 6.2 and Theorem 6.7: The spectrum is not empty when X is a complex Banach space and $T \in \mathcal{B}(X)$, and $\lambda \to R(\lambda, T)$ is a operator-valued holomorphic function on the open subset $\varrho(T)$ of $\mathbb{C}$. However, if X is a real Banach sapce, all the definitions about spectrum can be done similarly.

From these definitions and the linearity of T we have : A necessary and sufficient condition for $\lambda_0 \in \sigma_p(T)$ is that the equation $Tx = \lambda_0 x$ has a solution $x \neq 0$. In this case λ_0 is called an *eigenvalue* of T, and x the corresponding *eigenvector*. The null space $N(\lambda_0 I - T)$ is called the *eigenspace* of T corresponding to the eigenvalue λ_0 of T. It consists of the vector 0 and the totality of eigenvectors corresponding to λ_0. The dimension of the eigenspace corresponding to λ_0 is called the *multiplicity* of the eigenvalue λ_0.

Theorem 6.1. *Let X be a complex Banach space, and T a closed linear operator with its domain $D(T)$ and range $R(T)$ both in X. Then, for any $\lambda_0 \in \varrho(T)$ the resolvent $(\lambda_0 I - T)^{-1}$ is an everywhere defined continuous linear operator: $(\lambda_0 I - T)^{-1} \in \mathcal{B}(X)$.*

Proof. Since λ_0 is in the resolvent set $\varrho(T), R(\lambda_0 I - T) = D\left((\lambda_0 I - T)^{-1}\right)$ is dense in X in such a way that there exists a positive constant c for which

$$\|(\lambda_0 I - T)\, x\| \geq c\|x\| \text{ whenever } x \in D(T)\,.$$

We have to show that $R\left(\lambda_0 I - T\right) = X$. But, if $\left(\lambda_0 I - T\right) x_n \to y$, then, by the above inequality, $\{x_n\}$ converges , and so, by the closure property of T, we must have $\left(\lambda_0 I - T\right) x = y$. Hence, by the assumption that $R\left(\lambda_0 I - T\right)$ is dense in X, we must have $R\left(\lambda_0 I - T\right) = X$. $\square$

Henceforth, without specific statement, we always suppose that X is a complex Banach space and T a closed linear operator from $D(T) \subset X$ into X. (And at most times, the domain of T is the whole space X.) In this case, $\varrho(T) = \{\lambda \in \mathbb{C} : (\lambda I - T)^{-1} \in \mathcal{B}(X)\}$. We note that the resolvent set $\varrho(T)$ consists of all those $\lambda \in \mathbb{C}$ for which the equation

$$\lambda x - Tx = y$$

has a unique solution for each y which furthermore depends continuously on the right hand side y.

Our first main result about the spectrum of closed linear operators is

Theorem 6.2. *Let T be a closed linear operator with domain and range both in a complex Banach space X. Then the resolvent set $\varrho(T)$ is an open set of the complex plane $\mathbb{C}$. In each component (the maximal connected sets) of $\varrho(T)$, $R(\lambda; T)$ is a holomorphic function of λ.*

Proof. By the theorem of the preceding section, $R(\lambda; T)$ for $\lambda \in \varrho(T)$ is an everywhere defined continuous operator. Let $\lambda_0 \in \varrho(T)$ and consider

$$S(\lambda) = R\left(\lambda_0; T\right) \left\{ I + \sum_{n=1}^{\infty} \left(\lambda_0 - \lambda\right)^n R\left(\lambda_0; T\right)^n \right\}$$

The series is convergent in the operator norm whenever $|\lambda_0 - \lambda| \, \|R\left(\lambda_0; T\right)\| < 1$, and within this circle of the complex plane, the series defines a holomorphic function of λ. Multiplication by $(\lambda I - T) = \left(\lambda - \lambda_0\right) I + \left(\lambda_0 I - T\right)$ on the left or right gives I so that the series $S(\lambda)$ actually represents the resolvent $R(\lambda; T)$. Thus we have proved that a circular neighbourhood of λ_0 belongs to $\varrho(T)$ and $R(\lambda; T)$ is holomorphic in this neighbourhood. $\square$

Theorem 6.3 (Resolvent Equation)**.** *Let T be a closed linear operator with domain and range both in a complex Banach space X. If λ and μ both belong to $\varrho(T)$, then the resolvent equation holds:*

$$\begin{aligned} R(\lambda;T) - R(\mu;T) &= (\mu-\lambda)R(\lambda;T)R(\mu;T) \\ &= (\mu-\lambda)R(\mu;T)R(\lambda;T)\,. \end{aligned}$$

particularly, $R(\lambda;T)$ and $R(\mu;T)$ commute.

Proof. We have

$$\begin{aligned} R(\lambda;T) &= R(\lambda;T)(\mu I - T)R(\mu;T) \\ &= R(\lambda;T)\{(\mu-\lambda)I + (\lambda I - T)\}R(\mu;T) \\ &= (\mu-\lambda)R(\lambda;T)R(\mu;T) + R(\mu;T)\,. \end{aligned}$$

Then the desired result follows. □

Proposition 6.4. *Let T be a closed linear operator with domain and range both in a complex Banach space X. Let $\lambda_1, \cdots, \lambda_n$ be distinct eigenvalues of T. If $x_1, \cdots, x_n$ are corresponding non-trivial eigenvectors, then $\{x_1, \cdots, x_n\}$ are linear independent. In particular, the (internal) direct sum*

$$N(\lambda_1 I - T) \oplus \cdots \oplus N(\lambda_n I - T)$$

is well-defined.

Proof. We prove by induction. When $n = 1$, the propostion is trivial. Assume the propostion holds for $n-1$. If $\alpha_j \in \mathbb{C}$, $j = 1, \cdots, n$ so that

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

Let T act on the equality, then we get

$$\alpha_1 \lambda_1 x_1 + \cdots + \alpha_n \lambda_n x_n = 0$$

Thus we get

$$\alpha_1(\lambda_n - \lambda_1)x_1 + \cdots + \alpha_{n-1}(\lambda_n - \lambda_{n-1})x_{n-1} = 0\,.$$

Thus $\alpha_1 = \cdots = \alpha_{n-1} = 0$ since $\{x_1, \cdots, x_{n-1}\}$ are linear independent by induction and $\lambda_1, \cdots, \lambda_n$ are distinct. Then calerly $\alpha_n = 0$ and the desired result follows. □

Example 6.1. If the space X is of finite dimension, then any bounded linear operator T is represented by a matrix (t_{ij}). It is known that the eigenvalues of T are obtained as the roots of the algebraic equation, the so-called characteristic equation of the matrix (t_{ij}),

$$\det(\lambda\delta_{ij} - t_{ij}) = 0$$

where $\det(A)$ denotes the determinant of the matrix A. Moreover, $\sigma(T) = \sigma_p(T)$ is the set of eigenvalues and $\#\sigma(T) \leq n$.

Example 6.2. Let $X = L^2(\mathbb{R}, \mathbb{C})$ and let T be defined by

$$(Tx)(t) = tx(t) \text{ for } t \in \mathbb{R};$$

where, $D(T) = \{x(t) : x(t) \text{ and } tx(t) \in L^2(\mathbb{R}, \mathbb{C})\}$. Then every real number λ_0 is in $\sigma_c(T)$.

To see this, note that the condition $(\lambda_0 I - T)\, x = 0$ implies $(\lambda_0 - t)\, x(t) = 0$ a.e. and so $x(t) = 0$ a.e.. Thus $(\lambda_0 I - T)^{-1}$ exists. The domain $D\left((\lambda_0 I - T)^{-1}\right)$ comprises those $y(t) \in L^2(\mathbb{R}, \mathbb{C})$ which vanish identically in the neighbourhood of $t = \lambda_0$; the neighbourhood may vary with $y(t)$. Hence $D\left((\lambda_0 I - T)^{-1}\right)$ is dense in $L^2(\mathbb{R}, \mathbb{C})$. It is easy to see that the operator $(\lambda_0 I - T)^{-1}$ is not bounded on the totality of such $y(t)$'s.

Example 6.3 (Spectrum of the Shift Operators)**.** Let X be the Hilbert space $\ell^2 = \ell^2(\mathbb{N}, \mathbb{C})$, and define the operators $A, B : \ell^2 \to \ell^2$ by

$$Ax := (x_2, x_3, x_4, \ldots), \quad Bx := (0, x_1, x_2, x_3, \ldots)$$

for $x = (x_i)_{i \in \mathbb{N}} \in \ell^2$. Then

$$\sigma(A) = \sigma(B) = D$$

is the closed unit disc in $\mathbb{C}$ and

$$\begin{aligned}
&\sigma_p(A) = \mathrm{int}(D), \quad &&\sigma_r(A) = \emptyset, &&\sigma_c(A) = S^1 \,. \\
&\sigma_p(B) = \emptyset, &&\sigma_r(B) = \mathrm{int}(D), \quad &&\sigma_c(B) = S^1 \,.
\end{aligned}$$

Example 6.4. Let $X = \ell^2\ell^2(\mathbb{N}, \mathbb{C})$ and let $(\lambda_i)_{i\in\mathbb{N}}$ be a bounded sequence of complex numbers. Define the bounded linear operator $A : \ell^2 \to \ell^2$ by

$$Ax := (\lambda_i x_i)_{i\in\mathbb{N}} \qquad \text{for } x = (x_i)_{i\in\mathbb{N}} \in \ell^2 \,.$$

Then

$$\sigma(A) = \overline{\{\lambda_i \mid i \in \mathbb{N}\}}, \quad \sigma_p(A) = \{\lambda_i \mid i \in \mathbb{N}\}\,, \quad \sigma_r(A) = \emptyset \,.$$

Thus *every nonempty compact subset of* $\mathbb{C}$ *is the spectrum of a bounded linear operator on an infinite-dimensional Hilbert space.*

Example 6.5. Let X be a complex Banach space and $P \in \mathcal{B}(X)$ is a projection, i.e., $P^2 = P$. Suppose P is not trivial, that is $P \neq 0$ and $P \neq I$. Then

$$\sigma(P) = \sigma_p(P) = \{0, 1\}\,.$$

Clearly $\{0, 1\} \subset \sigma_p(P)$. It suffices to show that for $\lambda \in \mathbb{C}\backslash\{0, 1\}$, $\lambda \in \varrho(P)$. Firstly, we claim that $N(\lambda I - P) = \{0\}$. If not, there exists $x \neq 0$ so that $Px = \lambda x \neq 0$. Hence $P^2 x = Px = \lambda Px$, and we deduce that $\lambda x = 1$, which is a contradiction. Secondly, we show that $R(\lambda I - P) = X$, then the desired result follows from the Banach inverse operator theorem. Take any $y \in X$, if $\lambda x - Px = y$, then $\lambda Px - Px = Py$. So $Px = (\lambda - 1)^{-1} Py$ and

$$x = \frac{1}{\lambda}(y - Px) = \frac{1}{\lambda}\left(y - \frac{1}{\lambda - 1} Py\right) \,.$$

Example 6.6. Let A be a self-adjoint operator in a Hilbert space H. Then the resolvent set $\varrho(A)$ of A comprises all the complex numbers λ with $\mathrm{Im}(\lambda) \neq 0$, and the resolvent $R(\lambda; A)$ is a bounded linear operator with the estimate

$$\|R(\lambda; A)\| \le \frac{1}{|\,\mathrm{Im}(\lambda)|} \,. \tag{6.1}$$

Moreover,

$$\mathrm{Im}\langle(\lambda I - A)x, x\rangle = \mathrm{Im}(\lambda)\|x\|^2, \qquad \text{for all } x \in H\,. \tag{6.2}$$

We now prove this. If $x \in H$ then $\langle Ax, x\rangle$ is real since $\langle Ax, x\rangle = \langle x, Ax\rangle = \overline{\langle Ax, x\rangle}$. Therefore we have (6.2) and so, by Schwarz' inequality,

$$\|(\lambda I - A)x\| \cdot \|x\| \geq |((\lambda I - A)x, x)| \geq |\,\mathrm{Im}(\lambda)| \cdot \|x\|^2$$

which implies that

$$\|(\lambda I - A)x\| \geq |\,\mathrm{Im}(\lambda)| \cdot \|x\|, x \in H\,.$$

Hence the inverse $(\lambda I - A)^{-1}$ exists and is continuous if $\mathrm{Im}(\lambda) \neq 0$. Moreover, the range $R(\lambda I - A)$ is dense in X if $\mathrm{Im}(\lambda) \neq 0$. If otherwise, there would exist a $y \neq 0$ orthogonal to $R(\lambda I - A)$, i.e., $\langle(\lambda I - A)x, y\rangle = 0$ for all $x \in H$ and so $(x, (\bar{\lambda} I - A)y) = 0$ for all $x \in H$. We must have $(\bar{\lambda} I - A)y = 0$, that is, $Ay = \bar{\lambda} y$, contrary to the reality of the value $\langle Ay, y\rangle$. Therefore, by Theorem 6.1, we see that, for any complex number λ with $\mathrm{Im}(\lambda) \neq 0$, the resolvent $R(\lambda; A)$ is a bounded linear operator with the estimate (6.1).

Definition 6.3. Let X be a complex Banach space, and T a closed linear operator with its domain $D(T)$ and range $R(T)$ both in X. $\lambda \in \mathbb{C}$ is called an *approximate eigenvalue* if there exists a sequence $x_n \in D(T)$ with $\|x_n\| = 1$ so that $\|Tx_n - \lambda x_n\| \to 0$. We call the set

$$\sigma_{ap}(T) := \{\lambda : \lambda \text{ approximate eigenvalue of } T\}$$

the *approximate point spectrum* of T.

We note that every eigenvalue λ is also an approximate eigenvalue (as we may simply choose $x_n = x$ for an element of $N(\lambda I - T)$ that is normalised to $\|x\| = 1$), so we have

$$\sigma_p(T) \subset \sigma_{ap}(T) \subset \sigma(T)\,.$$

We also remark that λ is an approximate eigenvalue if and only if there exists NO $\delta > 0$ so that

$$\|Tx - \lambda x\| \geq \delta\|x\| \qquad \text{for all } x \in X\,, \tag{6.3}$$

as is equivalent to

$$\inf_{x\in X,\|x\|=1} \|Tx - \lambda x\| = 0\,.$$

Moreover, we get directly that $\sigma_c(T) \subset \sigma_{ap}(T)$ from (6.3).

6.2 Spectrum of Bounded Linear Operators

In this section, we always set X a complex Banach space and T a cintinuous linear operator from X into X.

Theorem 6.5. *Let X be a complex Banach space and $T \in \mathcal{B}(X)$. Then the following limit exists:*

$$\lim_{n\to\infty} \|T^n\|^{\frac{1}{n}} = \inf_{n\geq 1} \|T^n\|^{\frac{1}{n}} =: r_\sigma(T)\,.$$

It is called the spectral radius of T. If $|\lambda| > r_\sigma(T)$ (for example $\lambda > \|T\|$), then the resolvent $R(\lambda; T)$ exists and is given by the series

$$R(\lambda; T) = \sum_{n=1}^{\infty} \frac{1}{\lambda^n} T^{n-1}\,, \tag{6.4}$$

which converges in the norm of operators.

Proof. Set $r = \inf_{n\geq 1} \|T^n\|^{\frac{1}{n}} \geq 0$. It suffices to show that $\limsup_n \|T^n\|^{\frac{1}{n}} \leq r$. For any $\epsilon > 0$, choose k such that $\left\|T^k\right\|^{\frac{1}{k}} \leq r+\epsilon$. For arbitrary n write $n = pk+q$ where $0 \leq q \leq (k-1)$. Then, by $\|AB\| \leq \|A\|\|B\|$, we obtain

$$\|T^n\|^{\frac{1}{n}} \leq \left\|T^k\right\|^{\frac{p}{n}} \cdot \|T\|^{\frac{q}{n}} \leq (r+\epsilon)^{\frac{kp}{n}} \|T\|^{\frac{q}{n}}\,.$$

Since $(kp)/n \to 1$ and $q/n \to 0$ as $n \to \infty$, we have $\limsup_n \|T^n\|^{\frac{1}{n}} \leq r+\epsilon$. Since e was arbitrary, $\limsup_n \|T^n\|^{\frac{1}{n}} \leq r$.

The series is convergent in the norm of operators when $|\lambda| > r_\sigma(T)$. For, if $|\lambda| \geq r_\sigma(T)+\epsilon$, then

$$\left\|\frac{1}{\lambda^{n+1}} T^n\right\| \leq \frac{(r_\sigma(T)+\epsilon/2)^n}{(r_\sigma(T)+\epsilon)^{n+1}}\,.$$

Thus the series in (6.4) converges. Multiplication by $(\lambda I - T)$ on the left or right of this series gives I so that the series actually represents the resolvent $R(\lambda;T)$. $\square$

Corollary 6.6. *Let X be a complex Banach space and $T \in \mathcal{B}(X)$. Then resolvent set $\varrho(T)$ is not empty.*

One of the most important aspects of the following theorem is that every bounded operator has non-empty spectrum. Here we crucially use that the vector space is over $\mathbb{C}$. The claim is not true if we were to only consider the real spectrum as you already know from Linear Algebra.

Theorem 6.7 (Spectral Radius)**.** *Let X be a complex Banach space and $T \in \mathcal{B}(X)$. Then the spectrum $\sigma(T)$ is non-empty and compact subset of $\mathbb{C}$. In fact,*

$$r_\sigma(T) = \sup_{\lambda\in\sigma(T)} |\lambda| ,$$

and this is why we call $r_\sigma(T)$ the spectral radius of T.

Proof. If $\sigma(T) = \emptyset$, then $\varrho(T) = \mathbb{C}$ and $\lambda \mapsto R(\lambda, T)$ is holomorphic on the plane $\mathbb{C}$. Since for $\lambda > \|T\|$, $R(\lambda, T) = \sum_{n=1}^{\infty} \lambda^{-n} T^{n-1}$, we have

$$\|R(\lambda, T)\| \le \frac{1}{\lambda - \|T\|} .$$

Thus $\lambda \mapsto R(\lambda, T)$ is a bounded holomorphic vector-valued function on $\mathbb{C}$. It follows from Theorem 3.50 that $R(\lambda, T)$ is a constant, which is a contradiction!

By Theorem 6.5, we know that $r_\sigma(T) \ge \sup_{\lambda\in\sigma(T)} |\lambda|$. Hence we have only to show that $r_\sigma(T) \le \sup_{\lambda\in\sigma(T)} |\lambda|$. By Theorem 6.2, $R(\lambda;T)$ is holomorphic in λ when $|\lambda| > \sup_{\lambda\in\sigma(T)} |\lambda|$. Thus it admits a uniquely determined Laurent expansion in positive and non-positive powers of λ convergent in the operator norm for $|\lambda| > \sup_{\lambda\in\sigma(T)} |\lambda|$. By Theorem 6.5, this Laurent series must coincide with $\sum_{n=1}^{\infty} \lambda^{-n} T^{n-1}$. Hence $\lim_{n\to\infty} \|\lambda^{-n} T^n\| = 0$ if $|\lambda| > \sup_{\lambda\in\sigma(T)} |\lambda|$, and so for any $\epsilon > 0$,

$$\|T^n\| \le \left(\epsilon + \sup_{\lambda\in\sigma(T)} |\lambda|\right)^n \qquad \text{for large } n .$$

This proves that

$$r_\sigma(T) = \lim_{n\to\infty} \|T^n\|^{\frac{1}{n}} \le \sup_{\lambda\in\sigma(T)} |\lambda| \,.$$

Since $\sigma(T)$ is a bounded closed subset of $\mathbb{C}$, it's compact, as required. $\square$

Corollary 6.8. *The series $\sum_{n=1}^{\infty} \lambda^{-n}T^{n-1}$ diverges if $|\lambda| < r_\sigma(T)$.*

Proof. Let r be the smallest non-negative real number such that the series $\sum_{n=1}^{\infty} \lambda^{-n}T^{n-1}$ converges in the operator norm for $|\lambda| > r$. The existence of such an r is proved as for ordinary power series in λ^{-1}. Then, for $|\lambda| > r$, $\lim_{n\to\infty} \|\lambda^{-n}T^n\| = 0$ and so, as in the proof of $r_\sigma(T) \le \sup_{\lambda\in\sigma(T)} |\lambda|$, we must have $\lim_{n\to\infty} \|T^n\|^{1/n} \le r$. This proves that $r_\sigma(T) \le r$. In fact, $r_\sigma(T) = r$. $\square$

We furthermore record the following useful lemma. This lemma has been shown in Exercise 0.1.

Lemma 6.9. *Let X be a vector space, Let $S, T \in \mathcal{L}(X)$. Suppose that $ST = TS$. Then ST is bijective if and only if S and T are bijective.*

Theorem 6.10. *Let X be a complex Banach space, $T \in \mathcal{B}(X)$ and let p be a complex polynomial. Then*

$$\sigma(p(T)) = p(\sigma(T)) := \{p(\lambda) : \lambda \in \sigma(T)\}$$

Here we set $p(T) := \sum_{j=0}^{n} a_j T^j$ if the polynomial p is given by $p(z) = \sum_{j=0}^{n} a_j z^j$, with the usual convention that $T^0 = I$.

Proof. We first remark that if p is constant, say $p = c \in \mathbb{C}$, then the spectrum of $p(T) = cI$ is simply $\{c\}$. while the fact that $\sigma(T)$ is non-empty implies that also $p(\sigma(T)) = \{c\}$. So suppose that p has degree $n \ge 1$, let $\mu \in \mathbb{C}$ be any given number. As we are working in $\mathbb{C}$ we can factorise $p(\cdot) - \mu$ and write it as $p(z) - \mu = \alpha\,(z - \beta_1(\mu)) \ldots (z - \beta_n(\mu))$ for some $\alpha \ne 0$ and equally factorise

$$p(T) - \mu I = \alpha\,(T - \beta_1(\mu)I) \ldots (T - \beta_n(\mu)I) \tag{6.5}$$

where we note that all operators on the right hand side commute which will allow us to apply Lemma 6.9. Thus, $\mu \in \varrho(p(T)) \Leftrightarrow \beta_j(\mu) \in \varrho(T)$ for all j. In other words,

$$\mu \in \sigma(p(T)) \Leftrightarrow \exists\; j \text{ so that } \beta_j(\mu) \in \sigma(T)\,.$$

We now note that $\mu \in p(\sigma(T))$ if and only if the equation $p(z) - \mu = 0$ has a root in $\sigma(T)$, in other words,

$$\mu \in p(\sigma(T)) \Leftrightarrow \;\; j \text{ so that } \beta_j(\mu) \in \sigma(T)\,.$$

Then the desired result follows. □

This theorem can in particular be applied if a given operator can be written as a polynomial of a simpler operator.

As a final result of this section, we prove that there is the following close connection between the spectrum of an operator and the spectrum of its dual operator.

Theorem 6.11. *Let X be a complex Banach space, let $T \in \mathcal{B}(X)$ and let $T^* \in \mathcal{B}(X^*)$ be the corresponding dual operator of T. Then*

(a) $\sigma(T) = \sigma(T^*)$.

(b) The point, residual, and continuous spectra of T and T^* are related by

$$\begin{aligned}
&\sigma_p(T^*) \subset \sigma_p(T) \cup \sigma_r(T), && \sigma_p(T) \subset \sigma_p(T^*) \cup \sigma_r(T^*)\,;\\
&\sigma_r(T^*) \subset \sigma_p(T) \cup \sigma_c(T), && \sigma_r(T) \subset \sigma_p(T^*)\,;\\
&\sigma_c(T^*) \subset \sigma_c(T), && \sigma_c(T) \subset \sigma_r(T^*) \cup \sigma_c(T^*)\,.
\end{aligned}$$

(c) If X is reflexive, then $\sigma_c(A^*) = \sigma_c(A)$ and

$$\begin{aligned}
&\sigma_p(A^*) \subset \sigma_p(A) \cup \sigma_r(A), && \sigma_p(A) \subset \sigma_p(A^*) \cup \sigma_r(A^*)\,;\\
&\sigma_r(A^*) \subset \sigma_p(A), && \sigma_r(A) \subset \sigma_p(A^*)\,.
\end{aligned}$$

(d) $\sigma(T) = \sigma_{ap}(T) \cup \sigma_p(T^*)$.

Proof. To prove part (a), notice that for any $\lambda \in \mathbb{C}$, $(\lambda I_X - T)^* = \lambda I_{X^*} - T^*$, and then (a) follows from Theorem 5.28.

To prove part (b), note tice that by Theorem 5.27, we have

$$\overline{R(\lambda I_X - T)} = {}^{\perp}N(\lambda I_{X^*} - T^*)\,,\ \overline{R(\lambda I_{X^*} - T^*)}^{\sigma(X^*,X)} = N(\lambda I_X - T)^{\perp}\,.$$

Assume first that $\lambda \in \sigma_p(T^*)$. Then $N(\lambda I_{X^*} - T^*) \neq \{0\}$, so $\lambda I_X - T$ does not have a dense image, and hence $\lambda \in \sigma_p(T) \cup \sigma_r(T)$. Next assume $\lambda \in \sigma_r(T^*)$. Then $N(\lambda I_{X^*} - T^*) = \{0\}$, hence $\lambda I_X - T$ has a dense image, and hence $\lambda \in \sigma_p(T) \cup \sigma_c(T)$. Third, assume $\lambda \in \sigma_c(T^*)$. Then $\lambda I_{X^*} - T^*$ is injective and has a dense image and therefore also has a weak-star dense image. Thus $\lambda I_X - T$ is injective and has a dense image, so $\lambda \in \sigma_c(T)$. It follows from these three inclusions that $\sigma_p(T)$ is disjoint from $\sigma_c(T^*)$, that $\sigma_c(T)$ is disjoint from $\sigma_p(T^*)$, and that $\sigma_r(T)$ is disjoint from $\sigma_r(T^*) \cup \sigma_c(T^*)$. This proves part (b).

To prove part (c) observe that in the reflexive case a linear subspace of X^* is weak-star dense if and only if it is dense (Use Hahn-Banach theorem to show this!). Hence $\sigma_c(T) = \sigma_c(T^*)$ whenever X is reflexive. With this understood, the remaining assertions of part (c) follow directly from part (b).

To prove part (d), observe that $\sigma(T) \backslash \sigma_{ap}(T) \subset \sigma_r(T)$. In fact, by (6.3), if $\lambda \in \sigma(T) \backslash \sigma_{ap}(T)$, then there exists $\delta > 0$ with

$$\|(\lambda I - T)x\| \geq \delta \|x\| \quad \text{for all } x \in X\,.$$

Then it follows from Exercise 4.2 that $\lambda \in \sigma_r(T)$, and the desired result follows from (b). □

Chapter 7

Compact Operators

7.1 Fundamentals

One of the most important concepts in the study of bounded linear operators is that of a compact operator. In this section, we always suppose that X and Y are two Banach space.

Definition 7.1. Let X and Y be Banach spaces. Let B_X be the closed unit ball in X. A linear map $K : X \to Y$ is said to be *compact* if $K(B_X)$ is relatively compact in Y; i.e., the closure of $K(B_X)$ is compact in Y.

We will denote by $\mathcal{C}(X,Y)$ all the compact linear operator from X into Y, and we write $\mathcal{C}(X)$ for $\mathcal{C}(X,X)$. Since Y is a complete metric space, the relatively compact subsets of Y are precisely the totally bounded ones. So every compact linear operator is bounded, i.e., $\mathcal{C}(X,Y) \subset \mathcal{B}(X,Y)$. Since the sequential compactness and compactness are equivalent in metric space, the notion of a compact operator can be defined in several equivalent ways, as in the following proposition:

Proposition 7.1. *Let X and Y be Banach spaces. Let $K : X \to Y$ be a bounded linear operator. Then the following are equivalent.*

(a) $K \in \mathcal{C}(X,Y)$*, i.e.,* $K(B_X)$ *is a relatively compact subset of* Y*.*

(b) *If* $S \subset X$ *is bounded, then* $K(S)$ *is relatively compact.*

(c) *If* (x_n) *is a bounded sequence in* X*, then the sequence* (Kx_n) *has a convergent subsequence in* Y*.*

This lemma follows trivially from Theorem A.2 and hence we omit the proof.

Many of the operators that arise in the study of integral equations are compact. This accounts for their importance from the standpoint of applications. They are in some respects as similar to linear operators on finite dimensional spaces as one has any right to expect from operators on infinite-dimensional spaces. As we shall see, these similarities show up particularly strongly in their spectral properties.

Example 7.1. Let Ω be a compact subset of $\mathbb{R}^n$. Let $k(x,y)$ be a $\mathbb{F}$-valued continuous function defined on $\Omega \times \Omega$. Denote by $C(\Omega)$ all the continuous $\mathbb{F}$-valued on Ω. Then $C(\Omega)$ endowed with the supremum norm is a Banach space. Then the integral operator $K : C(\Omega) \to C(\Omega)$ defined by

$$(Kf)(x) = \int_\Omega k(x,y)f(y)dy$$

is a compact linear operator. To see this, clearly K mapping $C(\Omega)$ into $C(\Omega)$ is linear. Denote by B the closed unit ball in $C(\Omega)$. In order that $K(B)$ is relatively compact, by the Ascoli-Arzelà theorem, it suffices to show that $K(B)$ is uniformly bounded and equi-continuous.

Since k is continuous, there exists $M > 0$ with $|k(x,y)| \leq M$ for all x, y in Ω. Then $\|Kf\| \leq (b-a)M\|f\| \leq (b-a)M$ for $f \in B$, so that $K(B)$ is uniformly bounded. Observe that for every $f \in B$, $\|f\| \leq 1$,

$$|(Kf)(x_1) - (Kf)(x_2)| \leq \int_\Omega |K(x_1,y) - K(x_2,y)|\, dy\,.$$

Since k is uniformly continuous in Ω, $K(B)$ is equi-continuous. Therefore, by the Ascoli-Arzelà theorem, the set $K(S)$ is relatively compact in $C(\Omega)$.

Example 7.2. If $K : X \to Y$ is a bounded linear operator between Banach spaces whose image is a closed infinite-dimensional subspace of Y, then K is not compact. Namely, the image of the closed unit ball in X under K contains an open ball in $R(K)$ by the open mapping theorem, and hence does not have a compact closure by Theorem 1.12.

Proposition 7.2. *Let X,Y and Z be Banach spaces. Then the following holds.*

(a) $\mathcal{C}(X,Y)$ is a closed subspace of $\mathcal{B}(X,Y)$.

(b) If $K \in \mathcal{C}(X,Y)$, then $R(K)$ is separable.

(c) Let $A \in \mathcal{B}(X,Y)$ and $B \in \mathcal{C}(X,Y)$ and one of them is compact, then $BA \in \mathcal{C}(X,Z)$. In particular, $\mathcal{C}(X)$ is two-side ideal in the algebra $\mathcal{B}(X)$.

(d) If $K \in \mathcal{C}(X,Y)$ and Z is a subspace of X. Then $K|_Z \in \mathcal{C}(Z,Y)$.

Proof. We prove part (a). Obviously, $\mathcal{C}(X,Y)$ is a linear subspace of $\mathcal{B}(X,Y)$. To show that $\mathcal{C}(X,Y)$ is closed, suppose $\{K_n\}$ is a convergent sequence in $\mathcal{C}(X,Y)$, and $K_n \to K$ in operator norm. Let $\{x_k\}$ be a bounded sequence in B_X. By the compact property of each K_n, we can choose, by the *diagonal method*, a subsequence $\{\hat{x}_k\}$ of $\{x_k\}$ such that $\{K_n\hat{x}_k\}$ converges for each fixed n. Then we shall show that $\{K\hat{x}_k\}$ is a Cauchy sequence, and hence is convergent. For fixed $n \geq 1$ and any $k,m \geq 1$, we have

$$\begin{aligned}\|K\hat{x}_m - K\hat{x}_k\| &\leq \|K\hat{x}_m - K_n\hat{x}_k\| + \|K_n\hat{x}_m - K_n\hat{x}_k\| + \|K_n\hat{x}_k - K\hat{x}_k\| \\ &\leq \|K - K_n\| + \|K_n\hat{x}_m - K_n\hat{x}_k\| + \|K_n - K\| \ .\end{aligned}$$

and so

$$\limsup_{m,k\to\infty} \|K\hat{x}_m - K\hat{x}_k\| \leq 2\,\|K - K_n\|$$

Since n is arbitrary, $\{K\hat{x}_k\}$ is a Cauchy sequence in the Y.

We prove part (b). Observe that

$$R(K) = K(X) = K\left(\bigcup_{n\geq 1} nB_X\right) = \bigcup_{n\geq 1} nK\left(B_X\right) \ .$$

Since $K(B_X)$ is totally bounded, $K(B_X)$ is separable, and hence $nK(B_X)$ is separable. Thus $R(K)$ is separable.

We prove part (c). Let (x_n) be a bounded sequence in X. If A is compact, then there exists a subsequence (x_{n_k}) such that the sequence (Ax_{n_k}) converges, and so does the subsequence (BAx_{n_k}). If B is compact, then, since the sequence (Ax_n) is bounded, there exists a subsequence (Ax_{n_k}) such that the sequence (BAx_{n_k}) converges. This proves (c).

We prove part (d). Since B_Z is a bounded subset in X and $K|_Z(B_Z) = K(B_Z)$, the desired result follows. □

Theorem 7.3 (Schauder)**.** *Let X and Y be Banach space. Then $K \in \mathcal{C}(X, Y)$ if and only if $K^* \in \mathcal{C}(Y^*, X^*)$.*

Proof. Necessity. Suppose that $K \in \mathcal{C}(X, Y)$. We show that $K^* \in \mathcal{C}(Y^*, X^*)$. Let $\{y_n^*\}$ be a sequence in B_{Y^*}. It suffices to show that $\{K^* y_n^*\}$ has a convergent sequence in X^*.

Recall that B_{Y^*} is sequentially compact relative to $\sigma(Y^*, Y)$ if Y is separable (see Theorem 5.15). If Y is not separable, let $Z = \overline{R(K)}$. Then Z is a separable Banach space. Let $\hat{K} : X \to Z$ defined by $\hat{K}x = Kx$ for all x. Then $\hat{K}^* : Z^* \to X^*$, satisfies that $\hat{K}^*(y^*|_Z) = K^* y^*$ for every $y^* \in Y^*$. Hence, without loss of generality we assume Y is separable and so B_{Y^*} is sequentially compact relative to $\sigma(Y^*, Y)$. Let $\{y_{n_k}^*\}$ be a weak-star convergent subsequence of $\{y_n^*\}$ and $y_{n_k}^* \to y^*$ in $\sigma(Y^*, Y)$. We shall show that $K^* y_{n_k}^* \to K^* y^*$ in norm.

$$\|K^* y_{n_k}^* - K^* y^*\| = \sup_{x \in B_X} |\langle Kx, y_{n_k}^* - y^* \rangle| = \sup_{y \in K(B_X)} |\langle y, y_{n_k}^* - y^* \rangle| .$$

Since $K(B_X)$ is totally bounded in Y, for any $\epsilon > 0$, there exists $\{y_1, \cdots, y_m\} \subset K(B_X)$ so that

$$K(B_X) \subset \bigcup_{j=1}^{m} B_Y(y_j, \epsilon) .$$

Since $y_{n_k}^* \to y^*$ in $\sigma(Y^*, Y)$,

$$\sup_{1 \le j \le m} |\langle y_j, y_{n_k}^* - y^* \rangle| \to 0 \quad \text{ as } k \to \infty .$$

Notice that

$$\|K^*y^*_{n_k} - K^*y^*\| = \sup_{y\in K(B_X)} |\langle y, y^*_{n_k} - y^*\rangle| \le \sup_{1\le j\le m} |\langle y_j, y^*_{n_k} - y^*\rangle| + 2\epsilon\,,$$

we get that

$$\limsup_{k\to\infty} \|K^*y^*_{n_k} - K^*y^*\| \le 2\epsilon\,.$$

Since $\epsilon > 0$ is arbitrary, $\|K^*y^*_{n_k} - K^*y^*\| \to 0$. The desired result follows.

Sufficiency. Conversely, suppose that K^* is compact. Then, by what we have just proved, the bidual operator $K^{**} \in \mathcal{C}(X^{**}, Y^{**})$. Since $K^{**}|_X = K$, this implies that K is compact. □

Exercise 7.1. For $y^*_n \in B_{Y^*}$ consider the continuous function

$$\varphi_n : \overline{K(B_X)} \to \mathbb{F}\,;\; y \to \langle y, y^*_n\rangle\,.$$

Use the Arzel'a-Ascoli theorem to prove Theorem 7.3.

Completely Continuous Operators. Let X and Y be Banach spaces. A bounded linear operator $K \in \mathcal{B}(X,Y)$ is said to be *completely continuous* if the image of every weakly convergent sequence in X under K converges in the norm topology on Y.

Theorem 7.4. *Let X and Y be Banach spaces. Then the following hold.*

(a) Every compact operator $K \in \mathcal{C}(X,Y)$ is completely continuous.

(b) If $K \in \mathcal{C}(X,Y)$ and X is reflexive. Then K is compact.

Proof. We prove part (a). Assume K is compact and let (x_n) be a sequence in X that converges weakly to $x \in X$. By the PUB, (x_n) is (norm) bounded so there exists a subsequence (x_{n_k}) of (x_n) so that (Kx_{n_k}) converges to $y \in Y$. On the other hand, $K : (X, \sigma(X,X^*)) \to (Y, \sigma(Y,Y^*))$ is continuous, so $Kx_{n_k} \xrightarrow{w} Kx$. Thus $y = Kx$ and $Kx_{n_k} \to Kx$. Note that this argument holds for every subsequence of (x_n), so every subsequence of (Kx_n) has a subsequence converging to Kx. Thus the desired result follows.

We prove part (ii). Assume X is reflexive and K is completely continuous. Let (x_n) be a bounded sequence in X. Since X is reflexive, there exists a weakly convergent subsequence (x_{n_k}) by Theorem 5.23. Let $x \in X$ be the limit of that subsequence. Since K is completely continuous, the sequence (Kx_{n_k}) converges strongly to Kx. Thus K is cpmpact. $\square$

Example 7.3. The hypothesis that X is reflexive cannot be removed in part (b) of Theorem 7.4. For example a sequence in ℓ^1 converges weakly if and only if it converges strongly (see Theorem 5.18). Hence the identity operator $\mathrm{id} : \ell^1 \to \ell^1$ is completely continuous. However, it is not a compact operator by Theorem 1.12.

Finite rank operator. Let X and Y be Banach spaces. A bounded linear operator $T : X \to Y$ is said to be of *finite rank* if its image $R(T)$ is a finite-dimensional subspace of Y. We will denote by $\mathcal{F}(X, Y)$ all the bound linear operator form X into Y with finite rank. Clearly, $\mathcal{F}(X, Y) \subset \mathcal{C}(X, Y)$, and we write $\mathcal{F}(X)$ for $\mathcal{F}(X, X)$.

Example 7.4. Let X and Y be Banach spaces. For each $x^* \in X^*$ and $y \in Y$, define $y \otimes x^*$ by

$$y \otimes x^* : X \to Y\,;\; x \mapsto \langle x, x^* \rangle y\,.$$

Clearly $y \otimes x^*$ is a finite rank operator. In fact $\dim R(y \otimes x^*) = 1$.

Lemma 7.5. *$T \in \mathcal{F}(X, Y)$ if and only if there exists $n \geq 1$, $y_j \in Y$ and $x_j^* \in X^*$ for $j = 1, \cdots, n$ so that*

$$T = \sum_{j=1}^{n} y_j \otimes x_j^*$$

Proof. Sufficiency is trivial, we only show the necessity. Let $n = \dim R(T)$ and let $\{y_j\}$ be a basis of $R(T)$. Then for any $x \in X$, there exists scalars $\{l_j(x)\}$ so that

$$Tx = \sum_{j=1}^{n} l_j(x) y_j\,.$$

Let x_j^* defined by $\langle x, x_j^* \rangle \mapsto l(x_j)$, then it's easy to check that $x_j^* \in X^*$ and so the desired result follows. $\square$

Clearly, the limit of a sequence of finite rank operators in the norm topology is a compact operator. It is a natural question to ask whether, conversely, every compact operator can be approximated in the norm topology by a sequence of finite rank operators; i.e.,

$$\overline{\mathcal{F}(X,Y)} = \mathcal{C}(X,Y)\,. \tag{7.1}$$

The answer to this question was an open problem in functional analysis for many years. It was eventually shown that the answer depends on the Banach space in question. For example, if $Y = H$ is a Hilbert space, then (7.1) holds.

To see this, note that for any given $K \in \mathcal{C}(X,H)$, $\overline{R(K)}$ is a separable Hilbert space. Let $\{e_n\}$ be a Hilbert basis of $\overline{R(K)}$. Let P_n be the orthogonal projection from H onto $\text{span}\{e_1, \cdots, e_n\}$. Define

$$K_n \coloneqq P_n \circ K \quad \text{for all } n \geq 1\,.$$

Clearly $K_n \to K$ in the strong operator topology. W show that

$$\|K_n \to K\| \to 0\,.$$

Given any $\epsilon > 0$, since $K(B_X)$ is totally bounded, there exists $x_1, \cdots, x_m$ in B_X so that for every $x \in B_X$, there exists x_j so that $\|x - x_j\| \leq \epsilon$, and

$$\begin{aligned}\|K_n x - Kx\| &\leq (\|K_n x - K_n x_j\| + \|K_n x_j - Kx_j\| + \|Kx_j - Kx\|) \\ &\leq 2\epsilon + \sup_{1\leq j\leq m} \|K_n x_j - Kx_j\|\,.\end{aligned}$$

Then the desired result follows.

Let Y be a Banach space. If for every Banach space X, (7.1) holds, we say that Y has the *approximation property*. Every Hilbert space has this property. The follows lemma due to Grothendieck implies that the whether (7.1) hols or not depends only on Y.

Lemma 7.6. *Let Y be a Banach space. Then Y has the approximation property if and only if for every compact subset $C \subset Y$ and every $\epsilon > 0$ there is $T \in \mathcal{F}(Y)$ such that $\|y - Ty\| \leq \epsilon$ for all $y \in C$.*

Remark 7.1. The above condition can be restated as: the identity operator $I : X \to X$ can be approximated, uniformly on every compact subset C of X, by linear operators of finite rank.

Proof. We prove the sufficiency. Given any Banach space X and any $K \in \mathcal{C}(X, Y)$. Since $\overline{K(B_X)}$ is compact, by assumption, for $\epsilon > 0$, there exists $T \in \mathcal{F}(Y)$ so that $\|Kx - TKx\| \leq \epsilon$ for all $x \in B_X$, and hence $\|K - TK\| \leq \epsilon$. Since $TK \in \mathcal{F}(X, Y)$ and ϵ is arbitrary, $K \in \overline{\mathcal{F}(X, Y)}$. The proof of necessity is not easy, see here. □

In fact, using Theorem 4.17 and the same argument we can show that:

Theorem 7.7. *Every Separable Banach space with a Schauder basis has the approximation property.*

Example 7.5. Fix a number $1 \leq p \leq \infty$ and a bounded sequence of scalars $\lambda = (\lambda_j)$. For $j \in \mathbb{N}$ let $e_j := (\delta_{ij}) \in \ell^p$. Clearly (e_j) is a Schauder basis in ℓ^p. Define the bounded linear operator $K_\lambda : \ell^p \to \ell^p$ by

$$K_\lambda x := (\lambda_j x_j) \qquad \text{for } x = (x_j) \in \ell^p$$

Then

$$K_\lambda \in \mathcal{C}(\ell^p) \Leftrightarrow \lim_{j \to \infty} \lambda_j = 0\,.$$

The condition $\lambda_j \to 0$ is necessary for compactness because, if there exist a constant $\delta > 0$ and a sequence $1 \leq n_1 < n_2 < n_3 < \cdots$ such that $|\lambda_{n_k}| \geq \delta$ for all $k \in \mathbb{N}$, then the sequence $Ke_{n_k} = \lambda_{n_k} e_{n_k}, k \in \mathbb{N}$, in ℓ^p has no convergent subsequence. The condition $\lambda_j \to 0$ implies compactness because then K can be approximated by a sequence of finite rank operators in the norm topology.

Example 7.6. If $(X, \mathcal{F}, \mu)$ is a measure space and $k \in L^2(X \times X, \mathcal{F} \times \mathcal{F}, \mu \times \mu)$ then

$$(Kf)(x) = \int k(x,y)f(y)\mu(dy) \tag{7.2}$$

is a compact operator and $\|K\| \leq \|k\|_2$.

The following lemma is useful for proving this proposition. The lemma is intuitive so the proof is omitted.

Lemma. *If $\{e_i\}_{i \in I}$ is a Hilbert basis for $L^2(X, \mathcal{F}, \mu)$ and*

$$\phi_{ij}(x,y) = e_j(x)\overline{e_i(y)} \quad \textit{for } i,j \in I \textit{ and } x,y \in X\,,$$

then $\{\phi_{ij}\}_{i,j \in I}$ is an orthonormal set in $L^2(X \times X, \mathcal{F} \times \mathcal{F}, \mu \times \mu)$. If k and K are as in (7.2), then $\langle k, \phi_{ij} \rangle = \langle Ke_j, e_i \rangle$.

It follows from Example 1.17 that K is a bounded linear operator on L^2 with $\|K\| \leq \|k\|_2$. Now let $\{e_i\}$ be a basis for $L^2(\mu)$ and define ϕ_{ij} as in the preceding lemma. Thus by Bessel inequality (see Theorem 2.14),

$$\|k\|^2 \geq \sum_{i,j} |\langle k, \phi_{ij} \rangle|^2 = \sum_{i,j} |\langle Ke_j, e_i \rangle|^2\,.$$

Since $k \in L^2(\mu \times \mu)$, there are at most a countable number of i and j such that $\langle k, \phi_{ij} \rangle \neq 0$; denote these by $\{\psi_{km} : 1 \leq k, m < \infty\}$. Note that $\langle Ke_j, e_i \rangle = 0$, unless $\phi_{ij} \in \{\psi_{km}\}$. Wiout loss of generality, let $\psi_{km}(x,y) = e_k(x)\overline{e_m(y)}$. Let P_n be the orthogonal projection onto span $\{e_k : 1 \leq k \leq n\}$, and put

$$K_n = KP_n + P_nK - P_nKP_n\,;$$

so K_n is a finite rank operator. We will show that $\|K - K_n\| \to 0$ as $n \to \infty$, thus showing that K is compact.

Let $f \in L^2(\mu)$ with $\|f\|^2 \le 1$; so $f = \sum_j \alpha_j e_j$. Hence

$$\begin{aligned}
\|Kf - K_n f\|^2 &= \sum_i |\langle Kf - K_n f, e_i\rangle|^2 \\
&= \sum_i \left| \sum_j \alpha_j \left\langle (K - K_n)\, e_j, e_i \right\rangle \right|^2 \\
&= \sum_k \left| \sum_m \alpha_m \left\langle (K - K_n)\, e_m, e_k \right\rangle \right|^2 \\
&\le \sum_k \left[\sum_m |\alpha_m|^2 \right] \left[\sum_m |\langle (K - K_n)\, e_m, e_k\rangle|^2 \right] \\
&= \|f\|^2 \sum_{k,m} |\langle (K - K_n)\, e_m, e_k\rangle|^2 \ ,
\end{aligned}$$

and

$$\begin{aligned}
&\sum_{k,m} |\langle (K - K_n)\, e_m, e_k\rangle|^2 \\
&= \sum_{k,m} \left| \, \langle K e_m, e_k\rangle - \langle K P_n e_m, P_n e_k\rangle - \langle K P_n e_m, P_n e_k\rangle + \langle K P_n e_m, P_n e_k\rangle \right|^2 \\
&= \sum_{k,m>n} |\langle K e_m, e_k\rangle|^2 = \sum_{k,m>n} |\langle k, \psi_{km}\rangle|^2 \ .
\end{aligned}$$

Since $\sum_{k,m} |\langle k, \psi_{km}\rangle|^2 < \infty, n$ can be chosen sufficiently large such that for any $\epsilon > 0$ this last sum will be smaller than ϵ^2. Thus $\|K - K_n\| \to 0$.

7.2 Riesz-Fredholm Theory

Let X be a Banach space over $\mathbb{F}$ and $T \in \mathcal{C}(X)$, i.e., T is a compact linear operator from X into X. Let λ be a scalar (we always set $\lambda \ne 0$). In this section, we shall discuss the following equation

$$(\lambda I - T)x = y \, .$$

We shall denote $T_\lambda \coloneqq \lambda I - T$.

Fredholm studied the following integral equations. Let $K(x,y) \in C([0,1] \times [0,1])$ be an integral kernel, and consider the inhomogeneous equation and homogeneous equation

$$x(t) - \int_0^1 K(t,s)x(s)ds = y(t)\,, \tag{7.3}$$

with it's dual equation

$$f(t) - \int_0^1 K(s,t)f(s)ds = g(t)\,, \tag{7.4}$$

where $x, y, f, g \in L^2[0,1] = L^2([0,1], \mathbb{R})$. He got the following result:

Result 1: About the equation (7.3)((7.4)), exactly one of the following holds:

(i) For every $y\,(g) \in L^2[0,1]$, the equation (7.3)((7.4)) has a unique solution.

(ii) For $y\,(g) = 0$ the equation (7.3)((7.4)) has a non-trivial solution.

Result 2: If (7.3) falls into the first case, then so is (7.4); If (7.3) falls into the second case, then so is (7.4). Besides, in the second case, the homogeneous equations of (7.3) and (7.4) have the same finite numbers of the linear independent solutions.

Result 3: The equation (7.3) has a solution if and only if

$$\int_0^1 f(t)y(t)dt = 0$$

for all f satisfying $f(t) - \int_0^1 K(s,t)f(s)ds = 0$. The equation (7.4) has a solution if and only if

$$\int_0^1 g(t)x(t)dt = 0$$

for all x satisfying $x(t) - \int_0^1 K(t,s)x(s)ds = 0$.

As we have pointed in Example 7.6, if we let T be the integral operator on $L^2[0,1]$ with kernel K, then T is a compact operator on $L^2[0,1]$, and the previous equation is exactly

$$(I - T)x = y$$

and

$$(I-T)^* f = g\,.$$

We will extend the above results to the general situation.

Lemma 7.8 (F.Riesz)**.** *Let X be a Banach space over $\mathbb{F}$ and $T \in \mathcal{C}(X)$. Then for any $\lambda \in \mathbb{F}\backslash\{0\}$, $R(T_\lambda) = R(\lambda I - T)$ is closed.*

Naturally, set $\lambda x_n - Tx_n \to y$, and we will check that if we can find x so that $\lambda x - Tx = y$. If we suppose that $\{x_n\}$ is (norm) bounded in X, then without loss of generality we assume that $\{Tx_n\}$ converges, and since $\lambda x_n = y - Tx_n$, $\lambda \neq 0$, so $\{x_n\}$ converges. Clearly, let $x_n \to x$, then $T_\lambda x_n \to T_\lambda x = y$ and the desired result holds.

Proof. Let $\widetilde{T_\lambda} : X/N(T_\lambda) \to R(T_\lambda)$ be given by

$$\widetilde{T_\lambda}\widetilde{x} \coloneqq T_\lambda x \quad \text{for every } \widetilde{x} \in X/N(T_\lambda)\,.$$

By Example 1.20, $\widetilde{T_\lambda}$ is a well-defined continuous linear operator. Moreover, $\widetilde{T_\lambda}$ is bijective. So we only need to show that $\widetilde{T_\lambda}^{-1}$ is continuous, then we can suppose that the sequence $\{x_n\}$ is bounded in norm. However, indeed, $\widetilde{T_\lambda}$ is a linear homomorphism between $X/N(T_\lambda)$ and $R(T_\lambda)$. Since $X/N(T_\lambda)$ is a Banach space, then $R(T_\lambda)$ muse be a closed subspace in Y.

If $\widetilde{T_\lambda}^{-1}$ is not continuous, by Exercise 4.2, there exists a sequence $\{\widetilde{x}_n\}$ in $X/N(T_\lambda)$ so that $\|\widetilde{x}_n\| = 1$ and $\|\widetilde{T}_\lambda x_n\| = \|\lambda x_n - Tx_n\| \to 0$. Without loss of generality, suppose $\{x_n\}$ is (norm) bounded and $Tx_n \to z$. Then $\lambda x_n \to z$, so $Tx_n \to \frac{1}{\lambda}Tz$, and hence $T_\lambda z = (\lambda z - Tz) = 0$. So $\widetilde{z} = \widetilde{0}$. But $x_n \to z$ implies $\widetilde{x}_n \to \widetilde{z} = \widetilde{0}$, which contradicts to $\|\widetilde{x}_n\| = 1$. We now complete the proof. $\square$

Theorem 7.9 (Fredholm Alternative)**.** *Let X be a Banach space over $\mathbb{F}$ and $T \in \mathcal{C}(X)$. Let $\lambda \in \mathbb{F}\backslash\{0\}$. Then $N(T_\lambda) = \{0\}$ if and only if $R(T_\lambda) = X$.*

Remark 7.2. The Fredholm alternative claim that either there is a non-zero solution $x \in X$ to the equation $Tx = \lambda x$ or the operator $T_\lambda = \lambda I - T$ has a continuous inverse $R(\lambda, T) = (\lambda I - T)^{-1}$ on X.

Proof. The strategy here is to try to contradict the compactness of T by exhibiting a bounded set whose image under T is not totally bounded.

Firstly, we assume that $N(T_\lambda) = \{0\}$. For each $n \in \mathbb{N}$, let

$$X_n = R(T_\lambda^n) = R((\lambda I - T)^n)\,,$$

and let $X_0 = X$. Obviously $X_{n+1} \subset X_n$ for $n \geq 0$. Suppose for contradiction that $X_1 = R(T_\lambda) \subsetneqq X$, then we have $X_{n+1} \subsetneqq X_n$ for every $n \geq 0$, since $N(T_\lambda) = \{0\}$ impleis $N(T_\lambda^n) = \{0\}$. Moreover, by Lemma 7.8, X_{n+1} is a closed of X_n, since $(\lambda I - T)^n = \lambda^n I + \sum_{j=1}^n \binom{n}{j}(-T)^j = \lambda^n I -$ (a compact operator).

$$X_0 \supsetneqq X_1 \supsetneqq X_2 \supsetneqq \cdots \supsetneqq X_n \supsetneqq X_{n+1} \supsetneqq \cdots$$

By Riesz's lemma (Lemma 1.11), there exists $x_n \in X_n$ with $\|x_n\| = 1$ and $d(x_n, X_{n+1}) > 1/2$.

Now let $n \geq 0$ and $p \geq 1$. By construction, x_{n+p}, $T_\lambda x_{n+p}$, $T_\lambda x_n$ all lie in X_{n+1}, and thus

$$\begin{aligned} Tx_n - Tx_{n+p} &= (\lambda I - T_\lambda)(x_n - x_{n+p}) \\ &= \lambda x_n + (-\lambda x_{n+p} - T_\lambda x_n + T_\lambda x_{n+p}) \in \lambda x_n + X_{n+1}\,. \end{aligned}$$

Since x_n lies at a distance at least $1/2$ from X_{n+1}, we conclude the separation proeprty

$$\|Tx_{n+p} - Tx_n\| \geq \frac{|\lambda|}{2}$$

But this implies that the sequence $\{Tx_n\}$ is not totally bounded, contradicting the compactness of T. Thus $N(T_\lambda) = \{0\}$ implies $R(T_\lambda) = X$.

Secondly, we assume that $R(T_\lambda) = X$. For each $n \in \mathbb{N}$, let

$$Y_n = N(T_\lambda^n) = N((\lambda I - T)^n)\,,$$

and let $Y_0 = \{0\}$. Obviously $Y_n \subset Y_{n+1}$ for $n \geq 0$. Suppose for contradiction that $\{0\} \subsetneqq N(T_\lambda) = Y_1$, then we have $Y_n \subsetneqq Y_{n+1}$ for every $n \geq 0$, since $R(T_\lambda) = X$ impleis $R(T_\lambda^n) = X$. Moreover, Y_{n+1} is a closed of Y_n.

$$Y_0 \subsetneqq Y_1 \subsetneqq Y_2 \subsetneqq \cdots \subsetneqq Y_n \subsetneqq Y_{n+1} \subsetneqq \cdots$$

By Riesz's lemma (Lemma 1.11), there exists $y_n \in Y_n$ with $\|y_n\| = 1$ and $d(y_n, Y_{n-1}) > 1/2$.

Now let $n \geq 0$ and $p \geq 1$. By construction, y_n, $T_\lambda y_{n+p}$, $T_\lambda y_n$ all lie in Y_{n+p-1}, and thus

$$\begin{aligned} Ty_{n+p} - Ty_n &= (\lambda I - T_\lambda)(y_{n+p} - y_n) \\ &= \lambda y_{n+p} + (-\lambda y_n - T_\lambda y_{n+p} + T_\lambda y_n) \in \lambda y_{n+p} + Y_{n+p-1}\,. \end{aligned}$$

Since y_{n+p} lies at a distance at least $1/2$ from Y_{n+p-1}, we conclude the separation proeprty

$$\|Ty_{n+p} - Ty_n\| \geq \frac{|\lambda|}{2}$$

But this implies that the sequence $\{Ty_n\}$ is not totally bounded, contradicting the compactness of T. Thus $R(T_\lambda) = X$ implies $N(T_\lambda) = \{0\}$. □

Remark 7.3. A hypothesis such as compactness is necessary; the shift operator U on the complex Hilbert space $\ell^2_{\mathbb{C}}$, for instance, has no eigenfunctions, but $\lambda I - U$ is not invertible for any unit complex number $\lambda I - U$ (see Example 6.3). The claim is also false when $\lambda = 0$; consider for instance the multiplication operator $(Tx)(n) := x(n)/n$ on $\ell^2_{\mathbb{C}}$, which is compact and has no eigenvalue at zero, but is not invertible (see Example 6.4).

Theorem 7.10 (Riesz-Fredholm). *Let X be a Banach space over $\mathbb{F}$ and $T \in \mathcal{C}(X)$. Denote by I the indentity operator on X. Then for every $\lambda \in \mathbb{F}\backslash\{0\}$,*

(a) $N(\lambda I - T) = \{0\}$ *if and only if* $R(\lambda I - T) = X$.

(b) $R(\lambda I - T) = {}^\perp N(\lambda I^* - T^*)$, *and* $R(\lambda I^* - T^*) = N(\lambda I - T)^\perp$.

(c) $\sigma(T) = \sigma(T^*)$ *and* $\dim N(\lambda I - T) = \dim N(\lambda I^* - T^*) < \infty$. *Moreover,* $\dim N(\lambda I - T) = \operatorname{codim} R(\lambda I - T)$.

Proof. As before, denote $T_\lambda = \lambda I - T$. Part (a) is the Riesz-Fredholm alternative. To show part (b), observing that $R(T_\lambda)$ is closed by Lemma 7.8, using Theorem 5.27, we get $R(T_\lambda) = {}^\perp N(T_\lambda^*)$. Besides, by Theorem 5.31, $R(T_\lambda)$ is closed implies that $R(T_\lambda^*)$ is weak-star closed in X^*. Then using Theorem 5.27 again, we get $R(T_\lambda^*) = N(T_\lambda^*)^\perp$.

We now prove part (c). By Theorem 6.11, $\sigma(T) = \sigma(T^*)$. Since $T|_{N(T_\lambda)} = \lambda I|_{N(T_\lambda)}$, the indentity operator from $N(T_\lambda)$ onto $N(T_\lambda)$ is compact. Thus the closed unit ball in $N(T_\lambda)$ is compact. By Theorem 1.12, there must holds that $\dim N(T_\lambda) < \infty$. Since T^* is also compact ($T^* \in \mathcal{C}(X^*)$) by Theorem 7.3, we get $\dim N(T_\lambda^*) < \infty$. We only need to show that $\dim N(T_\lambda) = \dim N(T_\lambda^*)$. By Theorem 5.27, $N(T_\lambda^*) = R(T_\lambda)^\perp$; and by Theorem 5.5, $R(T_\lambda)^\perp$ is isometrically isomorphic to $(X/R(T_\lambda))^*$. So it suffices to show that $\dim N(T_\lambda) = \dim(X/R(T_\lambda))^*$. Since $\dim N(T_\lambda) < \infty$, the desired result holds if and only if

$$\dim N(T_\lambda) = \dim X/R(T_\lambda) \equiv \operatorname{codim} R(T_\lambda) = \dim N(T_\lambda^*) .$$

To this end, we firstly show the following lemma:

Lemma 7.11. $\operatorname{codim} R(T_\lambda) \equiv \dim(X/R(T_\lambda)) \le \dim N(T_\lambda) < \infty.$

Proof of Lemma 7.11. Suppose for contradiction that

$$\operatorname{codim} R(T_\lambda) > \dim N(T_\lambda) = n .$$

Then we can find $\{\widetilde{x}_1, \cdots, \widetilde{x}_{n+1}\}$ in $X/R(T_\lambda)$ which is linear independent. Trivially, $\{x_1, \cdots, x_{n+1}\}$ is linear independent. Let $N = \operatorname{span}\{x_1, \cdots, x_n\}$. Clearly $N \cap R(T_\lambda) = \{0\}$ and $x_{n+1} \notin R(T_\lambda) \oplus N$. Let $V : N(T_\lambda) \to N$ be a linear isomorphism. Let $P : X \to N(P)$ be a projection on X, that is, $P \in \mathcal{B}(X)$, $P^2 = P$ and $R(P) = N(T_\lambda)$. (The existence of such a projection is contained in Theorem 2.6 and Theorem 4.16.) Now define

$$S := T_\lambda + VP = \lambda I - (T - VP) .$$

Firstly, $N(S) = \{0\}$ since $N \cap R(T_\lambda) = \{0\}$. Secondly, observe that $T - VP \in \mathcal{C}(X)$, then we get $R(S) = X$ by the Fredholm alternative. However, $x_{n+1} \notin R(S)$ since $R(S) \subset R(T_\lambda) \oplus N$, which is a contradiction! $\square$

Using Lemma 7.11, since $\dim X/R(T_\lambda) < \infty$, we have

$$\begin{aligned} \dim N(T_\lambda^*) &= \dim R(T_\lambda)^\perp = \dim(X/R(T_\lambda))^* \\ &= \dim X/R(T_\lambda) \le \dim N(T_\lambda) . \end{aligned}$$

For the same reason, $\dim N(T_\lambda^{**}) \le \dim N(T_\lambda^*)$. But $\dim N(T_\lambda) \le \dim N(T_\lambda^{**})$ since we can regard T_λ^{**} as an extension of T on X^{**} (see Remark 5.5), then the desired result follows. $\square$

7.3 Riesz-Schauder Theory

In this section, we will deal with the following two subjects: the spectrum of compact operators; the construction of compact operators. Corresponding to matrices, each question has a clear answer: 1, matrices have eigenvalues whose number $\le$ the dimension of the space. 2, using a series of invariant subspaces, a matrix can be transformed into a Jordan canonical form.

7.3.1 Spectrum of Compact Operators

The spectral theory of compact operators is considerably simpler than that of general bounded linear operators. In particular, every nonzero spectral value is an eigenvalue, the generalized eigenspaces are all finite-dimensional, and zero is the only possible accumulation point of the spectrum (i.e. each nonzero spectral value is an isolated point of the spectrum).

Theorem 7.12 (Riesz-Schauder)**.** *Let X be a Banach space over $\mathbb{F}$ and $T \in \mathcal{C}(X)$, i.e., T is a compact operator on X. Then the following holds*

(a) Every non-zero spectral point of T is an eigenvalue, in other words,

$$\sigma(T)\backslash\{0\} = \sigma_p(T)\,.$$

(b) If the dimension of X is not finite, then $\sigma(T)$ must contain 0.

(c) $\sigma(T)$ is at most countably infinite. If $\sigma(T)$ is countably infinite, let $\{\lambda_n\}$ be all the non-zero eigencalues, then $\lim_n \lambda_n = 0$.

Proof. Part (a) is the Fredholm alternative. Part (b) is the easy: If not, then $T : X \to X$ is a linear homeomorphism. Denote by B_X the closed unit ball of X. Then $T^{-1}(B_X)$ is a bounded subset of X. As a consequence of the compactness of T, $T(T^{-1}(B_X)) = B_X$ is compact , deducing that X has finite dimension. This is a contradiction.

We now show part (c). It sufficies to show that for each $\delta > 0$, there are only finite many eigenvalues outside $B(0, \delta)$. Suppose for contradiction that there is a sequence $\{\lambda_n\}$ su that $|\lambda_n| > \delta$ and $\lambda_n \neq \lambda_m$ for every $n \neq m$. Let $\{x_n\}$ be the corresponding non-trivial eigenvectors of $\{\lambda_n\}$. For each $n \in \mathbb{N}$, let

$$X_n = \operatorname{span}\{x_1, \cdots, x_n\}\,,$$

and $X_0 = \{0\}$. Then clearly $X_n \subsetneqq X_{n+1}$ for $n \geq 0$. By Riesz-Fredholm theory, X_n has finite dimension. So it follows from Lemma 1.11 that there exists $y_n \in X_n$ with

$$d(y_n, X_{n-1}) \geq \frac{1}{2} \quad \text{and} \quad \|y_n\| = 1\,.$$

Assume that $y_n = \sum_{j=1}^{n} \alpha_j^n x_j$, then $Ty_n = \sum_{j=1}^{n} \alpha_j^n \lambda_j x_j$, and hence

$$Ty_n - \lambda_n y_n \in X_{n-1}\,.$$

For $n > m \geq 0$, since $Ty_m \in X_m \subset X_{n-1}$, we have

$$Ty_n - Ty_m = \lambda_n y_n + [(Ty_n - \lambda_n y_n) - Ty_m] \in \lambda_n y_n + X_{n-1}\,.$$

Then

$$\|Ty_n - Ty_m\| \geq d(\lambda_n y_n, X_{n-1}) = |\lambda_n| d(y_n, X_{n-1}) \geq \frac{\delta}{2}\,.$$

But this implies that the sequence $\{Tx_n\}$ is not totally bounded, contradicting the compactness of T. We now complete the proof. $\square$

7.3.2 Construction of Compact Operators

The classical result for square matrices is the Jordan canonical form, which states the following:

Theorem. *Let A be an $n \times n$ complex matrix, i.e. A a linear operator acting on $\mathbb{C}^n$. If $\lambda_1, \cdots, \lambda_k$ are the distinct eigenvalues of A, then $\mathbb{C}^n$ can be decomposed into the invariant subspaces of A*

$$\mathbb{C}^n = N(\lambda_1 I - A)^{r_1} \oplus \cdots \oplus N(\lambda_k I - A)^{r_k}$$

where r_j is the algebraic multiplicity of eigenvalue λ_j satisfying $N\left(\lambda_i - A\right)^m = N\left(\lambda_i - A\right)^{m+1}$.

Theorem. *Let V be a finite dimensional vector sapce. Let L be a nilpotent transformation on V and q is the minimal positive integer satisfying $L^q = 0$. Then there exists positive integer r, integers $1 \leq q_1 \leq \cdots \leq q_r \leq q$, and vectors $x_1, \cdots, x_r$ such that*

$$\begin{array}{cccc}
\{x_1, & Lx_1, & \cdots, & L^{q_1-1}x_1 \\
x_2, & Lx_2, & \ldots, & L^{q_2-1}x_2 \\
\vdots & \vdots & & \vdots \\
x_r, & Lx_r, & \cdots, & L^{q_r-1}x_r\}
\end{array}$$

is a basis of V with $L^{q_1}x_1 = \cdots = L^{q_r}x_r = 0$. Under this basis, the matrix of L is $\operatorname{diag}\{J_1, \cdots, J_r\}$, where

$$J_k = \begin{pmatrix} 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ & & & 0 \end{pmatrix}_{q_k \times q_k}.$$

We will try to extend the result to compact linear operators on Banach space X. Firstly, we point out the following fact. Let $A \in \mathcal{B}(X)$ be a bounded linear transformation on X. As we all know, (set $A^0 = I$)

$$\begin{aligned}
\{0\} = N(A^0) \subset N(A) \subset N(A^2) \subset \cdots \subset N(A^n) \subset \cdots, \\
X = R(A^0) \supset R(A) \supset R(A^2) \supset \cdots \supset R(A^n) \supset \cdots.
\end{aligned}$$

If $N(A^r) = N(A^{r+1})$ for some integer $r \geq 1$, then $N(A^r) = N(A^{r+k})$ for all $k \in \mathbb{N}$. Similarly, if $R(A^r) = R(A^{r+1})$ for some integer $m \geq 1$, then $R(A^r) = R(A^{r+k})$ for all $k \in \mathbb{N}$. However, for general $A \in B(X)$, we can not guarantee the existence of such a r, unless $A = T_\lambda = \lambda I - T$ for some $T \in \mathcal{C}(X)$ and $\lambda \neq 0$, as the following lemma states.

Lemma 7.13. *Let X be a Banach space over $\mathbb{F}$. Let $T \in \mathcal{C}(X)$ be a compact operator. Let $\lambda \in \mathbb{F}\backslash\{0\}$. Then*

$$\begin{aligned} r &\coloneqq \inf\{m \geq 0 : N(\lambda I - T)^m = N(\lambda I - T)^{m+1}\} \\ &= \inf\{m \geq 0 : R(\lambda I - T)^m = R(\lambda I - T)^{m+1}\} < \infty\,. \end{aligned}$$

Proof. Let $p \coloneqq \inf\{m \geq 0 : N(\lambda I - T)^m = N(\lambda I - T)^{m+1}\}$ and $q = coloneqq \inf\{m \geq 0 : R(\lambda I - T)^m = R(\lambda I - T)^{m+1}\}$. It follows from the proof of Theorem 7.9 that p, q are finite. We shall show that $p \leq q$ and $q \leq p$, respectively.

To show that $p \leq q$, notice that $R(\lambda I - T)^q = R(\lambda I - T)^{q+1}$. By Theorem 7.10 (c), since $(\lambda I - T)^q = \lambda^q I + \sum_{j=1}^n \binom{q}{j}(-T)^j = \lambda^q I-$ (a compact operator), we have

$$\begin{aligned} \dim N(\lambda I - T)^q &= \operatorname{codim} R(\lambda I - T)^q \\ &= \operatorname{codim} R(\lambda I - T)^{q+1} = \dim N(\lambda I - T)^{q+1}\,. \end{aligned}$$

Because $\dim N(\lambda I - T)^q < \infty$, we have $N(\lambda I - T)^q = N(\lambda I - T)^{q+1}$. Thus it follows from the definition that $p \leq q$. Similarly, notice that $N(\lambda I - T)^p = N(\lambda I - T)^{p+1}$. By Theorem 7.10 (c) we have

$$\begin{aligned} \operatorname{codim} R(\lambda I - T)^p &= \dim N(\lambda I - T)^p \\ &= \operatorname{codim} N(\lambda I - T)^{p+1} = \operatorname{codim} R(\lambda I - T)^{p+1}\,. \end{aligned}$$

Because $\operatorname{codim} R(\lambda I - T)^q < \infty$, we have $R(\lambda I - T)^q = R(\lambda I - T)^{q+1}$. Thus it follows from the definition that $q \leq p$. Hence $p = q < \infty$. ☐

The subspace $E_\lambda \coloneqq E_\lambda(T) \coloneqq N(\lambda I - T)^r$ is called the *generalized eigenspace* of T associated to the non-zero eigenvalue λ. Recall that $\dim N(\lambda I - T)$ is called the *geometric multiplicity* of the eigenvalue λ and $\dim N(\lambda I - T)^r$ is called the *algebraic multiplicity* of the eigenvalue λ.

Theorem 7.14. *Let X be a Banach space and let $T \in \mathcal{C}(X)$ be a compact operator. Let λ be an non-zero eigenvalue of T. Then*

$$X = N(\lambda I - A)^r \oplus R(\lambda I - A)^r$$

where r is defined in Lemma 7.13.

Proof. Firstly, we show that $N(\lambda I - A)^r \cap R(\lambda I - A)^r = \{0\}$. In fact, if $y \in N(\lambda I - A)^r \cap R(\lambda I - A)^r$, then $(\lambda I - A)^r y = 0$ and there exists $x \in X$ so that $y = (\lambda I - A)^r x$. So $(\lambda I - A)^{2r} x = 0$. Since $N(\lambda I - A)^{2r} = N(\lambda I - A)^r$, we have $y = N(\lambda I - A)^r x = 0$.

Secondly, we show that $N(\lambda I - A)^r \oplus R(\lambda I - A)^r = X$. To this end, given any $x \in X$, we need to find some $z \in X$ with $x - (\lambda I - A)^r z \in N(\lambda I - A)^r$. That is

$$(\lambda I - A)^r x = (\lambda I - A)^{2r} z\,.$$

Since $R(\lambda I - A)^r = R(\lambda I - A)^{2r}$, such a z exists. We now complete the proof. $\square$

Remark 7.4. Since $(\lambda I - T)$ and T communicates, $N(\lambda I - T)^r, N(\lambda I - T)^r$ are invariant subspaces of T. Let $\widehat{T}_\lambda$ be the restriction of $\lambda I - T$ on $R(\lambda I - T)^r$ with it's range in $R(\lambda I - T)^r$. Then $\widehat{T}_\lambda$ is a linear homeomorphism.

Since $N(\lambda I - T)^r$, $R(\lambda I - T)^r$ both are closed subspace of X, let P_λ be the corresponding projection from X onto $N(\lambda I - T)^r$. Then we have a decomposition as following:

$$T = TP_\lambda + T(I - P_\lambda)\,.$$

Denote $S = TP_\lambda$ and $R = T(I - P_\lambda)$.

Corollary 7.15. *Let X be a Banach space and let $T \in \mathcal{C}(X)$ be a compact operator. Let λ be a non-zero eigenvalue of T. Then there exists a finite rank operator S on X and a compact operator R on X, satisfying the following assertions.*

(a) $T = S + R$ *and* $SR = RS = 0$.

(b) λ *is an eigenvalue of* S. *For* $\mu \neq 0$ *and* $\mu \neq \lambda$, $\mu \in \varrho(S)$.

(c) $\lambda \in \varrho(R)$. *For* $\mu \neq 0$ *and* $\mu \neq \lambda$, μ *is an eigenvalue of* R *if and only if* μ *is an eigenvalue of* T.

Proof. To show part (a), it suffices to show that T communicates with P_λ. This is trivial since $R(P_\lambda) = N(\lambda I - T)^r$ is an invariant subspace of T.

To show part (b), take any $\mu \neq 0$. If μ is an egivalue of S, then there exists $x \neq 0$ so that $(\mu I - S)x = 0$. Let $y = P_\lambda x$ and $z = (I - P_\lambda x)$, then we get

$$(\mu I - S)(y + z) = (\mu I - T)y + \mu z = 0\,.$$

Thus $z = 0$ and $Ty = \mu y$. Then $(\lambda I - T)^r y = (\lambda - \mu)^r y = 0$ forces that $\mu = \lambda$. Thus λ is the only non-zero eigenvalue of S.

We now show part (c). Take any $\mu \neq 0$. If μ is an egivalue of R, there exists $x \neq 0$ so that $(\mu I - R)x = 0$. Let $y = P_\lambda x$ and $z = (I - P_\lambda x)$, then we get

$$(\mu I - R)(y + z) = \mu y + (\mu I - T)z = 0\,.$$

Thus $y = 0, z \neq 0$ and $Tz = \mu z$. So we must have $\mu \in \sigma_p(T)$ and $\mu \neq \lambda$ (since if $\mu = \lambda$ then $z = 0$). Thus $\lambda \in \varrho(R)$, and for $\mu \neq 0$ and $\mu \neq \lambda$, if μ is an eigenvalue of R then μ is an eigenvalue of T.

On the other hand, if $\mu \neq 0$, $\mu \neq \lambda$ is an eigenvalue of T. There exists $x \neq 0$ so that $(\mu I - T)x = 0$. Let $y = P_\lambda x$ and $z = (I - P_\lambda x)$, then we get

$$(\mu I - T)(y + z) = (\mu - T)y + (\mu I - T)z = 0\,.$$

Then $Ty = \mu y$ and $Tz = Rz = \mu z$. Since $(\lambda I - T)^r y = (\lambda - \mu)^r y = 0$, we must have $y = 0$, so $z \neq 0$. Thus μ is an eigenvalue of R. $\square$

Chapter 8

Operators on Hilbert Space

A large area of current research interest is centered around the theory of operators on Hilbert space. There is a marked contrast here between Hilbert spaces and the Banach spaces. 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.

8.1 The Adjoint of an Operator

8.1.1 Sesquilinear form

Let H and K be Hilbert spaces over $\mathbb{F}$. 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 $|u(h, k)| \leq M\|h\|\|k\|$ for all h in H and k in K. When u is bounded, we say

$$\|u\| := \sup_{\|h\| \leq 1, \|k\| \leq 1} |u(h, k)| \tag{8.1}$$

is the *norm* of u, and u is bounded if and only if u is continous. In fact, if "only if" part is trivial; if u is continuous but not bounded, for any n, there exists $h_n \in H$ and $k_n \in K$ such that

$$\|h_n\| = \|k_n\| = 1\,, |u(h_n, k_n)| \geq n\,.$$

Then h_n/n and k_n/n tend to 0, but $|u(h_n/n, k_n/n)| \geq 1$ is a contradiction.

Example 8.1. Sesquilinear forms are used to study operators.

(a) If $A \in \mathcal{B}(H, K)$, then $u(h, k) := \langle Ah, k \rangle$ is a bounded sesquilinear form, and $\|u\| = \|A\|$.

(b) Also, if $B \in \mathcal{B}(K, H)$ $u(h, k) := \langle h, Bk \rangle$ 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 8.1. *Let H, K be Hilbert spaces. Let $u : H \times K \to \mathbb{F}$ be a continuous sesquilinear form. Then there are unique operators $A \in \mathcal{B}(H, K)$ and $B \in \mathcal{B}(K, H)$ such that*

$$u(h, k) = \langle Ah, k \rangle = \langle h, Bk \rangle \tag{8.2}$$

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,

$$|L_h(k)| \le \|u\|\|h\|\|k\|.$$

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 $\|f_h\| \le \|u\|\|h\|$.

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 **??**exa: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. ☐

A variant of F.Riesz's representation theorem, formulated by P. Lax and A.N.Milgram, is a useful tool for the discussion of the existence of solutions of linear partial differential equations of elliptic type.

Theorem 8.2 (The Lax-Milgram Theorem)**.** *Let H be Hilbert spaces. Let $u : H \times H \to \mathbb{F}$ be a continuous sesquilinear form. If u is also coercive, that is, there is some $c > 0$ so that*

$$u(x,x) \ge c\|x\|^2 \quad \textit{for each} \quad x \in H\,.$$

Then there exists a uniquely determined $S \in \mathcal{B}(H)$ such that $\langle x, y\rangle = u(x, Sy)$ whenever $x, y \in H$, and $\|S\| \le c^{-1}$.

Proof. By Theorem 8.1, there exists $B \in \mathcal{B}(H)$ so that $u(x,y) = \langle x, By\rangle$ whenever $x, y \in H$. If we can show that B is bijective, then letting $S = B^{-1}$, clearly we have $\langle x, y\rangle = u(x, Sy)$ whenever $x, y \in H$. Moreover, since for every $y \in H$,

$$\|Sy\|\|y\| \ge |\langle Sy, y\rangle| = |u(Sy, Sy)| \ge c\|Sy\|^2\,,$$

we get that $\|S\| \le c^{-1}$. We now have to show that B is bijective. Oviously, B is injective since if $By = 0$ then $0 = |\langle y, By\rangle| = |u(y,y)| \ge c\|y\|^2$, so $y = 0$. Next, we show that $R(B)^{\perp} = \{0\}$, from which we deduce that $R(B)$ is dense in H. Note that if $x \in R(B)^{\perp}$, then $0 = |\langle x, Bx\rangle| = |u(x,x)| \ge c\|x\|^2$, and hence $x = 0$. Finally, we show that $R(B)$ is closed, which follows trivially from that

$S = B^{-1} : R(B) \to H$ is a continuous lienar operator and hence $R(B)$ and H are lienar homogeneous. $\square$

Example 8.2. Here is a sketch of the typical application of Lax-Milgram to elliptic PDEs (we don' t need to know all notions in advance, just grab the main message). The task is always to identify the "good" Hilbert space of functions among which we look for solutions to a given PDE, and to check the validity of the assumptions on u. The sesquilinear form emerges naturally when testing the PDE against some test functions (in fact, one looks for *weak solutions*). Coercivity encodes some kind of Sobolev embedding.

Let Ω be a bounded region in $\mathbb{R}^d$. Let $f \in L^2(\Omega)$ be given. Let $a \geq 0$. Consider the boundary-value problem

$$\begin{cases} -\Delta u + au = f & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega \end{cases}. \tag{8.3}$$

The claim is: there exists a unique"weak" solutions u to (8.3) in the space $H_0^1(\Omega)$, the completion of $C_0^\infty(\Omega)$ with respect to the norm

$$\|v\|_{H_0^1} := \sqrt{\|v\|_2^2 + \sum_{j=1}^{n} \|\partial_j v\|_2^2}\,.$$

Note that the condition $u \in H_0^1(\Omega)$ encodes the vanishing of $u \in H_0^1(\Omega)$ at the boundary of Ω. By weak solution to (8.3) one means a function u that satisfies

$$\langle \nabla v, \nabla u \rangle + a\langle v, u \rangle = \langle v, f \rangle \quad \forall v \in H_0^1(\Omega) \tag{8.4}$$

where $\langle \cdot, \cdot \rangle$ denotes the scalar product in L^2 as usual. (∇u for $u \in H_0^1(\Omega)$ is a well defined function in L^2 via a limiting procedure- recall the definition of $H_0^1(\Omega)$). In fact, if u and v were smooth and vanished on $\partial\Omega$ then, owing to Green's identity,

$$\begin{aligned} \int_\Omega -(\Delta u)(x) v(x) dx &= \int_\Omega f(x) v(x) dx - \int_{\partial\Omega} \frac{\partial u}{\partial n} v d\sigma \\ &= \int_\Omega f(x) v(x) dx\,. \end{aligned}$$

Thus, in fact, (8.4) is the PDE (8.3) "tested" against v. Thus, (8.4) would be certainly satisfied by a "classical" solution u to (8.3) which might not exist in this case, though, because f is a priori not smooth enough. (8.4) suggests that the appropriate bilinear form in this case is

$$B(v,u) := \langle \nabla v, \nabla u\rangle + a\langle v,u\rangle\,.$$

Such a B is bounded on $H_0^1(\Omega)$ because $\forall u,v \in H_0^1(\Omega)$

$$|B(v,u)| \le \|\nabla v\|_2\|\nabla u\|_2 + a\|v\|_2\|u\|_2 \le (1+a)\|v\|_{H_0^1}\|u\|_{H_0^1}$$

and is coercive on $H_0^1(\Omega)$ because $\forall v \in H_0^1(\Omega)$

$$|B(v,v)| = \|\nabla v\|_2^2 + a\|v\|_2^2 \ge \|\nabla v\|_2^2 \ge c\|v\|_2$$

where the last step is the "deep" one and follows from the Poincaré's inequality for functions $g \in H_0^1(\Omega)$

$$\|g\|_{L^2(\Omega)} \le c_\Omega \|\nabla g\|_{L^2(\Omega)} \quad (\text{ when } |\Omega| < \infty)$$

and $c := \min\left\{\frac{1}{2}, \frac{1}{2c_\Omega^2}\right\} > 0$. Therefore Lax-Milgram says that there exists uniquely a $u \in H_0^1(\Omega)$ such that $B(v,u) = \langle v, f\rangle$ which means precisely that there is a unique weak solution u to (8.3).

8.1.2 Adjoints

Definition 8.1. If $A \in \mathcal{B}(H,K)$, then the unique operator B in $\mathcal{B}(K,H)$ satisfying (8.2) is called the *adjoint* of A, and is denoted by $B = A^*$.

Remark 8.1. 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 8.3. P is a projection operater on H, then $P^* = P$.

Example 8.4. 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 8.1: 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 8.3. *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 8.4. *If $A \in \mathcal{B}(H)$, then*

$$(\operatorname{ran} A)^\perp = \ker A^* \text{ and } \overline{\operatorname{ran} A} = (\ker A^*)^\perp \ . \tag{8.5}$$

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 8.5. *$A \in \mathcal{B}(H)$, then $H = \ker A \oplus \overline{\operatorname{ran} A}$.*

Examples.

Example 8.5. 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 8.6. 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 8.7. 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 8.8. 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.

8.2 Three Oerators on Hilbert space

Definition 8.2. 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 8.9.

- 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.

8.2.1 Self-adjoint operator

Definition 8.3. 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 8.1 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 8.6 (Polar indenity)**.** *u is a sesquiliear form on $\mathbb{C}$-Hilbert space H. Then the following "polar indentity" holds for any $h_1, h_2 \in H$.*

$$\begin{aligned} u(h_1,h_2) = & \frac{u(h_1+h_2,h_1+h_2) - u(h_1-h_2,h_1-h_2)}{4} \\ & + i \times \frac{u(h_1+ih_2,h_1+ih_2) - u(h_1-ih_2,h_1-ih_2)}{4} . \end{aligned} \tag{8.6}$$

Remark 8.2. If u is a sesquiliear hermitian form on $\mathbb{F}$-Hilbert space H, the "polar indentity" holds without doubt.

Proposition 8.7. *u is a sesquilinear form on $\mathbb{C}$-Hilbert space H. Then u is hermitian if and only if*

$$u(h,h) \in \mathbb{R}\,, \quad \text{for all } h \in H\,. \tag{8.7}$$

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. $\square$

Example 8.10 (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 8.8. *$A \in \mathcal{B}(H)$ and A is self-adjoint. Then*

$$\|A\| = \sup_{\|h\|=1} |\langle Ah, h\rangle|\,. \tag{8.8}$$

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

$$M := \sup_{\|h\|=1} |v(h)| \geq \|A\| = \|u\|\,.$$

For any $h_1, h_2 \in H$ such that $\|h_1\|, \|h_2\| \leq 1$. 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

$$|\operatorname{Re} u(h_1, h_2)| \leq \frac{M}{4}(\|h_1 + h_2\|^2 + \|h_1 - h_2\|^2) = \frac{M}{2}(\|h_1\|^2 + \|h_2\|^2) \leq M\,.$$

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,

$$|u(h_1, h_2)| = u(e^{-i\theta}h_1, h_2) = \operatorname{Re} u(e^{-i\theta}h_1, h_2) \leq M\,.$$

Therefore, $\|A\| = \|u\| \leq M$. $\square$

Corollary 8.9. *$A \in \mathcal{B}(H)$ is self-adjoint.If $\langle Ah, h\rangle = 0$ for all h, then $A = 0$.*

8.2.2 Normal operator

Theorem 8.10. *$A \in \mathcal{B}(H)$, then A is normal if and only if*

$$\|Ah\| = \|A^*h\| \ \textit{for all} h \in H. \tag{8.9}$$

If H is a $\mathbb{C}$-Hilbert space and $A \in \mathcal{B}(H)$, then let

$$B = \frac{A + A^*}{2}\,, \text{ and } C = \frac{A - A^*}{2i}\,. \tag{8.10}$$

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 8.11. *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.*

8.2.3 Unitary

Proposition 8.12. *$A \in \mathcal{B}(H)$,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 8.13. *$A \in \mathcal{B}(H)$,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.*

8.3 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 $M \leq H$ 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 8.4. 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 8.11. 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 8.14.

(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 $M, N \leq H, M \cap N = \{0\}$, 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 8.15. *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) *$\langle Eh, h\rangle \geq 0$ for all h in H.*

Proof. $3 \Rightarrow 1$. For any $x \in \operatorname{ran} E$ and $y \in \ker E$, we have $\|E(x+y)\| = \|x\| \leq \|x+y\|$, thus $Re\langle x, y\rangle \geq -\frac{1}{2}\|y\|^2$.Take any $y' \in H$ and $t > 0$, then $Re\langle x, y'\rangle \geq -\frac{t}{2}\|y'\|^2$. Let $t \to 0$, we have $Re\langle x, y'\rangle \geq 0$. 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 8.4, $(\operatorname{ran} E)^\perp = \ker E^*$, so E is a projection.

$6 \Rightarrow 1$. For any $x \in \operatorname{ran} E$ and $y \in \ker E$, we have $\langle E(x+y), x+y\rangle = \langle x, x+y\rangle \geq 0$, then $\langle x, y\rangle \geq -\|x\|^2$. Let $x = tx'$ where $t > 0$ and $x' \in H$. ☐

投影算子的运算.

Proposition 8.16. *Let $\{M_i : i \in I\}$ be a collection of closed subspaces of H. Then* $\cap\left\{M_i^\perp : i \in I\right\} = (\operatorname{cspan}\{M_i : i \in I\})^\perp$ *and* $(\cap\{M_i : i \in I\})^\perp = \operatorname{cspan}\left\{M_i^\perp : i \in I\right\}$.

Proposition 8.17. *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 8.18. *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*

$$\bigoplus_{i\in I} M_i := \left\{ \sum_{i\in I} x_i : x_i \in M_i,\ \sum_{i\in I} \|x_i\|^2 < \infty \right\}$$

then $\operatorname{ran} P = \bigoplus_{i\in I} M_i$*, or $P = P_{\oplus_{i\in I} M_i}$.*

Proposition 8.19. *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 8.5. 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 8.20. *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 8.1. $\{P_n\}$ are projections and is monotonous, i.e. $P_n \leq P_{n+1}$ for all n or $P_n \geq P_{n+1}$ for all n, then there exists a projection P such that $P_n \to P$ (s).

Invariant and Reducing Subspaces.

Definition 8.6. Let $A \in \mathcal{B}(H)$ and $M \leq H$, 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 8.21. *If $A \in \mathcal{B}(H), M \leq H$, then M is invariant for A if and only if $P_M A P_M = A P_M$.*

Proof. □

Proposition 8.22. *If $A \in \mathcal{B}(H), M \leq H$, 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 8.3. If $M \leq H$, 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.

Chapter 9

Operator Semigroups

Introduction

Generally speaking, a dynamical system is a family $(T(t))_{t\geq 0}$ of mappings on a set X satisfying

$$\begin{cases} T(t+s) = T(t)T(s) \text{ for all } t, s \geq 0 \\ T(0) = \mathrm{id} \end{cases}$$

Here X is viewed as the set of all states of a system, $t \in \mathbb{R}_+ := [0, \infty)$ as time and $T(t)$ as the map describing the change of a state $x \in X$ at time 0 into the state $T(t)x$ at time t. In the linear context, the state space X is a vector space, each $T(t)$ is a linear operator on X, and $(T(t))_{t\geq 0}$ is called a (one-parameter) *operator semigroup*.

The standard situation in which such operator semigroups naturally appear are so-called *Abstract Cauchy Problems* (ACP).

$$\begin{cases} \dot{u}(t) = Au(t) \quad \text{for } t \geq 0 \\ u(0) = x \end{cases}$$

where A is a linear operator on a Banach space X. Here, the problem consists in finding a differentiable function $u : \mathbb{R}_+ \to X$ such that (ACP) holds. If for each initial value $x \in X$ a unique solution $u(\cdot, x)$ exists, then

$$T(t)x := u(t,x), \quad t \geq 0, x \in X$$

defines an operator semigroup. For the "working mathematician," (ACP) is the problem, and $(T(t))_{t\geq 0}$ the solution to be found. The opposite point of view also makes sense: given an operator semigroup (i.e., a dynamical system $(T(t))_{t\geq 0}$, under what conditions can it be "described" by a differential equation (ACP), and how can the operator A be found?

In some simple and concrete situations the relation between $(T(t))_{t\geq 0}$ and A is given by the formulas

$$T(t) = \mathrm{e}^{tA} \quad \text{and} \quad A = \frac{d}{dt}T(t)\bigg|_{t=0}.$$

In general, a comparably simple relation seems to be out of reach. However, miraculously as it may seem, a simple continuity assumption on the semigroup produces, in the usual Banach space setting, a rich and beautiful theory with a broad and almost universal field of applications. It is the aim of this chapter to develop this theory.

9.1 Strongly Continuous Semigroups

The following is our basic definition.

Definition 9.1. A family $(T(t))_{t\geq 0}$ of bounded linear operators on a Banach space X is called a *strongly continuous (one-parameter) semigroup* if it satisfies the functional equation

$$\begin{cases} T(t+s) = T(t)T(s) & \text{for all } t, s \geq 0 \\ T(0) = I \end{cases} \tag{FE}$$

and is strongly continuous in the following sense. For every $x \in X$ the orbit maps

$$\xi_x : t \mapsto \xi_x(t) := T(t)x \tag{SC}$$

are continuous from $\mathbb{R}_+$ into X for every $x \in X$.

The property (SC) can also be expressed by saying that the map

$$t \mapsto T(t)$$

is continuous from $\mathbb{R}_+$ into the space $(\mathcal{B}(X), \mathcal{T}_s)$ of all bounded operators on X endowed with the strong operator topology. Finally, if these properties hold for $\mathbb{R}$ instead of $\mathbb{R}_+$, we call $(T(t))_{t\in\mathbb{R}}$ a *strongly continuous (one-parameter) group* on X.

9.1.1 Basic Properties

Our first goal is to facilitate the verification of the strong continuity (SC) required in Definition 9.1. This is possible thanks to the uniform boundedness principle, which implies the following frequently used equivalence.

Lemma 9.1. *Let X be a Banach space over $\mathbb{F}$. Let F be a mapping from a compact set $I \subset \mathbb{R}$ into $\mathcal{B}(X)$. Then the following assertions are equivalent.*

(a) F is continuous for the strong operator topology; i.e., the mappings $t \mapsto F(t)x\,;\, I \to X$ are continuous for every $x \in X$.

(b) $\{F(t) : t \in I\}$ is uniformly bounded, and the mappings $t \mapsto F(t)x\,;\, I \to X$ are continuous for all x in some dense subset D of X.

(c) For every compact subset K of X, the map $(t, x) \mapsto F(t)x\,;\, I \times K \to X$ is uniformly continuous .

Proof. (a) $\Rightarrow$ (b). Since for any fixed x, $t \to F(t)x$ is a continuous on the compact set I, $\{F(t)x : t \in I\}$ is norm bounded in X. It follows from PUB that $\{F(t) : t \in I\}$ is uniformly bounded.

(b) $\Rightarrow$ (c). We shall show that for any given $\epsilon > 0$, there exists some $\delta = \delta_\epsilon > 0$, so that for any $(t_1, x_1), (t_2, x_2)$ in $I \times K$ provided that $|t_1 - t_2| + \|x_1 - x_2\| \le \delta$, we have

$$\|F(t_1)x_1 - F(t_2)x_2\| \le \epsilon\,.$$

Since K is compact, we can find a $M \in \mathbb{R}_+$ satisfying $\sup_{x\in K} \|x\| \le M$. Since $\{F(t) : t \in I\}$ is uniformly bounded, without loss of generality we suppose that

$\sup_{t\in I}\|F(t)\| \le M$ too. Then, for $|t_1 - t_2| + \|x_1 - x_2\| \le \delta$, taking $x_0 \in D \cap B(x_1, \delta) \cap B(x_2, \delta)$ and using triangular inequality we get

$$
\begin{aligned}
&\|F(t_1)x_1 - F(t_2)x_2\| \\
&\le \|F(t_1)x_1 - F(t_1)x_0\| + \|F(t_1)x_0 - F(t_2)x_0\| + \|F(t_2)x_0 - F(t_2)x_2\| \\
&\le M\delta + M\|F(t_1) - F(t_2)\| + M\delta\,.
\end{aligned}
$$

Since the mapping $t \to F(t)x_0$ is continuous on the compact set I, it must be uniformly continuous. So there sufficiently small $\delta = \delta_\epsilon < \epsilon/(3M)$ such that $M\|F(t_1) - F(t_2)\| \le \epsilon/3$. Then the desired result follows.

(b) $\Rightarrow$ (c). This is trivial. □

As an easy consequence of this lemma, in combination with the functional equation (FE), we obtain that the continuity of the orbit maps

$$\xi_x : t \mapsto T(t)x$$

at each $t \ge 0$ and for each $x \in X$ is already implied by much weaker properties.

Proposition 9.2. *For a semigroup $(T(t))_{t\ge 0}$ on a Banach space X, the following assertions are equivalent.*

(a) $(T(t))_{t\ge0}$ *is strongly continuous.*

(b) *There exist $\delta > 0, M \ge 1$, and a dense subset $D \subset X$ such that*

  (bi) $\|T(t)\| \le M$ *for all $t \in [0, \delta]$,*

  (bii) $\lim_{t\downarrow 0} T(t)x = x$ *for all $x \in D$.*

(c) $\lim_{t\downarrow 0} T(t)x = x$ *for all $x \in X$.*

Proof. (a) $\Rightarrow$ (b). Claerly (a) $\Rightarrow$ (bi). Take any $\delta > 0$, then $t \mapsto T(t);\ [0, \delta] \to \mathcal{B}(X)$ is continuous for strong operator topology. By Lemma 9.1, (a) implies (bii).

(b) $\Rightarrow$ (c). This follows from Lemma 9.1 trivially. Indeed, take $I = [0, \delta]$ and $K = \{x\}$ in Lemma 9.1, then the desired result follws.

(c) $\Rightarrow$ (a). Let $t_0 > 0$ and let $x \in X$. We will show that $t \to T(t)x$ is continuous at t_0. To show the right continuity, observe that

$$\lim_{h \downarrow 0} \|T(t_0+h)x - T(t_0)x\| \leq \|T(t_0)\| \cdot \lim_{h \downarrow 0} \|T(h)x - x\| = 0 .$$

For $h < 0$, the estimate

$$\|T(t_0+h)x - T(t_0)x\| \leq \|T(t_0+h)\| \cdot \|x - T(-h)x\|$$

implies left continuity whenever $\|T(t)\|$ remains uniformly bounded for $t \in [0, t_0]$. However, combine the pointwise boundedness implied by the continuity of $t \to T(t)x$ at 0 and the PUB, there must exists a $\delta > 0$ so that $\{T(t) : t \in [0, \delta]\}$ is uniformly bounded. □

Because in many cases the uniform boundedness of the operators $T(t)$ for $t \in [0, t_0]$ is obvious, one obtains strong continuity by checking (right) continuity of the orbit maps ξ_x at $t = 0$ for a dense set of "nice" elements $x \in X$ only. We demonstrate the advantage of this procedure in the examples discussed below (e.g., in Paragraph ??).

We repeat that for a strongly continuous semigroup $(T(t))_{t \geq 0}$ the finite orbits

$$\{T(t)x : t \in [0, t_0]\}$$

are continuous images of a compact interval, hence compact and therefore bounded for each $x \in X$. So by the uniform boundedness principle each strongly continuous semigroup is uniformly bounded on each compact interval, a fact that implies exponential boundedness on $\mathbb{R}_+$.

Proposition 9.3. *For every strongly continuous semigroup $(T(t))_{t \geq 0}$, there exist constants $w \in \mathbb{R}$ and $M \geq 1$ such that*

$$\|T(t)\| \leq M \mathrm{e}^{wt} \quad \text{for all } t \geq 0 . \tag{9.1}$$

Proof. Choose $M \geq 1$ such that $\|T(s)\| \leq M$ for all $0 \leq s \leq 1$ and write $t \geq 0$ as $t = s + n$ for $n \in \mathbb{N}$ and $0 \leq s < 1$. Then

$$\begin{aligned} \|T(t)\| &\leq \|T(s)\| \cdot \|T(1)\|^n \leq M^{n+1} \\ &= M \mathrm{e}^{n \log M} \leq M \mathrm{e}^{wt} \end{aligned}$$

holds for $w := \log M$ and each $t \geq 0$. $\square$

The infimum of all exponents ω for which an estimate of the form (9.1) holds for a given strongly continuous semigroup plays an important role in the sequel. We therefore reserve a name for it.

Definition 9.2. For a strongly continuous semigroup $\mathfrak{T} = (T(t))_{t\geq 0}$, we call

$$\omega_0 := \omega_0(\mathfrak{T}) := \inf\left\{ w \in \mathbb{R} : \begin{array}{c} \text{there exists } M_w \geq 1 \text{ such that} \\ \|T(t)\| \leq M_w \mathrm{e}^{wt} \text{ for all } t \geq 0 \end{array} \right\}$$

its *growth bound* (or *type*). Moreover, a semigroup is called *bounded* if we can take $w = 0$ in (9.1) and *contractive* if $w = 0$ and $M = 1$ is possible. Finally, the semigroup $(T(t))_{t\geq 0}$ is called *isometric* if $\|T(t)x\| = \|x\|$ for all $t \geq 0$ and $x \in X$.

It becomes clear in the discussion below, but is presently left as a challenge to the reader that

- $\omega_0 = -\infty$ may occur,
- the infimum in (9.1) may not be attained; i.e, it might happen that no constant M exists such that $\|T(t)\| \leq M\mathrm{e}^{\omega_0 t}$ for all $t \geq 0$, and
- Constants $M > 1$ may be necessary; i.e., no matter how large $w \geq \omega_0$ is chosen, $\|T(t)\|$ will not be dominated by e^{wt} for all $t \geq 0$.

We close this subsection by showing that using the weak operator topology instead of the strong operator topology in Definition 9.1 will not change our class of semigroups. This is a surprising result, and its proof needs more sophisticated tools from functional analysis than we have used up to this point. So the beginner may just skip the proof.

Theorem 9.4. *A semigroup $(T(t))_{t\geq 0}$ on a Banach space X is strongly continuous if and only if it is weakly continuous, i.e., if the mappings*

$$t \mapsto \langle T(t)x, x^* \rangle \ ; \ [0, \infty) \to \mathbb{F}$$

are continuous for each $x \in X, x^* \in X^*$.

Proof. We have only to show that weak implies strong continuity. As a first step, for any $\delta > 0$, $(T(t) : t \in [0, \delta])$ is bounded for the weak operator topology. By PUB, it's uniformly bounded. Using Proposition 9.2 (b), it suffices to show that

$$E := \left\{ x \in X : \lim_{t \downarrow 0} \|T(t)x - x\| = 0 \right\}$$

is a (strongly) dense subspace of X. To this end, we define for $x \in X$ and $r > 0$ a linear form x_r on X^* by

$$\langle x_r, x^* \rangle := \frac{1}{r} \int_0^r \langle T(s)x, x^* \rangle \, ds \qquad \text{for } x^* \in X^*$$

Then x_r is bounded and hence $x_r \in X^{**}$. On the other hand, the set

$$\{T(s)x : s \in [0, r]\}$$

is the continuous image of $[0, r]$ in the space X endowed with the weak topology, hence is weakly compact in X. Krein's theorem implies that its closed convex hull

$$\overline{\mathrm{co}}\{T(s)x : s \in [0, r]\}$$

is still weakly compact in X. Because x_r is a $\sigma(X^{**}, X^*)$ -limit of Riemann sums, it follows that

$$x_r \in \overline{\mathrm{co}}\{T(s)x : s \in [0, r]\}$$

whence $x_r \in X$. (See also [Rudin, Thm. 3.27].) It is clear from the definition that the set

$$D := \{x_r : r > 0, x \in X\}$$

is weakly dense in X. On the other hand, for $x_r \in D$ we obtain

$$\begin{aligned}
\|T(t)x_r - x_r\| &= \sup_{\|x^*\| \le 1} \left| \frac{1}{r} \int_t^{t+r} \langle T(s)x, x^* \rangle \, ds - \frac{1}{r} \int_0^r \langle T(s)x, x^* \rangle \, ds \right| \\
&\le \sup_{\|x^*\| \le 1} \left(\left| \frac{1}{r} \int_r^{r+t} \langle T(s)x, x^* \rangle \, ds \right| + \left| \frac{1}{r} \int_0^t \langle T(s)x, x^* \rangle \, ds \right| \right) \\
&\le \frac{2t}{r} \|x\| \sup_{0 \le s \le r+t} \|T(s)\| \to 0
\end{aligned}$$

as $t \downarrow 0$; i.e., $D \subset E$. We conclude that E is weakly dense in X. However, as a subspace of X, E is convex. So E is strongly dense in X. $\square$

9.1.2 Standard Constructions

We now explain how one can construct in various ways new strongly continuous semigroups from a given one. This might seem trivial and/or boring, but there will be many occasions to appreciate the toolbox provided by these constructions. In any case, it is a series of instructive exercises.

In the following, we always assume $\mathfrak{T} = (T(t))_{t\geq 0}$ to be a strongly continuous semigroup on a Banach space X.

Example 9.1 (Similar Semigroups)**.** Given another Banach space Y and a linear homeomorphism V from Y onto X, we obtain (as in Lemma??) a new strongly continuous semigroup $\mathfrak{S} = (S(t))_{t\geq 0}$ on Y by defining

$$S(t) := V^{-1}T(t)V \quad \text{for } t \geq 0\,.$$

Without explicit reference to the linear homeomorphism V, we call the two semigroups $\mathfrak{T}$ and $\mathfrak{S}$ *similar*. Two such semigroups have the same topological properties; e.g., $\omega_0(\mathfrak{T}) = \omega_0(\mathfrak{S})$.

Example 9.2 (Rescaled Semigroups)**.** For any numbers $\beta \in \mathbb{C}$ and $\alpha > 0$, we define the rescaled semigroup $(S(t))_{t\geq 0}$ by

$$S(t) := \mathrm{e}^{\beta t}T(\alpha t)$$

for $t \geq 0$. For example, taking $\beta = -\omega_0$ (or $\beta < -\omega_0$) and $\alpha = 1$ the rescaled semigroup will have growth bound equal to (or less than) zero. This is an assumption we make without loss of generality in many situations.

Example 9.3 (Subspace Semigroups)**.** If Y is a closed subspace of X such that $T(t)Y \subseteq Y$ for all $t \geq 0$, i.e., if Y is $(T(t))_{t\geq 0}$ -invariant, then the restrictions

$$T(t)|_Y$$

form a strongly continuous semigroup $(T(t)|_Y)_{t\geq 0}$, called the subspace semigroup, on the Banach space Y.

Example 9.4 (Quotient Semigroups)**.** For a closed $(T(t))_{t\geq 0}$-invariant subspace Y of X, we consider the quotient Banach space X/Y with canonical quotient map $x \mapsto \tilde{x}$. The quotient operators $T(t)$, given by

$$\tilde{T}(t)\tilde{x} \coloneqq \widetilde{T(t)x} \quad \text{for } x \in X \text{ and } t \geq 0$$

are well-defined and form a strongly continuous semigroup, called the quotient semigroup $(\tilde{T}(t))_{t\geq 0}$ on X/Y.

Example 9.5 (Adjoint Semigroups)**.** The adjoint semigroup $(T(t)^*)_{t\geq 0}$ consisting of all adjoint operators $T(t)^*$ on the dual space X^* is, in general, NOT strongly continuous. An example is provided by the (left) translation group on $L^1(\mathbb{R})$. Its adjoint is the (right) translation group on $L^\infty(\mathbb{R})$, which is not strongly continuous (see the proposition ??). However, it is easy to see that $(T(t)^*)_{t\geq 0}$ is always weak*-continuous in the sense that the maps

$$t \mapsto \langle x, T(t)^* x^* \rangle = \langle T(t)x, x^* \rangle$$

are continuous for all $x \in X$ and $x^* \in X^*$. Because on the dual of a reflexive Banach space weak and weak* topology coincide, the adjoint semigroup on such spaces is weakly, and hence by Theorem 9.4 strongly, continuous.

Proposition. *The adjoint semigroup of a strongly continuous semigroup on a reflexive Banach space is again strongly continuous.*

Example 9.6 (Product Semigroups)**.** Let $(S(t))_{t\geq 0}$ be another strongly continuous semigroup commuting with $(T(t))_{t\geq 0}$; i.e., , $S(t)T(t) = T(t)S(t)$ for all $t \geq 0$. Then the operators

$$U(t) := S(t)T(t)$$

form a strongly continuous semigroup $(U(t))_{t\geq 0}$, called the product semigroup of $(T(t))_{t\geq 0}$ and $(S(t))_{t\geq 0}$.

Clearly, $U(0) = I$. In order to show the semigroup property for $(U(t))_{t\geq 0}$, we first show that $T(s)$ and $S(r)$ commute for all $s, r \geq 0$. To this end, we first take

$r = p_1/q$ and $s = p_2/q \in \mathbb{Q}_+$. Then

$$\begin{aligned} S(r)T(s) &= S(1/q)^{p_1} \cdot T(1/q)^{p_2} \\ &= T(1/q)^{p_2} \cdot S(1/q)^{p_1} = T(s)S(r) \end{aligned}$$

i.e., $F(r,s) = G(r,s)$ for all $r, s \in \mathbb{Q}_+$, where

$$F : [0,\infty) \times [0,\infty) \to \mathcal{L}(X), \quad F(r,s) := S(r)T(s)\,,$$

and

$$G : [0,\infty) \times [0,\infty) \to \mathcal{L}(X), \quad G(r,s) := T(s)S(r)\,.$$

Now, for fixed $x \in X$, the functions $F(\cdot,\cdot)x$ and $G(\cdot,\cdot)x$ are continuous in each coordinate and coincide on $\mathbb{Q}_+ \times \mathbb{Q}_+$; hence we conclude that $F = G$ This shows that

$$S(r)T(s) = T(s)S(r)$$

for all $s, r \geq 0$, and the semigroup property $U(r+s) = U(r)U(s)$ for $s, r \geq 0$ follows immediately. Finally, the strong continuity of $(U(t))_{t\geq 0}$ is trivial.

9.2 Examples

In order to create a feeling for the concepts introduced so far, we discuss first the case in which the semigroup $(T(t))_{t\geq 0}$ can be represented as an operator-valued exponential function $\left(\mathrm{e}^{tA}\right)_{t\geq 0}$. Due to this representation, we later consider this case as rather trivial.

9.2.1 Finite-Dimensional Systems: Matrix Semigroups

We start with a reasonably detailed discussion of the finite-dimensional situation; i.e., $X = \mathbb{F}^n$. Here, $\mathcal{B}(X)$ is identified with the space $\mathrm{M}_n(\mathbb{F})$ of all $n \times n$ matrices on the scalar feild $\mathbb{F}$. Because on $\mathrm{M}_n(\mathbb{F})$ all vector topologies coincide, we simply speak of continuity of a semigroup $(T(t))_{t\geq 0}$ on X. We want to determine all continuous semigroups on $X = \mathbb{F}^n$ and start by looking at the natural examples in the form of matrix exponentials.

Proposition 9.5. *For any $A \in \mathrm{M}_n(\mathbb{F})$ and $t \geq 0$, the series*

$$\mathrm{e}^{tA} := \sum_{k=0}^{\infty} \frac{t^k A^k}{k!} \tag{9.2}$$

converges absolutely. Moreover, the mapping $t \mapsto \mathrm{e}^{tA}$; $[0,\infty) \to \mathrm{M}_n(\mathbb{F})$ is continuous and satisfies

$$\begin{cases} \mathrm{e}^{(t+s)A} = \mathrm{e}^{tA}\mathrm{e}^{sA} & \text{for } t, s \geq 0\,. \\ \mathrm{e}^{0A} = I\,. \end{cases} \tag{FE}$$

Proof. Because the series $\sum_{k=0}^{\infty} t^k \|A\|^k / k!$ converges, one can show, as for the Cauchy product of scalar series, that

$$\begin{aligned} \sum_{k=0}^{\infty} \frac{t^k A^k}{k!} \cdot \sum_{k=0}^{\infty} \frac{s^k A^k}{k!} &= \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{t^{n-k} A^{n-k}}{(n-k)!} \cdot \frac{s^k A^k}{k!} \\ &= \sum_{n=0}^{\infty} \frac{(t+s)^n A^n}{n!} \end{aligned}$$

This proves (FE). In order to show that $t \mapsto \mathrm{e}^{tA}$ is continuous, we first observe that by (FE) one has

$$\mathrm{e}^{(t+h)A} - \mathrm{e}^{tA} = \mathrm{e}^{tA}\left(\mathrm{e}^{hA} - I\right)$$

for all $t, h \in \mathbb{R}$. Therefore, it suffices to show that $\lim_{h \to 0} \mathrm{e}^{hA} = I$. This follows from the estimate

$$\begin{aligned} \left\|\mathrm{e}^{hA} - I\right\| &= \left\| \sum_{k=1}^{\infty} \frac{h^k A^k}{k!} \right\| \\ &\leq \sum_{k=1}^{\infty} \frac{|h|^k \cdot \|A\|^k}{k!} = \mathrm{e}^{|h| \cdot \|A\|} - 1\,. \end{aligned}$$

We now complete the proof. □

We call $\left(\mathrm{e}^{tA}\right)_{t \geq 0}$ the (one-parameter) semigroup generated by the matrix $A \in \mathrm{M}_n(\mathbb{F})$. In fact, there is no need to restrict the (time) parameter t to $\mathbb{R}_+$.

The definition, the continuity, and the functional equation (FE) hold for any real and even complex t. Then the map

$$T(\cdot): t \mapsto \mathrm{e}^{tA}$$

extends to a continuous (even analytic) homomorphism from the additive group $(\mathbb{R},+)$(or , $(\mathbb{F},+)$) into the multiplicative group $\mathrm{GL}(n, \mathbb{F})$ of all invertible, scalar $n \times n$ matrices. We call $\left(\mathrm{e}^{tA}\right)_{t \in \mathbb{R}}$ the (one-parameter) group generated by A.

Example 9.7. Before proceeding with the abstract theory, we might appreciate some examples of matrix semigroups.

(i) The (semi) group generated by a diagonal matrix $A=\operatorname{diag}\left(a_1, \ldots, a_n\right)$ is given by

$$\mathrm{e}^{tA}=\operatorname{diag}\left(\mathrm{e}^{t a_1}, \ldots, \mathrm{e}^{t a_n}\right) .$$

(ii) Less trivial is the case of a $k \times k$ Jordan block

$$A=\begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & 0 & \lambda \end{pmatrix}_{k \times k}$$

with eigenvalue $\lambda \in \mathbb{F}$. Decompose A into a sum $A=D+N$ where $D=\lambda I$ Then the k th power of N is zero, and the power series (9.2) (with A replaced by N) becomes

$$\mathrm{e}^{tN}=\begin{pmatrix} 1 & t & \frac{t^2}{2} & \cdots & \frac{t^{k-1}}{(k-1)!} \\ 0 & 1 & t & \cdots & \frac{t^{k-2}}{(k-2)!} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & t \\ 0 & \cdots & \cdots & 0 & 1 \end{pmatrix}_{k \times k}$$

Because D and N commute, we obtain

$$e^{tA} = e^{t\lambda}e^{tN} .$$

For arbitrary matrices A, the direct computation of e^{tA}(using the above definition) is very difficult if not impossible. Fortunately, thanks to the existence of the Jordan normal form, the following lemma shows that in a certain sense the preceding example suffice.

Lemma 9.6. *Let $B \in \mathrm{M}_n(\mathbb{F})$ and take an invertible matrix $S \in \mathrm{M}_n(\mathbb{F})$. Then the (semi) group generated by the matrix $A := S^{-1}BS$ is given by*

$$e^{tA} = S^{-1}e^{tB}S$$

Proof. Because $A^k = S^{-1}B^kS$ for all $k \in \mathbb{N}$ and because S, S^{-1} are continuous operators, we obtain

$$\begin{aligned} e^{tA} &= \sum_{k=0}^{\infty} \frac{t^k A^k}{k!} = \sum_{k=0}^{\infty} \frac{t^k S^{-1} B^k S}{k!} \\ &= S^{-1}\left(\sum_{k=0}^{\infty} \frac{t^k B^k}{k!}\right) S = S^{-1}e^{tB}S \, , \end{aligned}$$

as required. □

The content of this lemma is that similar matrices generate similar (semi) groups. Because we know that any complex $n \times n$ matrix is similar to a direct sum of Jordan blocks, we conclude that any matrix (semi) group is similar to a direct sum of (semi) groups as in Exercise 9.7 (ii).

Returning to one of the questions posed at the very beginning of this text, namely if a given semigroup can be described by a differential equation, we now proceed in two more steps. First, we show that in the case $T(t) = e^{tA}$ we even have differentiability of the map $t \mapsto T(t)$ (from $\mathbb{R}_+$ to $M_n(\mathbb{F})$), and that $T(t)$ solves the differential equation

$$\begin{cases} \frac{d}{dt}T(t) = AT(t) \text{ for } t \geq 0 \, , \\ T(0) = I \, . \end{cases} \tag{DE}$$

Next, we show that a general continuous operator semigroup on $\mathbb{F}^n$ is even differentiable in $t = 0$ and is the exponential of its derivative at $t = 0$.

Proposition 9.7. *Let $T(t) := \mathrm{e}^{tA}$ for some $A \in \mathrm{M}_n(\mathbb{F})$. Then the function $T(\cdot) : \mathbb{R}_+ \to \mathrm{M}_n(\mathbb{F})$ is differentiable and satisfies the differential equation (DE). Conversely, every differentiable function $T(\cdot) : \mathbb{R}_+ \to \mathrm{M}_n(\mathbb{F})$ satisfying (DE) is already of the form $T(t) = \mathrm{e}^{tA}$ for $A := T'(0) \in \mathrm{M}_n(\mathbb{F})$.*

Proof. We only show that $T(\cdot)$ satisfies (DE). Because the functional equation (FE) in implies

$$\frac{T(t+h) - T(t)}{h} = \frac{T(h) - I}{h} \cdot T(t)$$

for all $t, h \in \mathbb{R}$, (DE) is proved if $\lim_{h \to 0} \frac{T(h)-I}{h} = A$. This, however, follows because

$$\begin{aligned}\left\| \frac{T(h) - I}{h} - A \right\| &\leq \sum_{k=2}^{\infty} \frac{|h|^{k-1} \cdot \|A\|^k}{k!} \\ &= \frac{\mathrm{e}^{|h| \cdot \|A\|} - 1}{|h|} - \|A\| \to 0 \qquad \text{as } h \to 0\end{aligned}$$

Conversely, if $\{T(t)\}$ solves the differential equation, then we have

$$\begin{aligned}\frac{d}{dt}\left(T(t)e^{-tA}\right) &= \left(\frac{d}{dt}T(t)\right)e^{-tA} + T(t)\left(\frac{d}{dt}e^{-tA}\right) \\ &= T(t)Ae^{-tA} + M(t)(-A)e^{-tA} = 0\,.\end{aligned}$$

Combine this with the initial value condition, the desired result follows. ☐

Theorem 9.8. *Let $T(\cdot) : \mathbb{R}_+ \to \mathrm{M}_n(\mathbb{F})$ be a continuous function satisfying (FE). Then there exists $A \in \mathrm{M}_n(\mathbb{F})$ such that*

$$T(t) = \mathrm{e}^{tA} \qquad \textit{for all } t \geq 0\,.$$

Proof. Because $T(\cdot)$ is continuous and $T(0) = I$ is invertible, the matrices

$$V(t) := \int_0^t T(s)ds$$

are invertible for sufficiently small $t \geq 0$. (use that $\lim_{t\downarrow 0} V(t)/t = T(0) = I$). Taking $t_0 > 0$ so that $V(t_0)$ is invertible, the functional equation (FE) now yields

$$\begin{aligned} T(t) &= V(t_0)^{-1} V(t_0) T(t) = V(t_0)^{-1} \int_0^{t_0} T(t+s)ds \\ &= V(t_0)^{-1} \int_t^{t+t_0} T(s)ds = V(t_0)^{-1} \left(V(t+t_0) - V(t)\right) \end{aligned}$$

for all $t \geq 0$. Clearly, $V(t)$ is differentiable. Hence, $T(\cdot)$ is differentiable with derivative

$$\begin{aligned} \frac{d}{dt}T(t) &= \lim_{h\downarrow 0} \frac{T(t+h) - T(t)}{h} \\ &= \lim_{h\downarrow 0} \frac{T(h) - T(0)}{h} T(t) = T'(0)T(t) \quad \text{for all } t \geq 0\,. \end{aligned}$$

This shows that $T(\cdot)$ satisfies (DE) with $A = T'(0)$. □

9.2.2 Uniformly Continuous Operator Semigroups

We now desire to extend the above results to semigroups $(T(t))_{t\geq 0}$ on an infinite-dimensional Banach space X over $\mathbb{F}$. To this purpose, it suffices to assume continuity of the map $t \mapsto T(t) \in \mathcal{B}(X)$ in the operator norm. Then we can replace the matrix $A \in \mathrm{M}_n(\mathbb{F})$ by a bounded operator $A \in \mathcal{B}(X)$ and argue as in the last subsection.

For $A \in \mathcal{B}(X)$ we define

$$\mathrm{e}^{tA} := \sum_{n=0}^{\infty} \frac{t^n A^n}{n!} \tag{9.3}$$

for each $t \geq 0$. It follows from the completeness of X that e^{tA} is a well-defined bounded operator on X.

The proof of the following two propositions can be adapted from the last subsection and is omitted.

Proposition 9.9. $\left(\mathrm{e}^{tA}\right)_{t\geq 0}$ *is a semigroup on X such that $t \mapsto \mathrm{e}^{tA}\,;\, \mathbb{R}_+ \to (\mathcal{B}(X), \|\cdot\|)$ is continuous.*

Semigroups having the continuity property stated above are called *uniformly continuous* (or *norm-continuous*) .

Proposition 9.10. *For $A \in \mathcal{B}(X)$ define $\left(\mathrm{e}^{tA}\right)_{t\geq 0}$ by (9.3). The map $t \mapsto T(t) := \mathrm{e}^{tA}\,; \mathbb{R}_+ \to (\mathcal{B}(X), \|\cdot\|)$ is differentiable and satisfies the differential equation*

$$\begin{cases} \frac{d}{dt}T(t) = AT(t) \text{ for } t \geq 0\,, \\ T(0) = I\,. \end{cases} \tag{DE}$$

Conversely, every differentiable function $T(\cdot) : \mathbb{R}_+ \to (\mathcal{B}(X), \|\cdot\|)$ satisfying (DE) is already of the form $T(t) = \mathrm{e}^{tA}$ for $A = T(0) \in \mathcal{B}(X)$.

Theorem 9.11. *Every uniformly continuous semigroup $(T(t))_{t\geq 0}$ on a Banach space X is of the form*

$$T(t) = \mathrm{e}^{tA}, \quad t \geq 0$$

for some bounded operator $A \in \mathcal{B}(X)$.

Proof. Because the following arguments were already used in the matrixvalued cases , a brief outline of the proof should be sufficient. For a uniformly continuous semigroup $(T(t))_{t\geq 0}$ the operators

$$V(t) := \int_0^t T(s)ds, \quad t \geq 0$$

are well-defined, and $V(t)/t$ converges (in norm!) to $T(0) = I$ as $t \downarrow 0$. Hence, for $t > 0$ sufficiently small, the operator $V(t)$ becomes invertible. Repeat now the computations from the proof of Theorem 9.8 in order to obtain that $t \mapsto T(t)$ is differentiable and satisfies (DE). Then Proposition 9.10 yields the assertion. □

Before adding some comments on and further properties of uniformly continuous semigroups we state the following question leading directly to the main objects of this text.

> Do there exist ”natural” one-parameter semigroups of linear operators on Banach spaces that are not uniformly continuous, hence not of the form $\left(\mathrm{e}^{tA}\right)_{t\geq 0}$ for some bounded operator A?

Remark 9.1. The operator A in Theorem 9.11 is determined uniquely as the derivative of $T(\cdot)$ at zero; i.e., $A = \dot{T}(0)$. We call it the *generator* of $(T(t))_{t\geq 0}$.

Remark 9.2. Because definition for e^{tA} works also for $t \in \mathbb{R}$ and even for $t \in \mathbb{C}$ it follows that each uniformly continuous semigroup can be extended to a uniformly continuous group $\left(\mathrm{e}^{tA}\right)_{t\in\mathbb{R}}$, or to $\left(\mathrm{e}^{tA}\right)_{t\in\mathbb{C}}$, respectively.

Remark 9.3. From the differentiability of $t \mapsto T(t)$ it follows that for each $x \in X$ the orbit map $t \mapsto T(t)x\,;\, \mathbb{R}_+ \to X$ is differentiable as well. Therefore, the map $x(t) := T(t)x$ is the unique solution of the X-valued initial value problem (or abstract Cauchy problem)

$$\begin{cases} \dot{x}(t) = Ax(t) \text{ for } t \geq 0\,. \\ x(0) = x\,. \end{cases} \tag{ACP}$$

9.3 More Semigroups

In order to convince usthat new and interesting phenomena appear for semigroups on infinite-dimensional Banach spaces, we first discuss several classes of one-parameter semigroups on concrete spaces. These semigroups are not uniformly continuous anymore and hence, unlike those in the last section, not of the form $\left(\mathrm{e}^{tA}\right)_{t\geq 0}$ for some bounded operator A. On the other hand, they are not "pathological" in the sense of being completely unrelated to any analytic structure. In addition, these semigroups accompany us through the further development of the theory and provide a source of illuminating examples and counterexamples.

9.3.1 Multiplication Semigroups on $\mathrm{C}_0(\Omega)$

Multiplication operators can be considered as an infinite-dimensional generalization of *diagonal matrices*. They are extremely simple to construct, and most of their properties are evident. Nevertheless, their value should not be underestimated. They appear, for example, naturally in the context of Fourier analysis or when one applies the spectral theorem for self-adjoint operators on Hilbert spaces (see Theorem ??). We therefore strongly recommend that any first attempt to illustrate a result or disprove a conjecture on semigroups should be made using multiplication semigroups.

We start from a locally compact Hausdorff space Ω and define the Banach space (endowed with the sup-norm) $\mathrm{C}_0(\Omega)$ of all continuous, $\mathbb{F}$-valued functions on Ω that vanish at infinity. As a typical example we might take Ω to be a (unbounded) connected open set in $\mathbb{R}^d$.

With any continuous function $q : \Omega \to \mathbb{F}$ we associate a linear operator M_q on $\mathrm{C}_0(\Omega)$ defined on its "maximal domain" $D(M_q)$ in $\mathrm{C}_0(\Omega)$. Specifically, let $M_q f \coloneqq q \cdot f$, for all f in the domain

$$D(M_q) \coloneqq \{f \in \mathrm{C}_0(\Omega) : q \cdot f \in \mathrm{C}_0(\Omega)\} .$$

The main feature of these multiplication operators is that most operator-theoretic properties of M_q can be characterized by analogous properties of the function q. In the following proposition we give some examples for this correspondence.

Proposition 9.12. *Let M_q with domain $D(M_q)$ be the multiplication operator induced on $\mathrm{C}_0(\Omega)$ by some continuous function q. Then the following assertions hold.*

(i) The operator $(M_q, D(M_q))$ is closed and densely defined.

(ii) The operator M_q is bounded (with $D(M_q) = \mathrm{C}_0(\Omega)$) if and only if the function q is bounded. In that case, one has

$$\|M_q\| = \|q\| := \sup_{s\in\Omega} |q(s)|$$

(iii) The operator M_q has a bounded inverse if and only if the function q has a bounded inverse $1/q$; i.e., $0 \notin \overline{q(\Omega)}$. In that case, one has

$$M_q^{-1} = M_{1/q}$$

(iv) The spectrum of M_q is the closed range of q; i.e.,

$$\sigma\left(M_q\right)=\overline{q(\Omega)}$$

Appendix A

Basic Topology

A.1 Zorn's Lemma

Sometimes one wants to prove the existence of a mathematical object (which can be viewed as a maximal element in some partially ordered oset). One could try proving the existence of such an object by assuming there is no maximal element and using transfinite induction and the assumptions of the situation to get a contradiction. Zorn's lemma tidies up the conditions a situation needs to satisfy in order for such an argument to work. Therefore Zorn's lemma enables mathematicians to not have to repeat the transfinite induction argument by hand each time, but just check the conditions of Zorn's lemma.

Definition A.1. A binary relation $\preceq$ on a nonempty set P is a *partial order* if it satisfies the following properties: For all $x, y, z \in P$

(a) $\preceq$ is *reflexive*: $x \preceq x$;

(b) $\preceq$ is *antisymmetric*: if $x \preceq y$ and $y \preceq x$, then $x = y$;

(c) $\preceq$ is *transitive*: if $x \preceq y$ and $y \preceq z$, then $x \preceq z$.

In this case, say that $(P, \preceq)$ is a *partially ordered set.*

We give some examples for partial ordered set. Let $P = \mathbb{R}$ and take $\preceq$ to be $\leq$, the usual less than or equal to relation on $\mathbb{R}$; Let $P = \mathcal{P}(X)$ the power set of a set X and take $\preceq$ to be $\subseteq$, the usual set inclusion relation. Let $P = \mathcal{C}[0,1]$, the space of continuous real-valued functions on the interval [0,1] and take $\preceq$ to be the relation $\leq$ given by $f \leq g$ if and only if $f(x) \leq g(x)$ for each $x \in [0,1]$.

Definition A.2. Let C be a subset of a partially ordered set $(P, \preceq)$. An element $u \in P$ is an *upper bound* of C if $x \preceq u$ for every $x \in C$. An element $m \in C$ is said to be *maximal* if for any element $y \in C$, the relation $m \preceq y$ implies that $m = y$.

Definition A.3. Let $(P, \preceq)$ be a partially ordered set and $x, y \in P$. We say that x and y are *comparable* if either $x \preceq y$ or $y \preceq x$. Otherwise, x and y are *incomparable.* A partial order $\preceq$ is called a *total order* if any two elements of P are comparable. In this case we say that $(P, \preceq)$ is a or totally ordered set. A totally ordered set is also called a *chain.*

Theorem A.1 (Zorn's lemma)**.** *Let $(P, \preceq)$ be a partially ordered set. If each totally ordered subset of P has an upper bound, then P has a maximal element.*

Proof. A sketch of the proof of Zorn's lemma follows, assuming the axiom of choice. Suppose for contradiction that the lemma is false. Then for every element in P there is another element bigger than it. For every totally ordered subset T we may then define a bigger element $b(T)$, because T has an upper bound, and that upper bound has a bigger element. To actually define the function b, we need to employ the axiom of choice.

Using the function b, we are going to define elements $a_0 \prec a_1 \prec a_2 \prec a_3 \prec \dots$ in P (where $u \prec v$ means that $u \preceq v$ and $u \neq v$). This sequence is really long: the indices are not just the natural numbers, but all ordinals. In fact, the sequence is too long for the set P, there are too many ordinals (a proper class), more than there are elements in any set, and the set P will be exhausted before long and then we will run into the desired contradiction.

The a_i are defined by transfinite recursion: we pick a_0 in P arbitrary, and for any other ordinal w we set $a_w = b(\{a_v : v < w\})$. Because the a_v are totally ordered, this is a well-founded definition. $\square$

Remark A.1. This proof shows that actually a slightly stronger version of Zorn's lemma is true: If P is a poset in which every well-ordered subset has an upper bound, and if x is any element of P, then P has a maximal element greater than or equal to x. That is, there is a maximal element which is comparable to x.

A.2 Metric space

A.2.1 Compact Subset of a Metric Space

Definition A.4. 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.5. 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.6. 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.2. 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 $B = B(0, 1) = \{x \in X : \|x\| \leq 1\}$. B is bounded but not totally bounded.

Theorem A.2. *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.7. 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.3. *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.2, 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.2, (x_n) has a Cauchy subsequence (x_{n_k}). since K is complete, (x_{n_k}) converges in K. Hence K is sequentially compact. $\square$

Definition A.8. 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.4. *(X, d) is a metric space and $K \subset X$. Then K is compact if and only if K is sequentially compact.*

Proof. $\square$

A.2.2 Baire's Category Theorem

Definition A.9. 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.5. *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 to 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, $G = \bigcap_{n \in \mathbb{N}} G_n \neq \emptyset$. It is also clear that $y \in B(x, \epsilon)$. □

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.6 (Baire's Catagory Theorem)**.** *A complete metric space (X, d) is of second category in itself.*

A.3 不可数级数的和

Definition A.10. 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.3. 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

$$\|\sum_{i \in T} \alpha_i - \beta\| \leq \epsilon$$

Proposition A.7. *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.8. 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

$$\|\sum_{i\in J}\alpha_i\|\leq\epsilon$$

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.9. $\sum_{i\in I}\alpha_i$ 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

$$\|\sum_{i\in J}\alpha_i\|\leq\frac{1}{n}$$

令 $S=\bigcup_{n=1}^{\infty}S_n$. 可见只有 $i\notin S$ 时必然有 $\alpha_i=0$. □

Proposition A.10. 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 □

Proposition A.11. $x_i\in\mathbb{R}$, 且 $x_i\geq 0\, i\in I$. 则 $\sum_{i\in I}x_i$ 收敛于 $y\in\mathbb{R}$ 当且仅当

$$y=\sup\{\sum_{i\in S}x_i : S \text{ 是} I \text{ 有限子集}\}.$$