Lecture Notes on Functional Analysis

Xiaojing Ye

Contents

These notes outline the materials covered in class. Detailed derivations and explanations are given in lectures and/or the referenced books. The notes will be continuously updated with additional content and corrections. Questions and comments can be addressed to `xye@gsu.edu`.

1 Metric Space

1.1 Contractive mapping

Definition 1.1 (Metric space)**.** A space (X, d) is called a *metric space* if X is a set and $d : X \times X \to \mathbb{R}$ satisfies

1. $d(x, y) \geq 0$; and $d(x, y) = 0$ iff $x = y$.
2. $d(x, y) = d(y, x)$.
3. $d(x, z) \leq d(x, y) + d(y, z)$.

for any $x, y, z \in X$. Here d is called the *metric* (or *distance*) on X. We may drop d and simply write X if the metric is clear from the context.

Example 1.2. Let $X = \mathbb{R}^n$ and $d(x, y) = ((x_1 - y_1)^2 + \cdots + (x_n - y_n)^2)^{1/2}$ for $x, y \in \mathbb{R}^n$. Then (X, d) is a metric space.

Example 1.3. Let $X = C([a, b]) := \{x : [a, b] \to \mathbb{R} : x \text{ is continuous}\}$ and $d(x, y) = \max_{a \leq t \leq b} |x(t) - y(t)|$. Then (X, d) is a metric space.

With metric defined, we can consider the concept of "convergence", as below.

Definition 1.4 (Convergence)**.** Let $\{x_k\}$ be a sequence in X, then $\{x_k\}$ is said to *converge to* x if $d(x_k, x) \to 0$ as $k \to \infty$. Namely, for any $\epsilon > 0$, there exists $K \in \mathbb{N}$, such that $d(x_k, x) < \epsilon$ for all $k \geq K$. We may also write this as $\lim_k d(x_k, x) = 0$ or $x_k \to x$.

Example 1.5. The convergence in Example 1.3 is the "uniform convergence" of continuous functions.

Definition 1.6 (Closed set)**.** The subset E of X is called *closed* if $\{x_k\} \subset E$ and $x_k \to x$ imply that $x \in E$.

Definition 1.7 (Cauchy sequence)**.** A sequence $\{x_k\} \subset X$ is called *Cauchy* if for any $\epsilon > 0$, there exists $K \in \mathbb{N}$, such that $d(x_k, x_j) < \epsilon$ for all $k, j \geq K$.

Definition 1.8 (Complete metric space)**.** A metric space (X, d) is called *complete* if every Cauchy sequence in X is convergent.

Remark. Let (X, d) be complete and $Y \subset X$. Then (Y, d) is complete iff Y is closed in X.

Example 1.9. The Euclidean space $(\mathbb{R}^n, d)$ is complete.

Example 1.10. The space $(C([a, b]), d)$ defined in Example 1.3 is complete.

Proof. Let $\{x_k\}$ be Cauchy, then for any $\epsilon > 0$, there exists $K \in \mathbb{N}$, such that

$$d(x_k, x_j) = \max_{a \leq t \leq b} |x_k(t) - x_j(t)| < \epsilon$$

for any $k, j \geq K$. Hence, for any t, $\{x_k(t)\}$ is a Cauchy sequence in $\mathbb{R}$ and thus convergent. Let $x(t) := \lim_k x_k(t)$. Then

$$|x_k(t) - x(t)| = \lim_{j \to \infty} |x_k(t) - x_j(t)| \leq \epsilon$$

for any $t \in [a,b]$. This implies that $x_k \to x$ uniformly. Hence x is continuous (as $[a,b]$ is compact and $\{x_k\}$ are continuous) and thus $x \in C([a,b])$. □

Definition 1.11 (Continuous mapping)**.** Let (X,d) and (Y,ρ) be two metric spaces. A mapping $T : (X,d) \to (Y,\rho)$ is called *continuous* if $\rho(Tx_k, Tx) \to 0$ in Y whenever $d(x_k, x) \to 0$ in X as $k \to \infty$. The set of continuous mappings from X to Y is denoted by $C(X;Y)$. If $Y = \mathbb{R}$ we also write it as $C(X)$.

Theorem 1.12. *Let (X,d) and (Y,ρ) be metric spaces. Then $T : X \to Y$ is continuous iff for any $\epsilon > 0$ and $x \in X$, there exists $\delta := \delta(x,\epsilon) > 0$, such that $\rho(Ty, Tx) < \epsilon$ whenever $d(y,x) < \delta$.*

Proof. ($\Rightarrow$) If not, then there exist $x \in X$, $\epsilon_0 > 0$, and a sequence $\{x_k\}$ in X, such that $d(x_k, x) < 1/k$ but $\rho(Tx_k, Tx) \geq \epsilon_0$ for all $k \in \mathbb{N}$, which contradicts to the continuity of T.

($\Leftarrow$) Let $x \in X$ be arbitrary. If the sequence $\{x_k\}$ is such that $d(x_k, x) \to 0$, then for any $\epsilon > 0$, there exists $K \in \mathbb{N}$ such that $d(x_k, x) < \delta := \delta(\epsilon, x)$ for all $k \geq K$. Thus $\rho(Tx_k, Tx) < \epsilon$, for all $k \geq K$. This implies that T is continuous. □

Definition 1.13 (Contractive mapping)**.** Let (X,d) be a metric space. A mapping $T : X \to X$ is called *contractive* if there exists $\theta \in (0,1)$ such that $d(Tx, Ty) \leq \theta d(x,y)$ for all $x, y \in X$. Contractive mappings are also called *contractions*.

Example 1.14. Let $T : [0,1] \to [0,1]$ be C^1 and $|T'(x)| \leq \theta < 1$ for all $x \in [0,1]$. Then T is a contractive mapping.

Proof. Note that $X = [0,1]$ and $d(x,y) = |x - y|$ for all $x, y \in X$. Hence, for any $x, y \in X$,

$$d(Tx, Ty) = |T(x) - T(y)| = |T'(c)(x - y)| \leq \theta|x - y| = \theta d(x,y)$$

where $c \in (x,y)$ due to the mean value theorem. Therefore T is contractive. □

Theorem 1.15. *If $T : X \to X$ is contractive, then T is continuous.*

Proof. Suppose $d(x_k, x) \to 0$, then $d(Tx_k, Tx) \leq \theta d(x_k, x) \to 0$ as $k \to \infty$. □

Theorem 1.16 (Banach fixed point theorem)**.** *Let (X,d) be complete. If $T : X \to X$ is contractive, then T has a unique fixed point x in X.*

Proof. Pick any $x_0 \in X$ and generate a sequence $\{x_k\}$ by $x_{k+1} := Tx_k$ for all $k \in \mathbb{N}$. Then

$$d(x_{k+1}, x_k) = d(Tx_k, Tx_{k-1}) \leq \theta d(x_k, x_{k-1}) \leq \cdots \leq \theta^k d(x_1, x_0).$$

Hence, for any $p \in \mathbb{N}$, we have

$$\begin{aligned} d(x_{k+p}, x_k) &= \sum_{i=1}^{p} d(x_{k+i}, x_{k+i-1}) \le \sum_{i=1}^{p} \theta^{k+i-1} d(x_1, x_0) \\ &\le \sum_{i=1}^{\infty} \theta^{k+i-1} d(x_1, x_0) = \frac{\theta^k}{1-\theta} d(x_1, x_0) \to 0 \end{aligned}$$

as $k \to \infty$. Therefore $\{x_k\}$ is Cauchy. As X is complete, we know there exists $x \in X$ such that $x_k \to x$. Moreover, we know $x_{k+1} = Tx_k \to x$, and hence $d(Tx, x) = \lim_k d(Tx_k, x_k) = 0$, which implies that $x = Tx$, i.e., x is a fixed point of T.

If both x and x' are fixed points of T but are different, then $d(x, x') > 0$ but

$$d(x, x') = d(Tx, Tx') \le \theta d(x, x') < d(x, x'),$$

which is a contradiction. Hence the fixed point of T is unique. □

Example 1.17 (Existence and uniqueness of the solution of ODE)**.** Consider the initial value problem of an ordinary differential equation (ODE):

$$\text{(ODE)} \qquad \begin{cases} x'(t) = f(t, x(t)), & t \in \mathbb{R}, \\ x(0) = \xi. \end{cases}$$

Suppose there exists $\delta > 0$ such that $f(t, x) : U \to \mathbb{R}$, where $U := [-h, h] \times [\xi - \delta, \xi + \delta]$, is continuous, and f is Lipschitz in x over U, i.e., there exists an $L > 0$ such that $|f(t, x) - f(t, y)| \le L|x - y|$ for any $t \in [-h, h]$ and $x, y \in [\xi - \delta, \xi + \delta]$. Denote $M := \max_{(t,x)\in U} |f(t, x)|$. If $0 < h < \min(\delta/M, 1/L)$, then the ODE has a unique solution on $[-h, h]$.

Proof. We denote $\bar{B}(\xi; \delta) := \{x : [-h, h] \to \mathbb{R} : |x(t) - \xi| \le \delta,\ \forall\, t \in [-h, h]\}$, which is a closed subset of $C([-h, h])$ (here $\bar{B}(\xi; \delta)$ should be interpreted as the closed ball in $C([a, b])$ with the constant function ξ as the center and δ as the radius). We define the mapping $T : C([-h, h]) \to C([-h, h])$ by

$$(Tx)(t) := \xi + \int_0^t f(s, x(s))\, \mathrm{d}s.$$

Hence x is a solution of the ODE iff $x = Tx$.

We first show that T maps $\bar{B}(\xi; \delta)$ to itself. To this end, for any $x \in \bar{B}(\xi; \delta)$, we know

$$|(Tx)(t) - \xi| = \left| \int_0^t f(s, x(s))\, \mathrm{d}s \right| \le M|t| \le Mh < \delta,$$

for all $t \in [-h, h]$, where the first inequality above is due to $x \in \bar{B}(\xi; \delta)$ (so $x(s) \in [\xi - \delta, \xi + \delta]$ for all s) and the definition of M, and the last inequality is due to the condition on h. Therefore $Tx \in \bar{B}(\xi; \delta)$.

Next we show that T is contractive on $\bar{B}(\xi;\delta)$. To this end, we first note that $Tx, Ty \in \bar{B}(\xi;\delta)$ for any $x, y \in \bar{B}(\xi;\delta)$ as shown above. Then

$$\begin{aligned} d(Tx,Ty) &= \max_{|t|\le h} |(Tx)(t) - (Ty)(t)| = \max_{|t|\le h} \left| \int_0^t (f(s,x(s)) - f(s,y(s)))\,\mathrm{d}s \right| \\ &\le \max_{|t|\le h} \int_0^t |f(s,x(s)) - f(s,y(s))|\,\mathrm{d}s \le \max_{|t|\le h} \int_0^t L|x(s)-y(s)|\,\mathrm{d}s \\ &\le Lh\,d(x,y). \end{aligned}$$

Since $Lh < 1$, we know T is contractive.

Finally, we see that $(\bar{B}(\xi;\delta), d)$ is complete as $\bar{B}(\xi;d)$ is a closed ball in a complete metric space $(C([-h,h]), d)$, and hence there exists a unique fixed point $x \in \bar{B}(\xi;\delta)$ by Theorem 1.16 (Banach fixed point). This implies that x is the unique solution of the ODE on $[-h,h]$. □

Example 1.18 (Implicit function theorem)**.** Let $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$. Suppose there exists $(x_0, y_0) \in \mathbb{R}^n \times \mathbb{R}^m$ such that f and $\partial_y f$ are continuous in the open neighborhood $U \times V$ of (x_0, y_0), where $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^m$, and

$$f(x_0, y_0) = 0, \quad \text{and} \quad \partial_y f(x_0, y_0) \text{ is invertible.}$$

Then there exist an open set $U_0 \times V_0$ such that $(x_0, y_0) \in U_0 \times V_0 \subset U \times V$ and a unique function $\phi : U_0 \to \mathbb{R}^m$ satisfying $\phi(x_0) = y_0$ and $f(x, \phi(x)) = 0$ for all $x \in U_0$.

Proof. For any $r > 0$ sufficiently small (we will specify the range of r later), we consider the mapping T defined by

$$(T\phi)(x) := \phi(x) - (\partial_y f(x_0, y_0))^{-1} f(x, \phi(x))$$

for any $\phi \in C(\bar{B}(x_0; r); \mathbb{R}^m)$. We also define the following metric in the space $C(\bar{B}(x_0; r); \mathbb{R}^m)$:

$$d(\phi, \psi) := \max_{x \in \bar{B}(x_0; r)} |\phi(x) - \psi(x)|_\infty$$

where $|y|_\infty := \max_{1 \le i \le m} |y_i|$ for any $y = (y_1, \dots, y_m) \in \mathbb{R}^m$.

Since $\partial_y f(x, y)$ is continuous in $U \times V$, we know there exists $\delta > 0$ such that the magnitudes of all entries of $I - (\partial_y f(x_0, y_0))^{-1} \partial_y f(x, y)$ are upper bounded by $1/(2m)$ for all $x \in \bar{B}(x_0; \delta)$ and $y \in \bar{B}(y_0; \delta)$.

Now let $\zeta(x) := \phi(x) - \psi(x)$. Note that by the mean value theorem, for any $x \in \bar{B}(x_0; r) \subset U$, there exists $z(x)$ lying between $\phi(x)$ and $\psi(x)$ (in the sense that there exists $\lambda(x) \in (0,1)$ such that $z(x) = (1-\lambda(x))\phi(x) + \lambda(x)\psi(x)$), such that

$$f(x, \phi(x)) - f(x, \psi(x)) = \partial_y f(x, z(x))\,\zeta(x), \quad \forall\, x \in \bar{B}(x_0; r).$$

Moreover, if $r < \delta$, then

$$\begin{aligned} d(T\phi, T\psi) &= \max_{x\in\bar{B}(x_0;r)} |\zeta(x) - (\partial_y f(x_0, y_0))^{-1}(f(x, \phi(x)) - f(x, \psi(x)))|_\infty \\ &= \max_{x\in\bar{B}(x_0;r)} |[I - (\partial_y f(x_0, y_0))^{-1}\partial_y f(x, z(x)))]\zeta(x)|_\infty \\ &< \max_{x\in\bar{B}(x_0;r)} \frac{1}{2}|\zeta(x)|_\infty = \frac{1}{2}d(\phi, \psi). \end{aligned}$$

Now we consider the closed subset X of $C(\bar{B}(x_0; r); \mathbb{R}^m)$ as follows,

$$X := \{\phi \in C(\bar{B}(x_0; r); \mathbb{R}^m) : \phi(x_0) = y_0,\ \phi(x) \in \bar{B}(y_0, \delta),\ \forall\, x \in \bar{B}(x_0; \delta)\}.$$

We already showed that T is contractive on X if $r < \delta$ above. Now we need to show that T is a mapping from X to itself for r small enough. To this end, we know by the continuity of f over $U \times V$ that there exists $\eta > 0$ such that $|(\partial_y f(x_0, y_0))^{-1}(f(x, y) - f(x, y_0))| < \delta/2$ for all $x \in \bar{B}(x_0; \eta)$ and $y \in \bar{B}(y_0; \eta)$. Furthermore, consider y_0 as the constant function taking value y_0, we have

$$\begin{aligned} d(T\phi, y_0) &\le d(T\phi, Ty_0) + d(Ty_0, y_0) \\ &< \frac{1}{2}d(\phi, y_0) + \max_{x\in\bar{B}(x_0;\eta)} |(\partial_y f(x_0, y_0))^{-1} f(x, y_0)|_\infty \\ &= \frac{1}{2}d(\phi, y_0) + \max_{x\in\bar{B}(x_0;\eta)} |(\partial_y f(x_0, y_0))^{-1}(f(x, y_0) - f(x_0, y_0))|_\infty \\ &< \frac{1}{2}d(\phi, y_0) + \frac{1}{2}\delta < \delta \end{aligned}$$

for $x \in B(x_0; \eta)$, where the equality is due to $f(x_0, y_0) = 0$. So $d(T\phi, y_0) < \delta$ if $0 < r < \min(\delta, \eta)$. Furthermore,

$$(T\phi)(x_0) = \phi(x_0) - (\partial_y f(x_0, y_0))^{-1} f(x_0, \phi(x_0)) = y_0$$

since $\phi(x_0) = y_0$ and $f(x_0, y_0) = 0$. Hence T is a mapping from X to itself.

Finally, since (X, d) is complete and $T : X \to X$ is contractive, we know there exists a unique $\phi \in X$ such that $\phi = T\phi$, which implies that $\phi(x_0) = y_0$ and $f(x, \phi(x)) = 0$ in $U_0 := \bar{B}(x_0; r)$. □

1.2 Complete metric space

Recall that completeness of (X, d) is necessary in Theorem 1.16 (Banach fixed point), otherwise the conclusion may not hold, as in the following example.

Example 1.19. Consider $T : [0, 1] \to [0, 1]$ defined by $Tx = \sqrt{x + 1}$. Then T is a contractive mapping and the unique fixed point is $x_0 = (\sqrt{17} + 1)/18$. It is easy to verify that T is a contractive mapping on $X := [0, 1] \setminus \{x_0\}$, but then it does not have a fixed point on X.

The completeness of a metric space also depends on the metric, as shown in the following example.

Example 1.20. Consider $(C([a,b]),\rho)$ where ρ is defined by

$$\rho(x,y) := \int_a^b |x(t) - y(t)|\,\mathrm{d}t.$$

Then it is easy to verify that ρ is a metric. However $(C([a,b]),\rho)$ is not complete. To see this, recall that $C([a,b])$ is dense in $L([a,b])$, but not all functions in $L([a,b])$ are continuous.

Definition 1.21 (Isometry)**.** Let (X,d_X) and (Y,d_Y) be two metric spaces. If there exists a mapping $T : X \to Y$ such that

1. T is surjective,
2. $d_X(x,x') = d_Y(Tx,Tx')$ for all $x,x' \in X$,

then we say that (X,d_X) and (Y,d_Y) are *isometric*, and T is called an *isometry*.

Remark. Note that item 2 above implies that T is injective: if $Tx = Tx'$, then $d_X(x,x') = d_Y(Tx,Tx') = 0$, which implies that $x = x'$. Therefore an isometry is also a one-to-one correspondence.

If (X,d_X) and (Y,d_Y) are isometric, then they have the same properties regarding metrics (distances). If (X,d_X) is isometric to (Y,d_Y), and (Y,d_Y) is a subspace of (Z,d_Z) such that $d_Z|_Y = d_Y$ (i.e., d_Z is identical to d_Y when restricted to Y), then we say (X,d_X) is isometrically embedded into (Z,d_Z).

Definition 1.22 (Dense subset)**.** Let (X,d) be a metric space. Then a subset E of X is called dense if for any $x \in X$ and $\epsilon > 0$, there exists $y \in E$ such that $d(x,y) < \epsilon$.

Example 1.23. Let $P([a,b])$ be the set of polynomials on $[a,b]$, then $P([a,b])$ is dense in $C([a,b])$ (due to Weierstrass Theorem) under the standard norm d in $C([a,b])$. Also, $C([a,b])$ is dense in $L([a,b])$ under the metric ρ defined in Example 1.20.

Definition 1.24 (Completion)**.** Let (X,d) be a metric space, and $\mathcal{F} := \{(Y,\rho) : X \subset Y,\ \rho|_X = d,\ Y \text{ is complete}\}$. Then $\bar{X} := \cap_{Y\in\mathcal{F}} Y$ is a complete metric space, called the *completion* of X. $\bar{X}$ is the smallest complete space containing X in the sense that $\bar{X} \subset Y$ for all $Y \in \mathcal{F}$.

Theorem 1.25. *If $X \subset Y$ where $Y \in \mathcal{F}$ and X is dense in Y, then Y is the completion of X.*

Proof. Let (X,d) be a metric space which is dense in (Y,ρ) and $\rho|_X = d$. Then we know for any $y \in Y$, there exists a sequence $(x_1,\dots,x_k,\dots)$ in X such that $x_k \to y$, i.e., $\rho(x_k,y) \to 0$ as $k \to \infty$. If $(Z,\varrho) \in \mathcal{F}$, then by noting that $\{x_k\}$ is Cauchy in Z, we know there exists $z \in Z$ such that $\varrho(x_k,z) \to 0$.

Now define $T : Y \to Z$ such that $Ty = z$. We shall show that T is an isometry. To this end, for any $y' \in Y$, there exists a sequence $(x'_1,\cdots,x'_k,\cdots)$ in X such that $\rho(x'_k,y') \to 0$. Hence

$$\rho(y,y') = \lim_{k\to\infty} \rho(x_k,x'_k) = \lim_{k\to\infty} \varrho(x_k,x'_k) = \varrho(Ty,Ty')$$

where the second equality is due to $\rho|_X = \varrho|_X = d$ and the third equality is due to the definition of T. Hence (Y, ρ) is a subspace of (Z, ϱ). Since Y is complete, we know (Y, ρ) is the completion of (X, d) by definition. □

Theorem 1.26. *Every metric space has a completion.*

Proof. 1. For a metric space (X, d), we consider the set

$$\tilde{Y} := \{x = (\xi_1, \xi_2, \dots) \text{ which is a Cauchy sequence in } X\}.$$

Define an equivalence relation $\sim$ in $\tilde{Y}$ as follows: $x \sim x'$ iff $\lim_k d(\xi_k, \xi'_k) = 0$. Let $Y := \tilde{Y}/\sim$ be the quotient space and $[x] \in Y$ stands for the equivalence class of x. Then define $\rho : Y \times Y \to \mathbb{R}_+$ by

$$\rho([x], [x']) = \lim_{k\to\infty} d(\xi_k, \xi'_k).$$

It is easy to verify that ρ is a metric on Y: in particular,

$$\begin{aligned}\rho([x], [x']) = \lim_{k\to\infty} d(\xi_k, \xi'_k) &\le \lim_{k\to\infty} d(\xi_k, \xi''_k) + d(\xi''_k, \xi'_k) \\ &= \rho([x], [x'']) + \rho([x''], [x']),\end{aligned}$$

for any $[x], [x'], [x''] \in Y$ where x, x', x'' are Cauchy sequences in X. This verifies the triangle inequality. Hence (Y, ρ) is a metric space.

2. We now show that (X, d) is dense in (Y, ρ). To this end, we identify $\xi \in X$ with $[x^\xi] \in Y$, where $x^\xi := (\xi, \xi, \dots)$ is the constant sequence in X. We denote this identification mapping $T : X \to Y$ by $T\xi = x^\xi$. Then clearly $T : X \to T(X)$ is surjective and $\rho(T\xi, T\eta) = \rho([x^\xi], [\xi^\eta]) = d(\xi, \eta)$. Hence $T : X \to T(X) \subset Y$ is an isometry, X and $T(X)$ are isometric, and (X, d) is embedded into (Y, ρ). For any $[y] \in Y$, where $y = (\eta_1, \eta_2, \dots) \in \tilde{Y}$ is a Cauchy sequence in X, we know $(T\eta_1, T\eta_2, \dots)$ is a sequence in $T(X)$ such that

$$\lim_{j\to\infty} \rho(T\eta_j, [y]) = \lim_{j\to\infty} \rho([x^{\eta_j}], [y]) = \lim_{j\to\infty} \lim_{k\to\infty} d(\eta_j, \eta_k) = 0.$$

Hence $T(X)$ is dense in Y.

3. Finally we need to show that (Y, d_Y) is complete. Suppose $([y_1], [y_2], \dots)$ is a Cauchy sequence in Y. For each $k \in \mathbb{N}$, there exists $\xi_k \in X$ such that $\rho(T\xi_k, [y_k]) < 1/k$ (since $T(X)$ is dense in Y). Then for any $\epsilon > 0$, there exists $K = K(\epsilon) \in \mathbb{N}$ such that $2/k < \epsilon/2$ and $\rho([y_{k+p}], [y_k]) < \epsilon/2$ for all $k \ge K$ and $p \in \mathbb{N}$. Hence

$$\begin{aligned}d(\xi_{k+p}, \xi_k) = \rho(T\xi_{k+p}, T\xi_k) &\le \rho(T\xi_{k+p}, [y_{k+p}]) + \rho([y_{k+p}], [y_k]) + \rho([y_k], T\xi_k) \\ &\le \frac{1}{k+p} + \rho([y_{k+p}], [y_k]) + \frac{1}{k} \le \frac{2}{K} + \frac{\epsilon}{2} < \epsilon\end{aligned}$$

for all $k \ge K$. So $x = (\xi_1, \xi_2, \dots)$ is Cauchy in X and hence $[x] \in Y$. It is then easy to show that $\rho([y_k], [x]) \to 0$.

Combining the conclusions in the previous three steps yields that (Y, ρ) is the completion of (x, d), by Theorem 1.25. □

Example 1.27. The completion of $(P([a,b]),d)$ is $(C([a,b]),d)$ where $d(x,y) := \max_{a\le t\le b}|x(t)-y(t)|$.

Example 1.28. The completion of $(C([a,b]),\rho)$ is $(L([a,b]),\rho)$ where $\rho(x,y) := \int_a^b |x(t)-y(t)|\,\mathrm{d}t$.

1.3 Sequentially compact set

Definition 1.29 (Bounded set)**.** Let (X,d) be a metric space, a subset E is called *bounded* if there exists $x_0 \in X$ and $r > 0$ such that $E \subset B(x_0; r)$.

Recall that in $(\mathbb{R}^n, d)$ (where d is the standard Euclidean distance), if an infinite subset E is bounded, then E has at least one limit point (also known as accumulation point); if a sequence $\{x_k\}$ is bounded, then it has at least one convergent subsequence. However, this is *not* true for general complete metric spaces, as shown in the following example.

Example 1.30. Consider $(C([0,1]),d)$ and the following sequence $\{x_k\}$:

$$x_k(t) := \begin{cases} 0, & t \ge 1/k, \\ 1-tk, & 0 \le t < 1/k, \end{cases}$$

then $\{x_k\}$ is bounded by 1 in $C([0,1])$, but $\{x_k\}$ does not have a convergent subsequence (the limit is not in $C([0,1])$).

Definition 1.31 (Sequentially precompact and compact)**.** Let (X,d) be a metric space, then a subset E is called *sequentially precompact* if every sequence in E has a Cauchy subsequence. If in addition all the Cauchy sequences converge with limits in E, then E is called *sequentially compact.* If X is sequentially compact, we call X a *sequentially compact space.*

Remark. It is obvious that a sequentially compact set is also sequentially precompact. In some texts, a sequentially precompact set is also called *Cauchy precompact* or *relatively compact.*

Theorem 1.32. *The following statements hold:*

- *If E is sequentially precompact and $F \subset E$, then F is sequentially precompact.*
- *If E is sequentially precompact, then $\bar{E}$ is sequentially compact (assuming the metric space is complete).*
- *If E is sequentially compact and F is a closed subset of E, then F is sequentially compact.*

Theorem 1.33. *A sequentially compact metric space is complete.*

Proof. Let (X,d) be a sequentially compact metric space and $\{x_k\}$ a Cauchy sequence in X. Then there exist $x \in X$ and a subsequence $\{x_{k_j}\}$ such that $x_{k_j} \to x$ as $j \to \infty$. Therefore $x_k \to x$ as $k \to \infty$ since $\{x_k\}$ is Cauchy. Hence X is complete. $\square$

We now introduce a condition stronger than boundedness.

Definition 1.34 (ϵ-net)**.** Let (X, d) be a metric space, $E \subset X$, and $\epsilon > 0$. We call $N^\epsilon \subset X$ an *ϵ-net of E* if for any $x \in E$, there exists $y \in N^\epsilon$, such that $d(x, y) < \epsilon$. In other words,

$$E \subset \bigcup_{y \in N^\epsilon} B(y; \epsilon).$$

If N^ϵ is finite, then we call N^ϵ a *finite ϵ-net of E*.

Definition 1.35 (Totally bounded)**.** Let (X, d) be a metric space and E a subset of X. Then E is called *totally bounded* if for any $\epsilon > 0$, E has a finite ϵ-net in X. Totally bounded sets are also called *precompact*.

Remark. In some texts, a set E is called totally bounded if for any $\epsilon > 0$, E has a finite ϵ-net *in E* (rather than in X as in Definition 1.35). This alternative definition is equivalent to Definition 1.35: it is trivial to show from this alternative definition to Definition 1.35; Conversely, if E is totally bounded under Definition 1.35, then for any $\epsilon > 0$, E has a finite $(\epsilon/2)$-net $N^{\epsilon/2} = \{x_1, \dots, x_{n_\epsilon}\}$ in X, pick any $y_i \in E \cap B(x_i; \epsilon/2)$ (WLOG it is nonempty) for each $i = 1, \dots, n_\epsilon$, then there is

$$E \subset \bigcup_{i=1}^{n_\epsilon} B\Big(x_i; \frac{\epsilon}{2}\Big) \subset \bigcup_{i=1}^{n_\epsilon} B(y_i; \epsilon),$$

which means that $\{y_1, \dots, y_{n_\epsilon}\} \subset E$ is a finite ϵ-net of E.

Remark. The following facts can be verified easily:

- If E is totally bounded, then any subset of E is totally bounded.
- If E is totally bounded, then $\bar{E}$ is also totally bounded. To see this, let $\epsilon > 0$ be arbitrary and $N^{\epsilon/2}$ a finite $(\epsilon/2)$-net of E, then $N^{\epsilon/2}$ is a finite ϵ-net of $\bar{E}$.

Theorem 1.36 (Hausdorff)**.** *Let (X, d) be a metric space and E a subset of X, then E is sequentially precompact iff E is totally bounded.*

Proof. ($\Rightarrow$) If not, then there exists $\epsilon_0 > 0$, such that E does not have a finite ϵ_0-net. Pick any $x_1 \in X$. Then pick $x_2 \in E \setminus B(x_1; \epsilon_0)$ (this is nonempty since E cannot be covered by $B(x_1; \epsilon_0)$). Then pick $x_3 \in E \setminus (B(x_1; \epsilon_0) \cup B(x_2; \epsilon_0))$, and so on. This process will never stop since there is no finite ϵ_0-net of E. Then we obtain $\{x_k\}$ which is a sequence in E but $d(x_k, x_j) \geq \epsilon_0$ for any $k \neq j$. Therefore $\{x_k\}$ does not have any Cauchy subsequence in E, which contradicts to E being sequentially precompact.

($\Leftarrow$) Let $\{x_k\}$ be a sequence in E. Since E has a finite 1-net, we know there exists $y_1 \in M$ such that $B(y_1; 1)$ contains infinitely many terms (thus a subsequence) of $\{x_k\}$, which we denote by $\{x_k^{(1)}\}$. Since E has a finite $(1/2)$-net, there exists $y_2 \in E$, such that $B(y_2; 1/2)$ contains a subsequence of $\{x_k^{(1)}\}$, which we denote by $\{x_k^{(2)}\}$. Continue doing so, we obtain $\{x_k^{(j)}\}$ for $j = 1, 2, \dots$, which is a sequence of sequences and forms an infinite matrix with $(x_1^{(j)}, x_2^{(j)}, \dots)$ as the jth row. Note that $x_k^{(j)} \in B(y_J; 1/J)$ for all $j \geq J$ and $J, k \in \mathbb{N}$. Now consider the sequence $\{x_k^{(k)}\}$ (by extracting the diagonal terms as a sequence), we know for any $\epsilon > 0$, if $k > 2/\epsilon$, then for any $p \in \mathbb{N}$, there is

$$d(x_{k+p}^{(k+p)}, x_k^{(k)}) \leq d(x_{k+p}^{(k+p)}, y_k) + d(x_k^{(k)}, y_k) < \frac{1}{k} + \frac{1}{k} = \frac{2}{k} < \epsilon.$$

Therefore, $\{x_k^{(k)}\}$ is a Cauchy subsequence of $\{x_k\}$. □

Remark. If in addition (X, d) is complete in Theorem 1.36, then E is sequentially compact iff E is totally bounded and closed:

($\Rightarrow$) Since E is sequentially compact, we know E is sequentially precompact and by Theorem 1.36 that E is totally bounded. Moreover, every Cauchy subsequence in E is convergent with limit in E, which implies that E is closed.

($\Leftarrow$) Since E is totally bounded, we know by Theorem 1.36 that E is sequentially precompact. As E is closed and X is complete, we know E is complete. Hence every Cauchy sequence in E also converges in E, from which we know E is sequentially compact.

Definition 1.37 (Separable space)**.** Let (X, d) be a metric space. We call X *separable* if X has a countable dense subset.

Theorem 1.38 (Totally bounded spaces are separable)**.** *Let (X, d) be a totally bounded metric space, then X is separable.*

Proof. Let $N_k \subset X$ be a finite $(1/k)$-net. Then $\cup_{k=1}^{\infty} N_k$ is a countable dense subset of X. □

Definition 1.39 (Compact)**.** Let (X, d) be a metric space, then a subset E of X is called *compact* if any open cover of E has a finite subcover.

Remark. Recall that a compact set is closed and bounded. However, the converse is true in Euclidean spaces but not in general metric spaces. The following theorem states that compactness and sequential compactness are equivalent in metric spaces.

Theorem 1.40. *Let (X, d) be a metric space and $E \subset X$. Then E is compact iff E is sequentially compact.*

Proof. ($\Rightarrow$) Suppose E is compact, then E is closed. If E is not sequentially compact, then there exists a sequence $\{x_k\}$ in E such that $\{x_k\}$ does not have a Cauchy (convergent) subsequence. Hence $S := \{x_k\}$, as a set, is closed. Let $S_k := S \setminus \{x_k\}$ (removing x_k from S), which is also closed. Then $X \setminus S_k$ is open, and thus

$$\bigcup_{k=1}^{\infty} (X \setminus S_k) = X \setminus \bigcap_{k=1}^{\infty} S_k = X \supset E,$$

i.e., $\{X \setminus S_k : k \in \mathbb{N}\}$ is an open cover of E. As E is compact, there exists a finite subcover, i.e., there exists $K \in \mathbb{N}$, such that $E \subset \cup_{k=1}^{K}(X \setminus S_k) = X \setminus \cap_{k=1}^{K} S_k$. However, since $x_{K+1} \notin X \setminus \cap_{k=1}^{K} S_k$ and $x_{K+1} \in E$, we arrive at a contradiction.

($\Leftarrow$) Suppose E is sequentially compact, and hence totally bounded. Assume E is not compact, then there exists an open cover $\{G_\alpha : \alpha \in \mathcal{A}\}$ of E, but it does not have a finite subcover for E. Since E is totally bounded, we know for any $k \in \mathbb{N}$, there exists a finite $(1/k)$-net N_k, such that $E \subset \cup_{y \in N_k} B(y; 1/k)$ (covered by finitely many balls). Hence there exists $y_k \in N_k$ such that $B(y_k; 1/k)$ has no finite subcover from $\{G_\alpha\}$.

Now consider $\{y_k\}$. Since E is sequentially compact, we know it contains a convergent subsequence, denoted by $\{y_{k_j}\}$, such that $y_{k_j} \to y$ for some $y \in E$. Then there exists $\alpha' \in \mathcal{A}$ and $\delta > 0$, such that $y \in B(y; \delta) \subset G_{\alpha'}$. Hence there exists $k_j \in \mathbb{N}$ large enough, such that $B(y_{k_j}; 1/k_j) \subset B(y; \delta) \subset G_{\alpha'}$, which contraditcts to the definition of y_{k_j} (that $B(y_{k_j}; 1/k_j)$ does not have a finite subcover from $\{G_\alpha\}$). $\square$

Now we consider a generalization of $C([a, b])$. Suppose (M, d) is a complete compact metric space. Define

$$C(M) := \{f : M \to \mathbb{R} : T \text{ is continuous}\}$$

and ρ by $\rho(u, v) := \max_{x \in M} |u(x) - v(x)|$.

Theorem 1.41. *$(C(M), \rho)$ is a metric space.*

Proof. We only need to verify that ρ is well defined, i.e., the maximum in the definition of ρ can be attained. Let $\{x_k\}$ be a sequence in M, such that $|u(x_k) - v(x_k)| \to \sup_{x \in M} |u(x) - v(x)|$. Since M is compact, we know there exists a subsequence x_{k_j} such that $x_{k_j} \to x_0$ for some $x_0 \in M$. On the other hand, as u and v are continuous, we know $|u(x_{k_j}) - v(x_{k_j})| \to |u(x_0) - v(x_0)|$. Hence $|u(x_0) - v(x_0)| = \sup_{x \in M} |u(x) - v(x)|$, i.e., the maximum can be attained. $\square$

Theorem 1.42. *$(C(M), \rho)$ is complete.*

Proof. Similar to the proof of completeness of $C([a, b])$ before. $\square$

Definition 1.43 (Uniformly bounded)**.** A subset F of $C(M)$ is called *uniformly bounded* if there exists $L > 0$ such that $|f(x)| \leq L$ for all $x \in M$ and $f \in F$.

Definition 1.44 (Equicontinuous)**.** A subset F of $C(M)$ is called *equicontinuous* if for any $\epsilon > 0$, there exists $\delta := \delta(\epsilon) > 0$, such that

$$|f(x_1) - f(x_2)| < \epsilon$$

for any $x_1, x_2 \in M$ satisfying $d(x_1, x_2) < \delta$ and any $f \in F$.

Theorem 1.45 (Arzelà-Ascoli)**.** *Let (M, d) be a compact metric space and $(C(M), \rho)$ be defined as above. Then a set $F \subset C(M)$ is sequentially precompact iff F is uniformly bounded and equicontinuous.*

Proof. Note that (M, d) is sequentially compact and hence also complete. Since $(C(M), \rho)$ is a metric space, we know that F is sequentially precompact iff F is totally bounded. Thus we prove the theorem with sequential compactness replaced by total boundedness.

($\Rightarrow$) Suppose F is totally bounded, then F is uniformly bounded. We also need to show that F is equicontinuous. To this end, for any $\epsilon > 0$, we know F has a finite $(\epsilon/3)$-net, i.e., there exist $\phi_1, \dots, \phi_K \in F$, such that $F \subset \cup_{k=1}^{K} B(\phi_k; \epsilon/3)$. For every $k \in \{1, \dots, K\}$, there exists $\delta_k := \delta_k(\epsilon) > 0$, such that $|\phi_k(x) - \phi_k(x')| < \epsilon/3$ for all $x, x' \in M$ satisfying $d(x, x') < \delta_k$. Let $\delta := \min\{\delta_1, \dots, \delta_K\} > 0$, then for any $x, x' \in M$ satisfying $d(x, x') < \delta$, and any $f \in F$, there exists $k \in \{1, \dots, K\}$, such that $f \in B(\phi_k; \epsilon/3)$, and hence

$$|f(x) - f(x')| \le |f(x) - \phi_k(x)| + |\phi_k(x) - \phi_k(x')| + |\phi_k(x') - f(x')| < 3 \cdot \frac{\epsilon}{3} = \epsilon.$$

Hence F is equicontinuous.

($\Leftarrow$) For any $\epsilon > 0$, we need to find a finite ϵ-net for F. Since F is equicontinuous, there exists $\delta = \delta(\epsilon) > 0$, such that for any $x, x' \in M$ satisfying $d(x, x') < \delta$ and any $f \in F$, there is $|f(x) - f(x')| < \epsilon/3$. As (M, d) is compact, we know M is sequentially compact and totally bounded. Let $\{x_1, \dots, x_n\}$ be a finite δ-net of M, i.e., $M \subset \cup_{i=1}^{n} B(x_i; \delta)$. Now consider the mapping $T : C(M) \to \mathbb{R}^n$ defined by $T(\psi) = (\psi(x_1), \dots, \psi(x_n))$ for any $\psi \in C(M)$. It is clear that T is surjective.

Since F is uniformly bounded, i.e., there exists $L > 0$, such that $|f(x)| \le L$ for all $x \in M$ and $f \in F$, we know that

$$|T(f)| = |(f(x_1), \dots, f(x_n))| \le \sqrt{n} L.$$

So $T(F)$ is bounded in $\mathbb{R}^n$. Thus $\overline{T(F)}$ is compact and thus totally bounded, and hence $T(F)$ is totally bounded and has a finite $(\epsilon/3)$-net. Since T is surjective, we know there exist $\phi_1, \dots, \phi_m \in F$, such that $\{T(\phi_1), \dots, T(\phi_m)\}$ is a finite $(\epsilon/3)$-net of $T(F)$.

For any $x \in M$, there exists $i \in \{1, \dots, n\}$, such that $x \in B(x_i; \delta)$, and $j \in \{1, \dots, m\}$, such that $T(f) \in B(T\phi_j; \epsilon/3)$. Hence

$$\begin{aligned} |f(x) - \phi_j(x)| &\le |f(x) - f(x_i)| + |f(x_i) - \phi_j(x_i)| + |\phi_j(x_i) - \phi_j(x)| \\ &\le \frac{\epsilon}{3} + |T(f) - T(\phi_j)| + \frac{\epsilon}{3} < \epsilon, \end{aligned}$$

which means $\{\phi_1, \dots, \phi_m\}$ is a finite ϵ-net of F. □

Remark. Under the setting of Theorem 1.45 (Arzelà-Ascoli), $(C(M), \rho)$ is complete. Therefore, if F is uniformly bounded and equicontinuous, then any sequence in F has a Cauchy subsequence (in the sense of ρ), and the subsequence converges to some $f \in C(M)$. If F is closed, then the limit $f \in F$.

Example 1.46. Let $\Omega \subset \mathbb{R}^n$ be open and convex, and $L_0, L_1 > 0$. Then the set $F := \{f \in C(\bar{\Omega}) \cap C^1(\Omega) : |f(x)| \le L_0, |\nabla f(x)| \le L_1, \forall\, x \in \Omega\}$ is sequentially precompact in $C(\bar{\Omega})$.

Proof. By the definition of F, we know F is uniformly bounded by L_0. By the mean value theorem, for any $x, x' \in \bar{\Omega}$, there exists $\lambda \in (0,1)$, such that

$$|f(x) - f(x')| = |\nabla f(\lambda x + (1-\lambda)x)(x - x')| \leq L_1|x - x'|$$

Hence f is L_1-Lipschitz continuous for all $f \in F$. Hence F is equicontinuous. By Theorem 1.45 (Arzelà-Ascoli), we know F is sequentially precompact. $\square$

Remark. Note that F in Example 1.46 is not closed, and hence not necessarily sequentially compact. For example, consider $F := \{f_k \in C([-1,1]) : f_k(x) = (x^2 + \frac{1}{k})^{1/2}\}$. Then f_k converges uniformly to $f(x) := |x|$, but $f \notin F$ since f is not differentiable at 0.

Example 1.47. Let (X, d) be a metric space and E a sequentially compact subset of X. Suppose $f : X \to E$ satisfies $d(f(x_1), f(x_2)) < d(x_1, x_2)$ for all distinct $x_1, x_2 \in X$. Show that f has a unique fixed point in X.

Proof. It is obvious that f is continuous. Denote $\delta = \inf\{d(x, f(x)) : x \in E\}$. Let $x_k \in M$ such that $\delta \leq d(x_k, f(x_k)) < \delta + 1/k$ for all $k \in \mathbb{N}$. Then by sequential compactness of M, there exists a subsequence $x_{k_j} \to x$ for some $x \in E$. Hence $d(x_{k_j}, f(x_{k_j})) \to d(x, f(x)) = \delta$ as $j \to \infty$ because f is continuous. If $\delta > 0$, then $d(f(x), f(f(x))) < d(x, f(x)) = \delta$ which contradicts to the definition of δ. So $d(x, f(x)) = \delta = 0$.

If x, x' are two fixed points of f but distinct, then

$$0 < d(x, x') = d(f(x), f(x')) < d(x, x'),$$

which is a contradiction. Hence $x = x'$. $\square$

2 Banach and Hilbert Space

2.1 Normed linear space and Banach space

We have discussed metric spaces which have topological structure due to the the metric (and thus we have the concepts of open sets, closed sets, compact sets etc.) However, this is often insufficient, and we also need to consider the algebraic structure of the spaces.

Definition 2.1 (Linear space)**.** Let X be a set and $\mathbb{K}$ be a field (either $\mathbb{R}$ or $\mathbb{C}$). We call X a *linear space* if it has summation and scalar multiplication defined, such that for any $x, y, z \in X$ and $\alpha, \beta \in \mathbb{K}$, the following statements hold.

1. $x + y = y + x$.
2. $(x + y) + z = x + (y + z)$.
3. There exists $0 \in X$ such that $x + 0 = 0 + x$ for all $x \in X$.
4. For any $x \in X$, there exists $x' \in X$, called $-x$, such that $x + x' = 0$.
5. $\alpha(\beta x) = (\alpha\beta)x$.
6. $1 \cdot x = x$ where $1 \in \mathbb{K}$.
7. $(\alpha + \beta)x = \alpha x + \beta x$.
8. $\alpha(x + y) = \alpha x + \beta y$.

Remark. The elements in X are also called *vectors*, and X is also called *vector space*.

Definition 2.2. We have the following definitions regarding linear spaces.

- (Isomorphism) Let X and Y be linear spaces, and $T : X \to Y$ be a linear mapping, i.e.,

  $$T(\alpha x + \beta x') = \alpha T(x) + \beta T(x')$$

  for all $x, x' \in X$ and $\alpha, \beta \in \mathbb{K}$. Then T is called an *isomorphism* if T is a one-to-one correspondence (injective and surjective, or bijective). In this case, X and Y are called *isomorphic* to each other.
- (Linear subspace) If Y is a subset of X and closed under the summation and scalar multiplication, then Y is called a *linear subspace* of X. Note that $\{0\}$ and X are trivial linear subspaces of X.
- (Linear manifold) Let Y be a linear subspace of X, then $x + Y := \{x + y : y \in Y\}$ is called a *linear manifold*, i.e., a translation of Y by x.
- (Linear independency) A set of vectors $\{x_1, \ldots, x_m\}$ of X are called *linearly independent* if $a_1 x_1 + \cdots + a_m x_m = 0$ implies that $a_1 = \cdots = a_m = 0 \in \mathbb{K}$. Otherwise they are called *linearly dependent*. A set E is called linearly independent if any finite subset of E is linearly independent.
- (Linear basis) A subset E of X is called a *linear basis* if E is linearly independent and any $x \in X$ can be written as linear combination of vectors in X (check that such combination is unique). The cardinality $|E|$ is called the dimension of X.
- (Linear span) Let $E = \{x_\alpha : \alpha \in \mathcal{A}\}$, then the span of E is defined by

  $$\text{span}(E) := \{a_1 x_{\alpha_1} + \cdots + a_n x_{\alpha_n} : a_i \in \mathbb{K},\ \alpha_i \in \mathcal{A},\ n \in \mathbb{N}\}.$$

- (Sum and direct sum) Let E_1, E_2 be linear subspaces of X, then the sum of E_1 and E_2 is $E_1 + E_2 := \{x_1 + x_2 : x_1 \in E_1, x_2 \in E_2\}$. If $E_1 \cap E_2 = \{0\}$, then the sum becomes the *direct sum*, denoted by $E_1 \oplus E_2$. In the case of direct sum, for any vector $x \in E_1 \oplus E_2$, there exist unique $x_1 \in E_1$ and $x_2 \in E_2$, such that $x = x_1 + x_2$ (check yourself).

Now we combine the algebraic structure and topological structure of (X, d) by requiring:

1. The metric d is translation invariant: $d(x + z, y + z) = d(x, y)$ for all $x, y, z \in X$.
2. Continuity in scalar multiplication:

$$\begin{aligned} d(x_k, x) \to 0 \quad &\Longrightarrow \quad d(ax_k, ax) \to 0, \quad \forall\, a \in \mathbb{K}. \\ a_k \to 0 \quad &\Longrightarrow \quad d(a_k x, ax) \to 0, \quad \forall\, x \in X. \end{aligned}$$

Note that by Item 1 above, $d(x, y) = d(x - y, 0)$.

Proposition 2.3. *If $d(x_k, x) \to 0$ and $d(y_k, y) \to 0$, then $d(x_k + y_k, x + y) \to 0$.*

Proof. We have that

$$\begin{aligned} d(x_k + y_k, x + y) &= d((x_k + y_k) - (x + y), 0) = d((x_k - x) + (y_k - y), 0) \\ &= d(x_k - x, y - y_k) \le d(x_k - x, 0) + d(y_k - y, 0) \\ &= d(x_k, x) + d(y_k, y) \to 0, \end{aligned}$$

which completes the proof. □

Now we can introduce $p : X \to \mathbb{R}_+$ by $p(x) = d(x, 0)$. Then it is easy to verify that the following properties of p hold:

1. (Positive definite) $p(x) \ge 0$; and $p(x) = 0$ iff $x = 0$.
2. (Symmetric) $p(x) = p(-x)$ for all $x \in X$.
3. (Subadditive) $p(x + y) \le p(x) + p(y)$ for all $x, y \in X$.
4. $p(x_k) \to 0$ implies $p(ax_k) \to p(ax)$ for all $a \in \mathbb{K}$.
5. $a_k \to 0$ implies $p(a_k x) \to p(ax)$ for all $x \in X$.

Items 1–3 are due to that d is a metric. Items 4–5 are due to the continuity in scalar multiplication.

Definition 2.4 (Frechét space)**.** We call (X, p) a *Frechét space* if the linear space X is complete and $p : X \to \mathbb{R}$ satisfies the items 1–5 above. Note that completeness means that for any Cauchy sequence is convergent in X in the sense of p.

Example 2.5. $\mathbb{R}^n$ with $p(x) = (x_1^2 + \cdots + x_n^2)^{1/2}$ is a Frechét space.

Example 2.6. Consider $C(M)$ where (M, d) is a compact metric space. Define $p : C(M) \to \mathbb{R}$ by

$$p(u) := \max_{x \in M} |u(x)|,$$

then $(C(M), p)$ is a Frechét space (check Items 1–5 above).

We need the following simple lemma for the next example.

Lemma 2.7. *For any $a, b \geq 0$, the following inequalties hold:*

1. *If $a \leq b$, then $\frac{a}{1+a} \leq \frac{b}{1+b}$.*
2. $\frac{a+b}{1+a+b} \leq \frac{a}{1+a+b} + \frac{b}{1+a+b} \leq \frac{a}{1+a} + \frac{b}{1+b}$.
3. $\frac{ab}{1+ab} \leq \max\{a, 1\} \frac{b}{1+b}$.

Proof. The first two can be verified directly. The last one is due to

$$\frac{ab}{1+ab} \leq \begin{cases} a \cdot \frac{b}{1+b}, & \text{if } a \geq 1, b \geq 0, \\ \frac{b}{1+b}, & \text{if } 0 \leq a < 1, b \geq 0. \end{cases}$$

from which the claimed inequality follows. □

Example 2.8. Consider $S := \{x = (x_1, x_2, \dots) : x_i \in \mathbb{R}\}$ where the summation and scalar multiplication are as usual. For any $x \in X$, define

$$p(x) = \sum_{k=1}^{\infty} \frac{1}{2^k} \frac{|x_k|}{1 + |x_k|}.$$

Then (S, p) is a Frechét space.

Proof. 1. It is straightforward to check that S is a linear space.

2. We need to check Items 1–5 of p. Items 1–2 are trivial. For Item 3, we have that

$$\begin{aligned} p(x+y) &= \sum_{k=1}^{\infty} \frac{1}{2^k} \frac{|x_k + y_k|}{1 + |x_k + y_k|} \leq \sum_{k=1}^{\infty} \frac{1}{2^k} \frac{|x_k| + |y_k|}{1 + |x_k| + |y_k|} \\ &\leq \sum_{k=1}^{\infty} \frac{1}{2^k} \left(\frac{|x_k|}{1 + |x_k|} + \frac{|y_k|}{1 + |y_k|} \right) = p(x) + p(y) \end{aligned}$$

where the first two inequalities are due to the first two statements in the previous lemma, respectively.

For Item 4, suppose $p(x^{(j)}) \to 0$ as $j \to \infty$, then

$$\begin{aligned} p(ax^{(j)}) &= \sum_{k=1}^{\infty} \frac{1}{2^k} \frac{|ax^{(j)}|}{1 + |ax^{(j)}|} \leq \max(a, 1) \cdot \sum_{k=1}^{\infty} \frac{1}{2^k} \frac{|x^{(j)}|}{1 + |x^{(j)}|} \\ &= \max(a, 1) \cdot p(x^{(j)}) \to 0, \end{aligned}$$

as $j \to \infty$, where the inequality is due to the third statement in the previous lemma.

For Item 5, suppose $a_j \to 0$ as $j \to \infty$. For any $\epsilon > 0$, there exists $K \in \mathbb{N}$ large enough, such that $1/2^K < \epsilon/2$, then choose $J \in \mathbb{N}$ large enough, such that

$a_j \max_{1\le k\le K} |x_k| < \epsilon/2$ for all $j \ge J$. Then we have

$$p(a_j x) = \sum_{k=1}^{K} \frac{1}{2^k} \frac{|a_j x_k|}{1+|a_j x_k|} + \sum_{k=K+1}^{\infty} \frac{1}{2^k} \frac{|a_j x_k|}{1+|a_j x_k|}$$
$$< \frac{\epsilon}{2} \sum_{k=1}^{K} \frac{1}{2^k} + \sum_{k=K+1}^{\infty} \frac{1}{2^k} < \frac{\epsilon}{2} + \frac{1}{2^K} < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

for all $j \ge J$. Hence $p(a_j x) \to 0$ as $j \to \infty$. □

3. Now we show that (S, p) is complete. Suppose $\{x^{(n)}\}$ is Cauchy in S, i.e., $p(x^{(n+p)} - x^{(n)}) \to 0$ as $n \to \infty$ for any $p \in \mathbb{N}$. Note that

$$p(x^{(n+p)} - x^{(n)}) = \sum_{k=1}^{\infty} \frac{1}{2^k} \frac{|x_k^{(n+p)} - x_k^{(n)}|}{1+|x_k^{(n+p)} - x_k^{(n)}|} \to 0,$$

which implies that $\{x_k^{(n)}\}$ is Cauchy for every k. Suppose $x_k^{(n)} \to x_k^*$. Let $x^* := (x_1^*, x_2^*, \cdots)$. Then for any $\epsilon > 0$, there exists $K \in \mathbb{N}$, such that $1/2^K < \epsilon/2$. Choose $N \in \mathbb{N}$ large enough, such that $|x_k^{(n)} - x_k^*| < \epsilon/2$ for all $n \ge N$ and $k = 1, \ldots, K$. Then

$$p(x^{(n)} - x^*) = \sum_{k=1}^{K} \frac{1}{2^k} \frac{|x_k^{(n)} - x_k^*|}{1+|x_k^{(n)} - x_k^*|} + \sum_{k=K+1}^{\infty} \frac{1}{2^k} \frac{|x_k^{(n)} - x_k^*|}{1+|x_k^{(n)} - x_k^*|}$$
$$< \frac{\epsilon}{2} \sum_{k=1}^{K} \frac{1}{2^k} + \sum_{k=K+1}^{\infty} \frac{1}{2^k} < \frac{\epsilon}{2} + \frac{1}{2^K} < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

As we can see, the p function of the Frechét spaces in Examples 2.5 and 2.6 are homogeneous: $p(ax) = |a|p(x)$ for all $a \in \mathbb{R}$ and $x \in X$, whereas the one in Example 2.8 is not. We formalize the definition of norm by requiring homogeneity as below.

Definition 2.9 (Norm)**.** The function $\|\cdot\| : X \to \mathbb{R}$ is called a *norm* if for any $x, y, z \in X$ and $a \in \mathbb{K}$ the following statements hold:
1. (Positive definite) $\|x\| \ge 0$; and $\|x\| = 0$ iff $x = 0$.
2. (Homogeneous) $\|ax\| = |a|\|x\|$.
3. (Triangle inequality) $\|x + y\| \le \|x\| + \|y\|$.

Remark. Homogeneity of norm also implies that $\|ax_k\| \to 0$ whenever $\|x_k\| \to 0$ and $\|a_k x\| \to 0$ whenever $a_k \to 0$.

Example 2.10 (Norm is continuous)**.** Suppose $x_k \to x$, then $\|x_k - x\| \to 0$, which implies that $|\|x_k\| - \|x\|| \le \|x_k - x\| \to 0$. Hence $\|x_k\| \to \|x\|$.

Definition 2.11 (Banach space)**.** Let X be a linear space and $\|\cdot\|$ be a norm. Then we call $(X, \|\cdot\|)$, or simply X is the norm is clear from the context, a *B^* space*. If X is complete, we call X a *Banach space*.

Example 2.12. Let $(X, \mathcal{M}, \mu)$ be a measure space where $\Omega \subset \mathbb{R}^n$, and $u : \Omega \to \mathbb{R}$ be a measurable function. Define $p \in [1, \infty)$ by

$$\|u\|_p := \Big(\int_X |u(x)|^p \, \mathrm{d}\mu(x) \Big)^{1/p}.$$

Then $L^p(\Omega, \mu) := \{u : \Omega \to \mathbb{R} : \|u\|_p < \infty\}$. We can verify that L^p is a complete metric space, namely,

1. $L^p(\Omega)$ is a linear space under summation and scalar multiplication.
2. $\| \cdot \|_p$ is a norm (the triangle inequality is due to Minkowski's inequality).
3. L^p is complete (Theorem 6.15 in Lecture Notes on Real Analysis).

Example 2.13. Let $S := \{x = (x_1, x_2, \dots) : x_k \in \mathbb{R}\}$. Then define

$$\|x\|_p := \Big(\sum_{k=1}^{\infty} |x_k|^p \Big)^{1/p}.$$

This is called the l^p space. Then we can verify that $(l^p, \|\cdot\|)$ is a complete metric space. Specifically, the triangle inequality is due to the generalized Minkowski's inequality, and the completeness as follows: let $\{x^n\}$ be a Cauchy sequence in l^p, then

$$\lim_{n,m\to\infty} \|x^{(m)} - x^{(n)}\|_p^p = \lim_{n,m\to\infty} \sum_{k=1}^{\infty} |x_k^{(m)} - x_k^{(n)}|^p = 0.$$

Hence for every $k \in \mathbb{N}$, $\{x_k^{(n)}\}$ is Cauchy. Let $x_k^* := \lim_{n\to\infty} x_k^{(n)}$. Then for any $\epsilon > 0$, there exists $N = N(\epsilon) \in \mathbb{N}$, such that $\|x^{(n)} - x^{(n+m)}\|_p^p < \epsilon$ for all $n \geq N$ and $m \in \mathbb{N}$. For any fixed $k \in \mathbb{N}$, we know

$$\begin{aligned} \sum_{k=1}^{K} |x_k^{(n)} - x_k^*|^p &= \lim_{m\to\infty} \sum_{k=1}^{K} |x_k^{(n)} - x_k^{(m)}|^p \leq \lim_{m\to\infty} \sum_{k=1}^{\infty} |x_k^{(n)} - x_k^{(m)}|^p \\ &= \lim_{m\to\infty} \|x^{(n)} - x^{(m)}\|_p^p \leq \epsilon. \end{aligned}$$

Hence we have

$$\|x^{(n)} - x^*\|_p^p = \sum_{k=1}^{\infty} |x_k^{(n)} - x_k^*|^p \leq \epsilon.$$

which means $x^{(n)} \to x^*$. Moreover, $\|x^*\|_p \leq \|x^{(n)} - x^*\|_p + \|x^{(n)}\|_p < \infty$, hence $x^* \in S$. Therefore, $(S, \|\cdot\|_p)$ is a Banach space.

Example 2.14. Let Ω be an open, bounded, connected subset of $\mathbb{R}^n$ and $k \in \mathbb{N}$. For a multi-index $\alpha = (\alpha_1, \dots, \alpha_n) \in (\mathbb{N} \cup \{0\})^n$, we denote $|\alpha| := \sum_{i=1}^n \alpha_i$ and

$$\partial^\alpha u(x) := \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}} u(x), \quad \forall\, x = (x_1, \dots, x_n) \in \Omega.$$

We define the norm

$$\|u\| := \max_{|\alpha| \leq k} \max_{x \in \bar{\Omega}} |\partial^\alpha u(x)|.$$

Then it is straightforward to verify that $\|\cdot\|$ is a norm on $C^k(\bar{\Omega})$. We now show that $C^k(\bar{\Omega})$ is complete under this norm. To this end, let $\{u_k\}$ be a Cauchy sequence in $C^k(\bar{\Omega})$. Then for any α where $|\alpha| \le k$, there exists v_α such that $\partial^\alpha u_k \rightrightarrows v_\alpha$ (converge uniformly) as $k \to \infty$ due to the definition of the norm. It remains to show that $\partial^\alpha u = v_\alpha$ (where $u := v_{(0,\dots,0)}$ is the limit of u_k for short) for all $|\alpha| \le k$. For any $x = (x_1, \dots, x_n) \in \Omega$, let $(x_1^0, x_2, \dots, x_n)$ be in the neighborbood of x,. Then for any k, we have

$$u_k(x_1, x_2, \dots, x_n) = u_k(x_1^0, x_2, \dots, x_n) + \int_{x_1^0}^{x_1} \partial_1 u_k(t, x_2, \dots, x_n)\,\mathrm{d}t.$$

As $u_k \rightrightarrows u$ and $\partial_1 u \rightrightarrows v_{(1,0,\dots,0)}$, we know

$$u(x_1, x_2, \dots, x_n) = u(x_1^0, x_2, \dots, x_n) + \int_{x_1^0}^{x_1} v_{(1,0,\dots,0)}(t, x_2, \dots, x_n)\,\mathrm{d}t.$$

Taking partial derivative with respect to x_1 on both sides yields that $\partial_1 u = v_{(1,0,\dots,0)}$. The cases for other α can be obtained similarly by induction.

Furthermore, we have that $\|u\| \le \|u_k - u\| + \|u_k\| < \infty$. Hence $(C^k(\bar{\Omega}), \|\cdot\|)$ is complete, and thus is a Banach space.

Example 2.15 (Sobolev space)**.** Let Ω be an open, bounded, and connected subset of $\mathbb{R}^n$, $k \in \mathbb{N}$, and $1 \le p < \infty$. Define

$$\|u\|_{W^{k,p}(\Omega)} := \Big(\sum_{|\alpha| \le k} \int_\Omega |\partial^\alpha u(x)|^p \,\mathrm{d}x \Big)^{1/p}.$$

Then $\|\cdot\|_{W^{k,p}}$ is a norm, but $C^k(\bar{\Omega})$ is not complete under this norm. To see this, recall that $(C^k(\bar{\Omega}), \rho)$ is not complete as in Example 1.20.

The completion of $C^k(\bar{\Omega})$ under the $W^{k,p}$ norm is called the *Sobolev space*, denoted by $(W^{k,p}(\Omega), \|\cdot\|_{W^{k,p}})$. It can be shown that $C^k(\bar{\Omega})$ is dense in $(W^{k,p}(\Omega), \|\cdot\|_{W^{k,p}})$. If $k = 2$, then $(W^{2,p}(\Omega), \|\cdot\|_{W^{2,p}})$ is called the *Hilbert space*, which is denoted for short by $(H^p(\Omega), \|\cdot\|_{H^p})$.

Suppose $\|\cdot\|_1$ and $\|\cdot\|_2$ are two norms in X. We are often interested in the convergence properties of sequence rather than the actual value of distance. Then we can treat the two norms equivalently if any sequence that converges in the sense of $\|\cdot\|_1$ is also convergent in the sense of $\|\cdot\|_2$ and vice versa.

Definition 2.16 (Equivalent norms)**.** Let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two norms in X. We say that $\|\cdot\|_2$ is *stronger* than $\|\cdot\|_1$ if $\|x_k\|_1 \to 0$ whenever $\|x_k\|_2 \to 0$. If $\|\cdot\|_1$ is also stronger than $\|\cdot\|_2$, then we say the two norms are *equivalent*.

Proposition 2.17. $\|\cdot\|_2$ *is stronger than* $\|\cdot\|_1$ *iff there exists a constant* $c > 0$ *such that*

$$\|x\|_1 \le c\|x\|_2, \quad \forall\, x \in X.$$

Proof. ($\Leftarrow$) Trivial. ($\Rightarrow$) If not, then for any $k \in \mathbb{N}$, there exists x_k such that $\|x_k\|_1 > k\|x_k\|_2$ (obviously $x_k \neq 0$). Let $y_k = x_k/\|x_k\|_1$, then $1 = \|y_k\|_1 > k\|y_k\|_2$. Hence $0 \leq \|y_k\|_2 < 1/k$ for all $k \in \mathbb{N}$. Therefore $\|y_k\|_2 \to 0$. However $\|y_k\|_1 = 1$ for all $k \in \mathbb{N}$, which contradicts to $\|\cdot\|_2$ being stronger than $\|\cdot\|_1$. $\square$

Corollary 2.18. *$\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent iff there exist constants $c_1, c_2 > 0$ such that*

$$c_1\|x\|_2 \leq \|x\|_1 \leq c_2\|x\|_2, \quad \forall\, x \in X.$$

If X is finite dimensional, say $\dim(X) = n$, then there exists a set of basis vectors, $\{e_1, \dots, e_n\} \subset X$, which are linearly independent. Moreover, for any $x \in X$, there exist unique coefficients $a_1, \dots, a_n \in \mathbb{R}$, such that

$$x = a_1e_1 + \cdots + a_ne_n.$$

Recall that two finite dimensional linear spaces X and Y are isomorphic iff $\dim(X) = \dim(Y) = n$. Now we want to exploit more connections between X and Y if they are also normed spaces.

For the fixed basis $\{e_1, \dots, e_n\}$ and any norm $\|\cdot\|$ on X, we consider the linear isomorphism $T : X \to \mathbb{R}^n$ as follows,

$$Tx = a := (a_1, \dots, a_n) \in \mathbb{R}^n$$

for every $x = a_1e_1 + \cdots + a_ne_n$. We want to establish the relation between $\|x\|$ and $|Tx| := |a|$. To this end, we consider the mapping $q : \mathbb{R}^n \to \mathbb{R}$ defined by

$$q(a) := \|T^{-1}a\| = \Big\|\sum_{i=1}^n a_ie_i\Big\|.$$

Then we can show that q is Lipschitz continuous:

$$\begin{aligned}|q(a) - q(b)| &= \Big|\|\sum_{i=1}^n a_ie_i\| - \|\sum_{i=1}^n b_ie_i\|\Big| \leq \|\sum_{i=1}^n (a_i - b_i)e_i\| \leq \sum_{i=1}^n |a_i - b_i|\|e_i\| \\ &\leq \Big(\sum_{i=1}^n |a_i - b_i|^2\Big)^{1/2}\Big(\sum_{i=1}^n \|e_i\|^2\Big)^{1/2} = L|a - b|,\end{aligned}$$

where $L := (\sum_{i=1}^n |a_i - b_i|^2)^{1/2}$ is a constant. The first equality above is by the definition of q, the first two inequalities are due to the triangle inequality, and the third inequality is due to the Cauchy-Schwarz inequality.

Furthermore, we can show that q is homogeneous: for any $a \neq 0$, we have

$$q(a) = \Big\|\sum_{i=1}^n a_ie_i\Big\| = |a| \cdot \Big\|\sum_{i=1}^n \frac{a_i}{|a|}e_i\Big\| = |a| \cdot q\left(\frac{a}{|a|}\right),$$

which implies that $q(a)/|a| = q(a/|a|)$ for all nonzero $a \in \mathbb{R}^n$.

Denote $S := \{a \in \mathbb{R}^n : |a| = 1\}$ as the unit sphere in $\mathbb{R}^n$. We know that S is compact. Hence $q : S \to \mathbb{R}$ attains maximum and minimum on S. Namely, there exist $c_1, c_2 \geq 0$ such that

$$c_1 \leq q(a) \leq c_2, \quad \forall\, a \in S.$$

We claim that $c_1 > 0$: if $c_1 = 0$, then there exists $a \in S$ such that $q(a) = 0$, i.e., $\| \sum_{i=1}^n a_i e_i \| = 0$, which implies that $\sum_{i=1}^n a_i e_i = 0$. As $\{e_1, \dots, e_n\}$ is a basis (and hence linearly independent), we know $a_i = 0$ for all i, but this contradicts to that $a \in S$. Hence we know for any nonzero $a \in \mathbb{R}^n$ (not necessarily on S), there is

$$c_1 \leq \frac{q(a)}{|a|} = q\Big(\frac{a}{|a|}\Big) \leq c_2.$$

This means that $c_1|a| \leq q(a) \leq c_2|a|$ for all $a \in \mathbb{R}^n$ (we did not consider the case $a = 0$ but it is obviously true as well). Therefore, by noting that $a = Tx$, we have

$$c_1|Tx| = c_1|a| \leq q(a) = \|T^{-1}a\| = \|T^{-1}Tx\| = \|x\| \leq c_2|Tx|.$$

We denote $\|x\|_T := |Tx|$, then $\|\cdot\|_T$ is another norm on X. Hence $\|\cdot\|$ and $\|\cdot\|_T$ are equivalent. Note that $\|\cdot\|_T$ depends on T and the basis only.

Theorem 2.19 (Norms on finite dimensional linear spaces are equivalent)**.** *Let $\|\cdot\|_1$ and $\|\cdot\|_2$ be norms on X, and $\dim(X) = n$, then there exist $c_1, c_2 > 0$ such that $c_1\|x\|_1 \leq \|x\|_2 \leq c_2\|x\|_1$ for all $x \in X$.*

Proof. Both norms are equivalent to $\|\cdot\|_T$, and hence they are equivalent to each other. □

Corollary 2.20. *A finite dimensional B^* space is a Banach space.*

Proof. Let the basis $\{e_1, \dots, e_n\}$ and T be defined as above. Let $\{x^{(k)}\}$ be a Cauchy sequence in X under some norm $\|\cdot\|$, then it is also Cauchy under $\|\cdot\|_T$ (since $\|\cdot\|_T$ is equivalent to $\|\cdot\|$). Note that

$$\|x^{(m)} - x^{(k)}\|_T = |Tx^{(m)} - Tx^{(k)}| = |a^{(m)} - a^{(k)}|,$$

where $a^{(k)} \in \mathbb{R}^n$ is the coefficients to represent $x^{(k)}$ using the given basis. Therefore $\{a^{(k)}\}$ is Cauchy in $\mathbb{R}^n$ and hence converges to some $a \in \mathbb{R}^n$. Let $x = T^{-1}a = \sum_i a_i e_i$, then $\|x^{(k)} - x\|_T = |Tx^{(k)} - Tx| = |a^{(k)} - a| \to 0$ as $k \to 0$. So $x^{(k)} \to x \in X$ where $x = \sum_i a_i x_i$. Hence X is complete. □

Corollary 2.21. *A finite dimensional linear subspace of a B^* space is a Banach space.*

Definition 2.22 (Sublinear functional)**.** Let X be a linear space, and $P : X \to \mathbb{R}$ satisfy

1. (Subadditive) $P(x + y) \leq P(x) + P(y)$ for all $x, y \in X$.
2. (Positive homogeneous) $P(\lambda x) = \lambda P(x)$ for all $\lambda > 0$ and $x \in X$.

Then P is called a *sublinear functional*.

Definition 2.23 (Semi-norm)**.** Let X be a linear space, and $P : X \to \mathbb{R}$ satisfy

1. (Nonnegative) $P(x) \geq 0$ for all $x \in X$.
2. (Subadditive) $P(x + y) \leq P(x) + P(y)$ for all $x, y \in X$.
3. (Homogeneous) $P(\lambda x) = |\lambda| P(x)$ for all $\lambda \in \mathbb{R}$ and $x \in X$.

Then P is called a *semi-norm*.

Example 2.24. The total variation norm is a semi-norm on $\mathrm{BV}([a, b])$.

Theorem 2.25. *Let P be a sublinear functional on a finite dimensional Banach space X. If $P(x) \geq 0$ for all $x \in X$ and $P(x) = 0$ iff $x = 0$, then there exist $c_1, c_2 > 0$ such that $c_1\|x\| \leq P(x) \leq c_2\|x\|$ for all $x \in X$.*

Proof. Let $\{e_1, \dots, e_n\}$ be a basis of X and $Tx = a$ where $x = \sum_i a_i e_i$ as before. Consider the $\|\cdot\|_T$ norm on x, then

$$\begin{aligned}
|P(x) - P(y)| \leq P(x - y) = P\Big(\sum_{i=1}^{n}(a_i - b_i)e_i\Big) &\leq \sum_{i=1}^{n} P((a_i - b_i)e_i) \\
= \sum_{i=1}^{n} |a_i - b_i| P(\mathrm{sign}(a_i - b_i)e_i) &\leq \sum_{i=1}^{n} |a_i - b_i|(P(e_i) + P(-e_i)) \\
\leq \Big(\sum_{i=1}^{n}(P(e_i) + P(-e_i)\Big)^{1/2} |a - b| &= L\|x - y\|_T
\end{aligned}$$

where the first two inequalities is due to the subadditivity of P, the second equality due to positive homogeneity, and the last equality due to $L := (\sum_{i=1}^{n} P(e_i) + P(-e_i))^{1/2}$ which is a constant. Hence P is Lipschitz continuous on X.

For $S := \{x \in X : \|x\|_T = 1\}$, we know there exist $c_1, c_2 \geq 0$ such that $c_1 \leq P(x) \leq c_2$ for all $x \in S$. Then we claim that $c_1 > 0$: if not, then there exists $x \in S$ such that $P(x) = 0$, which contradicts to that P being positive definite. Hence for any nonzero x, we know $c_1 \leq P(x/\|x\|_T) = P(x)/\|x\|_T \leq c_2$, which verifies the claimed inequalities. □

Given a set of functions $\phi_1, \dots, \phi_n$, how to approximate a given function f using a linear combination of $\{\phi_i\}$? For example, let $f : [0, 2\pi] \to \mathbb{R}$ be given and $\phi_i(x) = \cos(ix)$, then how to approximate f by $\sum_{i=1}^{n} a_i\phi_i$ in the sense of L^2 norm?

Specifically, given a B^* space $(X, \|\cdot\|)$ and $\{e_1, \dots, e_n\}$ (assume they are linearly independent) and $x \in X$, consider the problem

$$\min_{a \in \mathbb{R}^n} \Big\| x - \sum_{i=1}^{n} a_i e_i \Big\|$$

where $a = (a_1, \dots, a_n) \in \mathbb{R}^n$. Question: does such optimal a^* exist? Is it unique?

To answer these questions, we define $F : \mathbb{R}^n \to \mathbb{R}$ by

$$F(a) = \Big\| x - \sum_{i=1}^{n} a_i e_i \Big\|.$$

Then it is straightforward to verify that F is Lipschitz continuous:

$$|F(a) - F(b)| \le \Big\| \sum_{i=1}^{n} (a_i - b_i) e_i \Big\| \le L|a - b|,$$

where $L := (\sum_{i=1}^{n} \|e_i\|^2)^{1/2}$. By triangle inequality, we know

$$F(a) \ge \Big\| \sum_{i=1}^{n} a_i e_i \Big\| - \|x\|.$$

Define $P : \mathbb{R}^n \to \mathbb{R}$ by $P(a) := \| \sum_{i=1}^{n} a_i e_i \|$, then P is a norm on $\mathbb{R}^n$ (check yourself). Hence there exists $c_1 > 0$ such that $P(a) \ge c_1|a|$. From this we see that $F(a) \to \infty$ as $a \to \infty$ (such F is called *coercive*), so the infimum of F must be in a ball $\bar{B}(0; r)$ in $\mathbb{R}^n$ for some $r > 0$. Thus, the minimizer of F must be attained in $\bar{B}(0; r)$ since $\bar{B}(0; r)$ is compact and F is continuous. This conclusion is summarized in the following theorem.

Theorem 2.26. *Let X be a B^* space and $\{e_1, \dots, e_n\}$ be linearly independent. Given any $x \in X$, there exist $a \in \mathbb{R}^n$ such that $F(a) := \|x - \sum_{i=1}^{n} a_i e_i\|$ is minimized.*

Remark. Let $M := \operatorname{span}(\{e_1, \dots, e_n\})$, then M is a linear subspace. The approximation problem consider earlier can be rewritten as

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

The theorem above implies that a minimizer y^* can be obtained over a finite dimensional subspace M of the B* space X. We call the distance between x and the subspace M by

$$d(x, M) := \min_{y \in M} \|x - y\|.$$

Now we would like to study the uniqueness of the solution in the approximation problem above.

Definition 2.27. A B* space $(X, \|\cdot\|)$ is called *strictly convex* if for any $x, y \in X$ where $x \neq y$ and $\|x\| = \|y\| = 1$, there is $\|(1 - \lambda)x + \lambda y\| < 1$ for all $\lambda \in (0, 1)$.

Example 2.28. In $\mathbb{R}^n$, the norm $|\cdot|_2$ is strictly convex, but the norms $|\cdot|_1$ and $|\cdot|_\infty$ are not strictly convex (check yourself).

Theorem 2.29. *Let X be a strictly convex B^* space, $\{e_1, \dots, e_n\}$ be a set of linearly independent vectors. Then for any $x \in X$, there exists a unique $a \in \mathbb{R}^n$ such that $F(a) := \|x - \sum_{i=1}^{n} a_i e_i\|$ is minimized.*

Proof. We have proved the existence of a. If a and b are both minimizers of F but are different, then let $y = \sum_{i=1}^n a_i e_i$ and $z = \sum_{i=1}^n b_i e_i$, and denote $r := d(x, M)$. Hence $\|x - y\| = \|x - z\| = r$. Moreover, by strict convexity of X, we have

$$\frac{1}{r}\|x - ((1-\lambda)y + \lambda z)\| = \left\|(1-\lambda)\frac{x-y}{r} + \lambda\frac{x-z}{r}\right\| < 1.$$

That is, $(1-\lambda)y + \lambda z \in M$ but $\|x - ((1-\lambda)y + \lambda z)\| < r$, which contradicts to the definition of r. □

Example 2.30. $L^p(\Omega, \mu)$ is strictly convex if $p \in (1, \infty)$.

Proof. Recall that the Minkowski's inequality holds: $\|u + v\|_p \le \|u\|_p + \|v\|_p$; and the equality holds only if one of u and v is zero or $u = cv$ for some $c > 0$. Consider $u, v \in X$ where $\|u\| = \|v\| = 1$ but $u \ne v$. Then $(1-\lambda)u$ and λv cannot have a $c > 0$ such that $(1-\lambda)u = c\lambda v$ (otherwise $u = v$). Hence $\|(1-\lambda)u + \lambda v\| < (1-\lambda)\|u\| + \lambda\|v\| = 1$. □

Example 2.31. Let (M, d) be a compact metric space. The space $(C(M), \|\cdot\|)$ where $\|x\| := \max_{x\in M} \|u(x)\|$ is not strictly convex. Take $C([0,1])$ as an example: consider $x(t) \equiv 1$ and $y(t) = t$. Then $\|x\| = \|y\| = 1$, $x \ne y$. But $\|(x+y)/2\| < 1$.

Example 2.32. $L^1(\Omega, \mu)$ is not strictly convex. Take $L^1([0,1])$ as an example: consider $x(t) \equiv 1$ and $y(t) = 2t$, then $\|x\| = \|y\| = 1$ and $x \ne y$, but $\|(x+y)/2\| = 1$.

Theorem 2.33. *A B^* space X is finite dimensional iff the unit sphere $S := \{x \in X : \|x\| = 1\}$ is sequentially compact.*

Proof. ($\Rightarrow$) This is because X is homeomorphic to the unit sphere in $\mathbb{R}^n$ and hence is sequentially compact.

($\Leftarrow$) Assume not, then for any $x_1, \ldots, x_n \in S$, we know that their linear span $M_n := \operatorname{span}(x_1, \ldots, x_n)$ is a proper subspace of X. Hence there exists nonzero $y \notin M_n$ (since M_n is closed). Let $r := d(y, M_n) > 0$. Then there exists $x \in M_n$ such that $r = d(y, x) = d(y, M_n)$. Define $x_{n+1} := (y - x)/r$, then for any $i \in \{1, \ldots, n\}$, we have

$$\|x_{n+1} - x_i\| = \left\|\frac{y-x}{r} - x_i\right\| = \frac{1}{r}\|y - (x + rx_i)\| \ge 1,$$

where the inequality is due to that $x + rx_i \in X$ and the definition of r. Hence we obtain a sequence $\{x_1, x_2, \ldots,\}$ such that $\|x_n - x_k\| \ge 1$ for all $n \ne k$. Therefore S is not sequentially compact, which is a contradiction. □

Lemma 2.34 (Riesz)**.** *Let X be a B^* space and X_0 be a proper closed subspace (X_0 may be infinite dimensional). Then for any $\epsilon \in (0,1)$, there exists $y \in X$ such that $\|y\| = 1$ and $\|y - x\| \ge 1 - \epsilon$ for all $x \in X_0$.*

Proof. Let $z \in X \setminus X_0$. Then $r := \inf_{x \in X_0} \|z - x\| > 0$ as X_0 is closed. For any $\epsilon \in (0,1)$, there exist $\eta > 0$ (we will specify η later) and $x_0 \in X_0$ such that $r \leq \|z - x_0\| < r + \eta$. Let $y = (z - x_0)/\|z - x_0\|$, then

$$\|y - x\| = \left\| \frac{z - x_0}{\|z - x_0\|} - x \right\| = \frac{\|z - (x_0 - \|z + x_0\| x)\|}{\|z - x_0\|} \geq \frac{r}{r + \eta}.$$

Solving $r/(r + \eta) = 1 - \epsilon$ we obtain $\eta = r\epsilon/(1 - \epsilon)$, and hence we can choose $\eta \in (0, r\epsilon/(1 - \epsilon))$. □

2.2 Convex sets and fixed points

Definition 2.35 (Convex set)**.** Let X be a linear space. Then a subset E of X is called *convex* if $(1 - \lambda)x + \lambda y \in E$ for any $\lambda \in [0,1]$ and $x, y \in E$.

Remark. The interior and closure of a convex set are also convex.

Proposition 2.36. *Let $\{E_\alpha : \alpha \in \mathcal{A}\}$ be a (finite, countable or uncountable) family of convex sets, then $\cap_{\alpha \in \mathcal{A}} E_\alpha$ is also convex.*

Definition 2.37 (Convex hull)**.** Let X be a linear space and $E \subset X$. Consider the family of convex sets:

$$\mathcal{F} := \{F \subset X : E \subset F,\ F \text{ is convex}\}.$$

Then $\cap_{F \in \mathcal{F}} F$ is called the *convex hull of* E, denoted by $\mathrm{conv}(E)$ or $\mathrm{co}(E)$. In other words, $\mathrm{conv}(E)$ is the smallest convex set containing E.

Proposition 2.38. *Let X be a linear space, $E \subset X$. Then*

$$\mathrm{conv}(E) = \Big\{ \sum_{i=1}^n a_i x_i : \sum_{i=1}^n a_i = 1, a_i \geq 0, x_i \in E, n \in \mathbb{N} \Big\}$$

Proof. Denote the set on the right hand side by S. As S is convex, we know $E \subset S$ and hence $\mathrm{conv}(E) \subset S$. On the other hand, for any $s \in S$, we know $s \in F$ for any $F \in \mathcal{F}$. Hence $S \subset F$ and thus $S \subset \mathrm{conv}(E)$. Therefore $S = \mathrm{conv}(E)$. □

Remark. Similar to Definition 2.37, we define the smallest *closed* convex set containing E to be the *closed convex hull of* E, denoted by $\overline{\mathrm{conv}}(E)$.

We have $\overline{\mathrm{conv}}(E) = \overline{\mathrm{conv}(E)}$ for any E: ($\subset$) Note that $E \subset \mathrm{conv}(E) \subset \overline{\mathrm{conv}(E)}$, $\overline{\mathrm{conv}(E)}$ is closed (as a closure) and convex (as the closure of the convex set $\mathrm{conv}(E)$), we know $\overline{\mathrm{conv}}(E) \subset \overline{\mathrm{conv}(E)}$ since $\overline{\mathrm{conv}}(E)$ is the smallest closed convex set containing E. ($\supset$) Note that $\mathrm{conv}(E) \subset \overline{\mathrm{conv}}(E)$ and $\overline{\mathrm{conv}}(E)$ is closed, we know $\overline{\mathrm{conv}(E)} \subset \overline{\mathrm{conv}}(E)$ since the closure $\overline{\mathrm{conv}(E)}$ is the smallest closed set containing $\mathrm{conv}(E)$.

There is also $\mathrm{conv}(\bar{E}) \subset \overline{\mathrm{conv}(E)}$ for any E: as $E \subset \overline{\mathrm{conv}(E)}$, we know $\bar{E} \subset \overline{\mathrm{conv}(E)}$. Since $\overline{\mathrm{conv}(E)}$ is convex, we know $\mathrm{conv}(\bar{E}) \subset \overline{\mathrm{conv}(E)}$ because $\mathrm{conv}(\bar{E})$ is the smallest convex set containing E. Note that the converse may not be true: consider $E = \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \geq (1 + x_2^2)^{-1}\}$, then $\mathrm{conv}(\bar{E}) = \{(x_1, x_2) : x_2 > 0\}$ and $\overline{\mathrm{conv}(E)} = \{(x_1, x_2) : x_2 \geq 0\}$.

Definition 2.39 (Minkowski functional)**.** Let X be a linear space and C a convex subset of X containing 0. Define a function $P : X \to [0, \infty]$ by

$$P(x) = \inf \left\{ \lambda > 0 : \frac{x}{\lambda} \in C \right\}, \quad \forall x \in X.$$

Then P is called the *Minkowski functional* (or gauge) of C.

Proposition 2.40. *Let X be linear space and C a convex subset of X containing 0. Then the Minkowski functional of C satisfies the following properties:*

1. $P(x) \in [0, \infty]$ *and* $P(0) = 0$.
2. $P(\lambda x) = \lambda P(x)$ *for all* $\lambda > 0$ *and* $x \in X$.
3. $P(x + y) \leq P(x) + P(y)$ *for all* $x, y \in X$.

Proof. Items 1 and 2 are obvious. For Item 3, WLOG, assume $P(x), P(y) < \infty$. Then for any $\epsilon > 0$, let $\lambda_1, \lambda_2 > 0$ such that

$$P(x) \leq \lambda_1 < P(x) + \frac{\epsilon}{2}, \quad P(y) \leq \lambda_2 < P(y) + \frac{\epsilon}{2},$$

and $x/\lambda_1, y/\lambda_2 \in C$. Then we have

$$\frac{\lambda_1}{\lambda_1 + \lambda_2} \cdot \frac{x}{\lambda_1} + \frac{\lambda_2}{\lambda_1 + \lambda_2} \cdot \frac{y}{\lambda_2} = \frac{x + y}{\lambda_1 + \lambda_2} \in C.$$

Hence $P(x + y) \leq \lambda_1 + \lambda_2 \leq P(x) + P(y) + \epsilon$. As ϵ is arbitrary, we have $P(x + y) \leq P(x) + P(y)$. □

Definition 2.41. (Absorbing set) Let X be a linear space and C a convex subset of X containing 0. C is called *absorbing* if for any $x \in X$ there exists $\lambda > 0$ such that $x/\lambda \in C$.

Definition 2.42 (Symmetric set)**.** Let X be a linear space and C a convex subset of X containing 0. C is called *symmetric* if $-x \in C$ whenever $x \in C$.

Proposition 2.43. *Let X be a linear space and C a convex subset of X containing 0. Let P be the Minkowski functional. Then the following two statements hold.*

1. *C is absorbing iff $P(x) \in [0, \infty)$ for any $x \in X$.*
2. *C is symmetric iff $P(ax) = |a|P(x)$ for any $a \in \mathbb{R}$ and $x \in X$.*

We can obtain even stronger result if X is a B* space.

Proposition 2.44. *Let X be a linear space and C a closed convex subset of X containing 0. Let P be the Minkowski functional. Then the following three statements hold.*

1. $C = \{x \in X : P(x) \leq 1\}$.
2. *If C is bounded, then $P(x) = 0$ iff $x = 0$.*
3. *If $0 \in \text{int}(C)$, then C is absorbing and $P(x)$ is Lipschitz continuous.*

Proof. 1. Let S denote the set on the right hand side. If $x \in C$, then $P(x) \leq 1$, and therefore $x \in S$. Hence $C \subset S$. If $x \in S$, then $P(x) \leq 1$. Hence $x_k = x/(1+\frac{1}{k}) \in C$ for all $k \in \mathbb{N}$. So $x_k \to x$. Since C is closed, we know $x \in C$. Hence $S \subset C$.

2. It is clear that $P(0) = 0$ since $0 \in C$. If C is bounded, then there exists $r > 0$ such that $C \subset B(0;r)$, i.e., for all $x \in 0$, $\frac{r}{\|x\|}x \notin C$. Hence $0 < \|x\|/r \leq P(x)$. Hence $P(x) = 0$ implies $x = 0$.

3. As $0 \in \text{int}(C)$, there exists $\epsilon > 0$ such that $B(0;\epsilon) \in C$. Then for any $x \in X$, there is $\frac{\epsilon}{2}\frac{x}{\|x\|} \in B(0;\epsilon) \subset C$. Since $P(x) \leq \frac{2\|x\|}{\epsilon}$, we know

$$|P(x) - P(y)| \leq \max(P(x-y), P(y-x)) \leq \frac{2}{\epsilon}\|x-y\|$$

which means that P is Lipschitz continuous (and hence also uniformly continuous). $\square$

Corollary 2.45. *Suppose C is a compact convex set in $\mathbb{R}^n$. Then there exists $m \leq n$ such that C is homeomorphic to $B^m := \{x \in \mathbb{R}^m : |x| \leq 1\}$.*

Proof. 1. Let E be the smallest linear manifold containing C, and $\dim(E) = m$. Then there exist $e_1, \ldots, e_{m+1} \in C$ such that $e_1 - e_{m+1}, \ldots, e_m - e_{m+1}$ are linearly independent. Therefore

$$e_0 := \frac{1}{m+1}\sum_{i=1}^{m+1} e_i \in C,$$

since C is convex. Then $E - e_0$ is an m-dimensional linear subspace of $\mathbb{R}^n$. We claim that for any $y \in E$, there exist unique $a_1, \ldots, a_m$ such that $y = e_0 + \sum_{i=1}^m a_i(e_i - e_0)$. To see this, we only need to show that $e_1 - e_0, \ldots, e_m - e_0$ are linearly independent which implies that they form a basis of $E - e_0$: let $a_1, \ldots, a_m \in \mathbb{R}$ be such that

$$\begin{aligned}
0 &= a_1(e_1 - e_0) + \cdots + a_m(e_m - e_0) \\
&= \sum_{i=1}^m a_i e_i - \sum_{i=1}^m \frac{\sum_{i=1}^m a_i}{m+1} e_i - \frac{\sum_{i=1}^m a_i}{m+1} e_{m+1} \\
&= \sum_{i=1}^m \left(a_i - \frac{\sum_{i=1}^m a_i}{m+1}\right)(e_i - e_{m+1}).
\end{aligned}$$

Since $e_1 - e_{m+1}, \ldots, e_m - e_{m+1}$ are linearly independent, we know all coefficients are 0 and hence we can readily deduce that $a_i = 0$ for all i. Hence we can define a norm $\|y\| = |a|$ for every $y = e_0 + \sum_{i=1}^m a_i(e_i - e_0) \in E$.

2. We want to show that e_0 is an interior point of C relative to E, i.e., if $\|y\| = |a|$ is small enough (we will specify it later) then $y \in C$. To see this, we notice that

$$y = e_0 + \sum_{i=1}^m a_i(e_i - e_{m+1}) = \sum_{i=1}^m \left(a_i + \frac{1 - \sum_{i=1}^m a_i}{m+1}\right) e_i + \frac{1 - \sum_{i=1}^m a_i}{m+1} e_{m+1}.$$

Choose $\epsilon > 0$ such that $(1 - m\epsilon)/(m + 1) > \epsilon$, i.e., $\epsilon < 1/(2m + 1)$. Now as long as $|a_i| \le \epsilon$ for all i, we have

$$|a_i| \le \epsilon < \frac{1 - m\epsilon}{m + 1} \le \frac{1 - \sum_{i=1}^{m} a_i}{m + 1}.$$

In this case, all coefficients in the representation of y using $e_1, \ldots, e_{m+1}$ above are nonnegative and their sum is 1, which means $y \in \operatorname{conv}(e_1, \ldots, e_{m+1}) \subset C$. Thus e_0 is an interior point of C, namely 0 is an interior point of $C - e_0$.

3. Let P be the Minkowski functional of $C - e_0$, then P is nonnegative, positive homogeneous, and subadditive (since P is the Minkowski functional), $P(x) = 0$ iff $x = 0$ (since $C - e_0$ is bounded), and Lipschitz continuous (0 is interior point of $C - e_0$). By Theorem 2.25, there exist $c_1, c_2 > 0$ such that $c_1\|x\| \le P(x) \le c_2\|x\|$ for any $y = e_0 + x \in E$. For any $z \in B(0; 1)$, define $f : B(0; 1) \to C$ by

$$f(z) = \begin{cases} e_0 + \frac{\|z\|}{P(z)} z, & \text{if } z \ne 0, \\ e_0, & \text{if } z = 0. \end{cases}$$

Since $\frac{1}{c_2} \le \frac{\|z\|}{P(z)} \le \frac{1}{c_1}$ for all $z \ne 0$, we know f is continuous for all $z \ne 0$. Moreover, there is

$$z \to 0 \quad \Longrightarrow \quad \left\|\frac{\|z\|}{P(z)} z\right\| \le \left\|\frac{1}{c_1} z\right\| \to 0,$$

which implies that f is also continuous at $z = 0$. The inverse of f is given by

$$f^{-1}(y) = \begin{cases} \frac{P(x)}{\|x\|} x, & \text{if } x \ne 0, \\ 0, & \text{if } x = 0, \end{cases}$$

for any $y = e_0 + x \in E$. Therefore f is a homeomorphism. □

Recall a well-known fixed point theorem from topology.

Theorem 2.46 (Brouwer). *Let $\bar{B}(0; 1) \subset \mathbb{R}^n$ be the closed unit ball and $T : B \to B$ is continuous. Then there exists $x \in B$ such that $Tx = x$.*

Corollary 2.47. *Let $C \subset \mathbb{R}^n$ be compact convex and $T : C \to C$ be continuous, then T has a fixed point in C.*

Proof. There exists a homeomorphism $\phi : C \to B$, and therefore $\phi \circ T \circ \phi^{-1} : B \to B$ is continuous. Hence there exists $y \in B$ such that $\phi \circ T \circ \phi^{-1}(y) = y$, i.e., $T(x) = x$ for $x = \phi^{-1}(y)$. □

Now we consider to extend the Brouwer fixed point theorem to infinite dimensional case.

Theorem 2.48 (Schauder). *Let X be a Banach space and C a closed convex subset of X. Suppose $T : C \to C$ is continuous and $T(C)$ is sequentially precompact. Then T has a fixed point in C.*

Proof. 1. As $T(C)$ is sequentially precompact, we know it is totally bounded, i.e., for any $k \in \mathbb{N}$ there exists $(1/k)$-net $N_k = \{y_1, \dots, y_{n_k}\}$ such that

$$T(C) \subset \bigcup_{i=1}^{n_k} B\left(y_i; \frac{1}{k}\right).$$

Denote $C_k := \operatorname{conv}(N_k)$, which is closed and bounded, and $C_k \subset C$ since C is convex.

2. Define $I_k : T(C) \to C_k$ as follows. For any $y \in T(C)$ and any $i \in \{1, \dots, n_k\}$, let

$$m_i(y) = \begin{cases} 1 - k\|y - y_i\|, & \text{if } y \in B(y_i; 1/k), \\ 0, & \text{if } y \notin B(y_i; 1/k). \end{cases}$$

We know $m_i(y) \geq 0$ and not all 0 (as N_k is a $(1/k)$-net). Then set $I_k(y) = \sum_{i=1}^{n_k} \lambda_i y_i \in C_k$ where $\lambda_i = m_i(y)/\sum_{i=1}^{n} m_i(y)$ for $i = 1, \dots, n_k$. Moreover

$$\begin{aligned} \|I_k(y) - y\| &= \left\|\sum_{i=1}^{n_k} \lambda_i y_i - \left(\sum_{i=1}^{n_k} \lambda_i\right) y\right\| \leq \sum_{i=1}^{n_k} \lambda_i \|y_i - y\| \\ &= \sum_{\{i: m_i(y) > 0\}} \lambda_i \|y_i - y\| + \sum_{\{i: m_i(y) = 0\}} \lambda_i \|y_i - y\| < \frac{1}{k} + 0 = \frac{1}{k}. \end{aligned}$$

It is also easy to verify that I_k is continuous.

3. Define $T_k := I_k \circ T : C \to C_k$ which is continuous. We can restrict T_k to C_k. By Corollary 2.47, there exists $x_k \in C_k$ such that $T_k x_k = x_k$, i.e., x_k is a fixed point of T_k. As $T(C)$ is sequentially precompact, we know $\{T x_k\}$ has at least one Cauchy subsequence, say $\{T x_{k_j}\}$. Since $T(C) \subset C$ and C is closed, we know there exists $x \in C$ such that $T x_{k_j} \to x$ as $j \to \infty$. Now we observe that

$$\begin{aligned} \|x_k - x\| &= \|T_k x_k - x\| = \|T_k x_k - T x_k + T x_k - x\| \\ &\leq \|I_k T x_k - T x_k\| + \|T x_k - x\| \leq \frac{1}{k} + \|T x_k - x\|. \end{aligned}$$

Therefore $\|x_{k_j} - x\| \leq \frac{1}{k_j} + \|T x_{k_j} - x\| \to 0$ as $j \to \infty$, i.e., $x_{k_j} \to x$. By the continuity of T, we know $T x_{k_j} \to T x$. Combining with $T x_{k_j} \to x$, we claim that $Tx = x$. □

We now consider an application of the Schauder theorem.

Theorem 2.49 (Carathéodory existence theorem)**.** *Let $f(t, x)$ be continuous on $U := [t - h, t + h] \times [\xi - b, \xi + b]$, where $M = \max_{(t,x) \in U} |f(t, x)|$ and $h \leq b/M$. Then the ODE*

$$\begin{cases} x'(t) = f(t, x(t)), & t \in [-h, h] \\ x(0) = \xi \end{cases}$$

has solution on $[-h, h]$.

Proof. Consider the closed ball $\bar{B}(\xi;b)$ in $C([-h,h])$ (ξ is the constant function on $[-h,h]$) and the mapping

$$(Tx)(t) = \xi + \int_0^t f(s, x(s))\,\mathrm{d}s.$$

We first show $T : \bar{B}(\xi;b) \to \bar{B}(\xi;b)$:

$$\|(Tx)(t) - \xi\| = \left\|\int_0^t f(s,x(s))\,\mathrm{d}s\right\| = \max_{|t|\le h}\left|\int_0^t f(s,x(s))\,\mathrm{d}s\right| \le Mh \le b.$$

Note that $T(\bar{B}(\xi;b))$ is also continuous: by continuity of f, for any $\epsilon > 0$, there exists $\delta > 0$, such that for any $(t,\eta),(t',\eta') \in U$ and $|(t,\eta)-(t',\eta')| < \delta$, there is $|f(t,\eta) - f(t',\eta')| < \epsilon/h$. Therefore, for any $x, y \in \bar{B}(\xi;b)$ satisfying $\|x-y\| < \delta$, there is

$$\begin{aligned}\|Tx - Ty\| &= \left\|\int_0^t (f(s,x(s)) - f(s,y(s)))\,\mathrm{d}s\right\| \\ &\le \max_{|t|\le h}\int_0^t |f(s,x(s)) - f(s,y(s))|\,\mathrm{d}s \le h\frac{\epsilon}{h} = \epsilon,\end{aligned}$$

since $|(s,x(s)) - (s,y(s))| = |x(s) - y(s)| < \delta$.

We now show that $T(\bar{B}(\xi;b))$, the image of $\bar{B}(\xi;b)$ under T is sequentially precompact. By Theorem 1.45 (Arzelà-Ascoli), we only need to show that $T(\bar{B}(\xi;b))$ is uniformly bounded and equicontinous. To this end, for any $x \in \bar{B}(\xi;b)$, we notice that

$$\begin{aligned}|(Tx)(t)| &\le |\xi| + \int_0^t |f(s,x(s))|\,\mathrm{d}s \le |\xi| + Mh, \\ |(Tx)(t) - (Tx)(t')| &= \left|\int_t^{t'} f(s,x(s))\,\mathrm{d}s\right| \le M|t-t'|,\end{aligned}$$

which show that T is uniformly bounded and equicontinuous, respectively.

Finally, by Theorem 2.48 (Schauder), we know T has a fixed point x in $\bar{B}(\xi;b)$, which implies that the ODE has a solution x in $\bar{B}(\xi;b)$. □

2.3 Inner product space and Hilbert space

B* space have norms and thus have the concept of convergence. However, they do not have the concept of "angles" between vectors and thus no "orthogonality". We plan to introduce such concept into linear spaces. We temporarily allow complex values $\mathbb{K}$ which can be either $\mathbb{R}$ and $\mathbb{C}$.

Definition 2.50 (Sesquilinear functional)**.** Let X be a linear space. Then $a : X \times X \to \mathbb{K}$ is called *sesquilinear* if the following identities hold for all $x, x_1, x_2, y, y_1, y_2 \in X$ and $\alpha_1, \alpha_2 \in \mathbb{K}$:

1. $a(\alpha_1 x_1 + \alpha_2 x_2, y) = \alpha_1 a(x_1, y) + \alpha_2 a(x_2, y)$;

2. $a(x, \alpha_1 y_1 + \alpha_2 y_2) = \overline{\alpha_1} a(x, y_1) + \overline{\alpha_2} a(x, y_2)$.

We call $q(x) := a(x, x)$ the *quadratic form* induced by a. If $a : X \times X \to \mathbb{R}$, then a is called *bilinear.*

Proposition 2.51. *Let X be a linear space and a a sesquilinear functional on X, then $a(x, x) \in \mathbb{R}$ for all $x \in X$ iff $a(x, y) = \overline{a(x, y)}$ for all $x, y \in X$.*

Proof. ($\Leftarrow$) $a(x, x) = \overline{a(x, x)}$ implies that $a(x, x) \in \mathbb{R}$ for all x.

($\Rightarrow$)For any $x, y \in X$, we have $a(x + y, x + y) = \overline{a(x + y, x + y)}$. The left hand side is $a(x, x) + a(x, y) + a(y, x) + a(y, y)$, and the right hand side is $a(x, x) + \overline{a(x, y)} + \overline{a(y, x)} + a(y, y)$. Hence $a(y, x) + a(x, y) = \overline{a(y, x)} + \overline{a(x, y)}$. Replacing y by ιy (ι denotes the unit imaginary number, i.e., $\iota^2 = -1$), then $\iota a(y, x) - \iota a(x, y) = -\iota\overline{a(y, x)} + \iota\overline{a(x, y)}$. Multiplying ι and combining the two identities yield $a(x, y) = \overline{a(y, x)}$. □

Definition 2.52 (Inner product)**.** Let X be a linear space, then a sesquilinear functional $\langle \cdot, \cdot \rangle : X \times X \to \mathbb{K}$ is called an *inner porduct* if for all $x, y \in X$, there are

1. $\langle x, y \rangle = \overline{\langle y, x \rangle}$;
2. $\langle x, x \rangle \geq 0$; and $\langle x, x \rangle = 0$ iff $x = 0$.

Then $(X, \langle \cdot, \cdot \rangle)$ is called an *inner product space.*

Example 2.53. $(\mathbb{R}^n, \langle \cdot, \cdot \rangle)$ is an inner product where $\langle x, y \rangle := \sum_{i=1}^n x_i y_i$ for all $x, y \in \mathbb{R}^n$. $(\mathbb{C}^n, \langle \cdot, \cdot \rangle)$ is an inner product where $\langle x, y \rangle := \sum_{i=1}^n x_i \overline{y_i}$ for all $x, y \in \mathbb{C}^n$.

Example 2.54. $(l^2, \langle \cdot, \cdot \rangle)$ is an inner product where $\langle x, y \rangle := \sum_{i=1}^\infty x_i \overline{y_i}$ for all $x, y \in l^2$.

Example 2.55. $(L^2(\Omega), \langle \cdot, \cdot \rangle)$ is an inner product where $\langle u, v \rangle := \int_\Omega u(x)\overline{v(x)}\,\mathrm{d}x$ for all $u, v \in L^2(\Omega)$.

Example 2.56. $(C^k(\bar{\Omega}), \langle \cdot, \cdot \rangle)$ is an inner product where

$$\langle u, v \rangle := \sum_{|\alpha| \leq k} \int_\Omega \partial^\alpha u(x) \overline{\partial^\alpha v(x)}\,\mathrm{d}x$$

for all $u, v \in C^k(\bar{\Omega})$.

Theorem 2.57 (Cauchy-Schwarz inequality)**.** *Let $(X, \langle \cdot, \cdot \rangle)$ be an inner product space, and $\|x\| := \langle x, x \rangle^{1/2}$ be the norm induced by the inner product. Then $|\langle x, y \rangle| \leq \|x\|\|y\|$ for all $x, y \in X$. Moreover, the equality holds iff $x = 0$ or $y = 0$ or $x = \lambda y$ for some $\lambda \neq 0$ (> 0 if without the absolute value on the left).*

Theorem 2.58. *Let X be a linear space and a be a sesequilinear functional on X. Let q be the quadratic form induced by a. If q is positive definite, i.e., $q(x) \geq 0$ for all $x \in X$ and $q(x) = 0$ iff $x = 0$, then $|a(x, y)| \leq (q(x)q(y))^{1/2}$, and the equality holds iff x and y are linearly dependent, i.e., there exist $\lambda_1, \lambda_2 \in \mathbb{K}$ such that $\lambda_1 x + \lambda_2 y = 0$.*

Proof. WLOG, assume $y \neq 0$. For any $\lambda \in \mathbb{K}$, we know

$$q(x + \lambda y) = q(x) + \bar{\lambda} a(x, y) + \lambda a(y, x) + |\lambda|^2 q(y) \geq 0$$

Let $\lambda = -a(x, y)/q(y)$, then from the inequality above we obtain

$$q(x + \lambda y) = q(x) - \frac{|a(x, y)|^2}{q(y)} \geq 0,$$

which implies that $|a(x, y)|^2 \leq q(x)q(y)$.

If $x = -\lambda y$, then it is easy to check that the equality holds. On the other hand, if the equality holds, then $q(x + \lambda y) = 0$ which implies that $x = -\lambda y$. □

Proposition 2.59. *Let $(X, \langle \cdot, \cdot \rangle)$ be an inner product space. Define $\|x\| := \langle x, x \rangle^{1/2}$ for all $x \in X$. Then $(X, \| \cdot \|)$ is a B^* space.*

Proof. We just need to show that $\| \cdot \|$ defined above is a norm. The positive definiteness and symmetry are easy to show. For triangle inequality, we have

$$\begin{aligned} \|x + y\|^2 &= \langle x + y, x + y \rangle = \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle \\ &\leq \|x\|^2 + 2\|x\|\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2, \end{aligned}$$

where we used Cauchy-Schwarz inequality. □

Proposition 2.60. *Let X be an inner product space. Then $\langle \cdot, \cdot \rangle$ is continous with respect to $\| \cdot \|$, i.e., $x_k \to x$ and $y_k \to y$ in $\| \cdot \|$ then $\langle x_k, y_k \rangle \to \langle x, y \rangle$.*

Proof. As $x_k \to x$ and $y_k \to y$, we know the sequences and x, y are all bounded (say by L). Then we have

$$\begin{aligned} |\langle x_k, y_k \rangle - \langle x, y \rangle| &= |\langle x_k, y_k \rangle - \langle x_k, y_k \rangle + \langle x_k, y_k \rangle - \langle x, y \rangle| \\ &\leq |\langle x_k, y_k - y \rangle| + |\langle x, y_k \rangle| \leq L\|y_k - y\| + \|y\|\|x_k - x\| \to 0 \end{aligned}$$

as $k \to \infty$ which completes the proof. □

Proposition 2.61. *Let $(X, \langle \cdot, \cdot \rangle)$ be an inner product space. Then X is a strictly convex B^* space.*

Proof. For any $x, y \in X$, where $\|x\| = \|y\| = 1$ and $x \neq y$, $\lambda \in (0, 1)$, there is

$$\begin{aligned} \|\lambda x + (1 - \lambda)y\|^2 &= \lambda^2 \|x\|^2 + 2\lambda(1 - \lambda)\Re\langle x, y \rangle + (1 - \lambda)^2 \|y\|^2 \\ &< \lambda^2 \|x\|^2 + 2\lambda(1 - \lambda)\|x\|\|y\| + (1 - \lambda)^2 \|y\|^2 \\ &= \lambda^2 + 2\lambda(1 - \lambda) + (1 - \lambda) = 1 \end{aligned}$$

which completes the proof. □

We also want to know in what case a B* space has an inner product $\langle \cdot, \cdot \rangle$ such that $\|x\| = \langle x, x \rangle^{1/2}$ for all $x \in X$.

Proposition 2.62. *Let $(X, \|\cdot\|)$ be a B^* space. Then there exists an inner product $\langle\cdot,\cdot\rangle$ such that $\|x\| = \langle x, x\rangle^{1/2}$ for all $x \in X$ iff $\|\cdot\|$ satisfies*

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2), \quad \forall\, x, y \in X.$$

Proof. ($\Rightarrow$) We can check that

$$\begin{aligned}\|x+y\|^2 + \|x-y\|^2 &= \langle x+y, x+y\rangle + \langle x-y, x-y\rangle \\ &= \|x\|^2 + \langle y, x\rangle + \langle x, y\rangle + \|y\|^2 + \|x\|^2 - \langle y, x\rangle - \langle x, y\rangle + \|y\|^2 \\ &= 2(\|x\|^2 + \|y\|^2)\end{aligned}$$

($\Leftarrow$) We can define

$$\langle x, y\rangle := \begin{cases} \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2), & \text{if } \mathbb{K} = \mathbb{R}, \\ \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2 + \iota\|x+\iota y\|^2 - \iota\|x-\iota y\|^2), & \text{if } \mathbb{K} = \mathbb{C}. \end{cases}$$

Then one can verify that this is an inner product. □

Definition 2.63 (Hilbert space)**.** A complete inner product space is called *Hilbert space.*

Theorem 2.64 (Poincaré inequality)**.** *Let Ω be an open bounded set in $\mathbb{R}^n$. Denote*

$$C_0^k(\Omega) := \{u \in C^k(\bar{\Omega}) : u = 0 \text{ near } \partial\Omega\}.$$

Here by $u = 0$ near $\partial\Omega$ we meant that for any $x \in \partial\Omega$ there exists $\delta_x > 0$ such that $u(y) = 0$ for all $y \in B(x; \delta_x)$. Then for any $u \in C_0^k(\Omega)$, there is

$$\sum_{|\alpha|<k} \int_\Omega |\partial^\alpha u(x)|^2 \,\mathrm{d}x \le C \sum_{|\alpha|=k} \int_\Omega |\partial^\alpha u(x)|^2 \,\mathrm{d}x$$

where $C = C(\Omega, k)$ only depends on Ω and k.

Proof. Since Ω is bounded, we can enclose it by a cube $\Omega_1 := [0, a]^n$ for some $a > 0$ large enough. Then $u \in C^k(\bar{\Omega}_1)$ and $u = 0$ on $\partial\Omega_1$. For any $x \in \Omega_1$, we know

$$u(x) = \int_0^{x_1} \partial_{x_1} u(t, x_2, \ldots, x_n) \,\mathrm{d}t.$$

By Cauchy-Schwarz inequality, we know

$$\begin{aligned}|u(x)|^2 &\le \Big(\int_0^{x_1} 1 \,\mathrm{d}t\Big) \Big(\int_0^{x_1} |\partial_{x_1} u(t, x_2, \ldots, x_n)|^2 \,\mathrm{d}t\Big) \\ &\le a \int_0^a |\partial_{x_1} u(t, x_2, \ldots, x_n)|^2 \,\mathrm{d}t,\end{aligned}$$

which is independent of x_1. So taking the integral on Ω_1, we know

$$\int_\Omega |u(x)|^2 \,\mathrm{d}x \le a^2 \int_\Omega |\partial_{x_1} u(x)|^2 \,\mathrm{d}x \le a^2 \int_\Omega |\nabla u(x)|^2 \,\mathrm{d}x.$$

Applying this inductively to obtain the claimed inequality. □

Remark. If we define

$$\|u\|_k := \sum_{|\alpha|=k} \int_\Omega |\partial^\alpha u(x)|^2 \, \mathrm{d}x,$$

then $\|\cdot\|_k$ is equivalent to the standard norm $\|\cdot\|$ defined by

$$\|u\| := \sum_{|\alpha|\le k} \int_\Omega |\partial^\alpha u(x)|^2 \, \mathrm{d}x$$

Moreover, the completion of $C_0^k(\Omega)$ under $\|\cdot\|$ is $H_0^k(\Omega)$ which is a closed subspace of $H^k(\Omega)$.

Example 2.65. $H_0^k(\Omega)$ is a Hilbert space with $\langle\cdot,\cdot\rangle$ defined by

$$\langle u, v\rangle_k := \sum_{|\alpha|=k} \int_\Omega \partial^\alpha u(x)\overline{\partial^\alpha v(x)} \, \mathrm{d}x.$$

Remark. If $\partial\Omega$ is smooth (the outer normal $\vec{n}$ is a smooth function), then

$$u|_{\partial\Omega} = \frac{\partial u}{\partial \vec{n}}\Big|_{\partial\Omega} = \cdots = \Big(\frac{\partial}{\partial \vec{n}}\Big)^{k-1} u\Big|_{\partial\Omega} = 0.$$

Now we can define orthogonality.

Definition 2.66 (Orthogonal)**.** The angle between $x, y \in H$ is defined by

$$\theta(x, y) := \arccos \frac{\langle x, y\rangle}{\|x\|\|y\|}.$$

We call x and y *orthogonal* if $\langle x, y\rangle = 0$. If $M \subset X$ is nonempty and $\langle x, y\rangle = 0$ for all $y \in M$, we say x is "orthogonal to M", denoted by $x \perp M$. We denote the orthogonal complement of M by $M^\perp := \{x \in X : x \perp M\}$.

Proposition 2.67. *Let $(X, \langle\cdot,\cdot\rangle)$ be an inner product space and $M \subset X$ be nonempty. Then the following statements hold:*

1. *If $x \perp y_1$ and $x \perp y_2$, then $x \perp (\lambda_1 y_1 + \lambda_2 y_2)$ for all $\lambda_1, \lambda_2 \in \mathbb{K}$.*
2. *If $x = y + z$ and $y \perp z$ then $\|x\|^2 = \|y\|^2 + \|z\|^2$.*
3. *If $x \perp y_k$ for all $k \in \mathbb{N}$ and $y_k \to y$, then $x \perp y$.*
4. *If $x \perp M$, then $x \perp \operatorname{span}(M)$.*
5. *$M^\perp$ is a closed linear subspace of X.*

Proof. Item 1 is trivial. For item 2, we have

$$\|x\|^2 = \|y + z\|^2 = \|y\|^2 + \langle y, z\rangle + \langle z, y\rangle + \|z\|^2 = \|y\|^2 + \|z\|^2.$$

For item 3, note that $0 = \langle x, y_k\rangle \to \langle x, y\rangle$ and hence $\langle x, y\rangle = 0$ since $y_k \to y$. Item 4 is due to item 1. For item 5, for any $y \in M$, $\{x_k\} \subset M^\perp$, and $x_k \to x$, then $\langle x_k, y\rangle \to \langle x, y\rangle = 0$, which implies that $x \in M^\perp$, which means $M^\perp$ is closed. □

Definition 2.68. Let $(X, \langle\cdot,\cdot\rangle)$ be an inner product space, $S := \{e_\alpha : \alpha \in \mathcal{A}\}$ is called an *orthogonal set* if $e_\alpha = e_\beta$ for all $\alpha, \beta \in \mathcal{A}$ and $\alpha \neq \beta$. If in addition $\|e_\alpha\| = 1$ for all $\alpha \in \mathcal{A}$, then S is called an *orthonormal set*. If $S^\perp = \{0\}$, then S is called *complete*.

We now show that every inner product space as a complete orthonormal set. The proof requires the Zorn's lemma which is logically equivalent to the axiom of choice. Recall that a set is X called *partially ordered* if an *order relation*, denoted by $\leq$, is defined for some pairs $a, b \in X$, such that the order has transitivity, i.e., $a \leq b$ and $b \leq c$ imply $a \leq c$, and reflexivity, i.e., $a \leq a$ for all $a \in X$. A subset E of a partially ordered set X is called *totally ordered* if for every pair $x, y \in E$, either $x \leq y$ or $y \leq x$. Let $E \subset X$, then we call $u \in X$ an *upper bound* of E if $x \leq u$ for all $x \in E$. We call $m \in X$ a *maximal element of* X if $x \leq m$ for all $x \in X$.

Lemma 2.69 (Zorn)**.** *Let X be a partially ordered set. If every totally ordered set has an upper bound, then X has a maximal element.*

Proposition 2.70. *Let $(X, \langle\cdot,\cdot\rangle)$ be an inner product space and $X \neq \{0\}$, then X has a complete orthonormal set.*

Proof. Let $\mathcal{E} := \{E \subset X : E \text{ is an orthonormal set}\}$. Suppose $E_1 \subset E_2 \subset \cdots$, then E_i is upper bounded by $\cup_{i=1}^\infty E_i$ in $\mathcal{E}$. Hence there exists $S \in \mathcal{E}$ which is maximal in $\mathcal{E}$.

We claim that S is complete: if not, then there exists a nonzero $x_0 \in S^\perp$. Let $S_0 := \{x_0\} \cup S^\perp$, then S_0 is also an orthonormal set and S is a proper subset of S_0, but this contradicts to that S being maximal in $\mathcal{E}$. ☐

Definition 2.71. Let $(X, \langle\cdot,\cdot\rangle)$ be an inner product space, then an orthonormal set $S = \{e_\alpha : \alpha \in \mathcal{A}\}$ is called a *basis of* X if for any $x \in X$ there is

$$x = \sum_{\alpha\in\mathcal{A}} \langle x, e_\alpha\rangle e_\alpha,$$

where $\langle x, e_\alpha\rangle$ is called the *Fourier coefficient* of x with respect to e_α.

Theorem 2.72 (Bessel inequality)**.** *Let $(X, \langle\cdot,\cdot\rangle)$ be an inner product space. If $S = \{e_\alpha : \alpha \in \mathcal{A}\}$ is an orthonormal set of X, then for any $x \in X$, there is*

$$\sum_{\alpha\in\mathcal{A}} |\langle x, e_\alpha\rangle|^2 \leq \|x\|^2.$$

Proof. For any finite subset of A, say $e_1, e_2, \ldots, e_n$, we have

$$0 \leq \left\| x - \sum_{i=1}^n \langle x, e_i\rangle e_i \right\|^2 = \|x\|^2 - \sum_{i=1}^n |\langle x, e_i\rangle|^2.$$

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

Now consider $\mathcal{A}_k := \{\alpha \in \mathcal{A} : |\langle x, e_\alpha \rangle| \in (\frac{1}{k+1}, \frac{1}{k}]\}$ and $\mathcal{A}_0 := \{\alpha \in \mathcal{A} : |\langle x, e_\alpha \rangle| > 1\}$. Then every $\mathcal{A}_0, \mathcal{A}_1, \dots$ is finite: otherwise we can extract $n > k$ terms from $\mathcal{A}_k$ such that $\sum_{i=1}^n |\langle x, e_i \rangle|^2 > \|x\|^2$. Hence there are at most countably many α such that $\langle x, e_\alpha \rangle > 0$. From above we can show $\sum_{i=1}^\infty |\langle x, e_\alpha \rangle|^2 \le \|x\|^2$. □

Corollary 2.73. *Let $(X, \langle \cdot, \cdot \rangle)$ be a Hilbert space, $\{e_\alpha : \alpha \in \mathcal{A}\}$ be an orthonormal set. Then for any $x \in X$, there is $\sum_{\alpha \in \mathcal{A}} \langle x, e_\alpha \rangle e_\alpha \in X$ and*

$$\Big\| x - \sum_{\alpha \in \mathcal{A}} \langle x, e_\alpha \rangle e_\alpha \Big\|^2 = \|x\|^2 - \sum_{\alpha \in \mathcal{A}} |\langle x, e_\alpha \rangle|^2.$$

Proof. We have shown above that $\mathcal{A}_x := \{\alpha \in \mathcal{A} : \langle x, e_\alpha \rangle \neq 0\}$ is at most countable. Let $\mathcal{A}_x = \{e_1, e_2, \dots\}$, then

$$\sum_{\alpha \in \mathcal{A}} \langle x, e_\alpha \rangle e_\alpha = \sum_{i=1}^\infty \langle x, e_i \rangle e_i.$$

By the Bessel inequality, we know the series $\sum_{\alpha \in \mathcal{A}} |\langle x, e_\alpha \rangle|^2$ is convergent. Hence

$$\sum_{i=k+1}^{k+p} |\langle x, e_i \rangle|^2 = \Big\| \sum_{i=k+1}^{k+p} \langle x, e_i \rangle e_i \Big\| \to 0$$

as $k \to \infty$ and for any $p \in \mathbb{N}$. Let $x_k := \sum_{i=1}^k \langle x, e_i \rangle e_i$, then $\{x_k\}$ is a Cauchy sequence in X, and

$$\sum_{\alpha \in \mathcal{A}} \langle x, e_\alpha \rangle e_\alpha = \sum_{i=1}^\infty \langle x, e_i \rangle e_i = \lim_{k \to \infty} x_k \in X$$

which verifies the first claim.

Furthermore, we know $(x - \sum_{i=1}^\infty \langle x, e_i \rangle e_i) \perp (\sum_{i=1}^\infty \langle x, e_i \rangle e_i)$, so

$$\Big\| x - \sum_{i=1}^\infty \langle x, e_i \rangle e_\alpha \Big\|^2 = \|x\|^2 - \sum_{i=1}^\infty |\langle x, e_i \rangle|^2.$$

This completes the proof. □

Theorem 2.74. *Let $(X, \langle \cdot, \cdot \rangle)$ be a Hilbert space, $S := \{e_\alpha : \alpha \in \mathcal{A}\}$ be an orthonormal set. Then the following statements are equivalent:*

1. *S is a basis.*
2. *S is complete.*
3. *The Parseval equality holds: $\|x\|^2 = \sum_{\alpha \in \mathcal{A}} |\langle x, e_\alpha \rangle|^2$ for all $x \in X$.*

Proof. (1) ⇒ (2). If S is not complete, then there exists $x \in S \setminus \{0\}$, such that $\langle x, e_\alpha \rangle = 0$ for all $\alpha \in \mathcal{A}$. Hence, as S is a basis, we have $x = \sum_{\alpha \in \mathcal{A}} \langle x, e_\alpha \rangle e_\alpha = 0$, contradiction.

(2) $\Rightarrow$ (3). If the Parseval equality does not hold, we have

$$\|x - \sum_{\alpha\in\mathcal{A}} \langle x, e_\alpha\rangle e_\alpha\|^2 = \|x\|^2 - \sum_{\alpha\in\mathcal{A}} |\langle x, e_\alpha\rangle|^2 > 0.$$

Let $y := x - \sum_{\alpha\in\mathcal{A}} \langle x, e_\alpha\rangle e_\alpha$, then $y \neq 0$ and $y \in S^\perp$, which contradicts to that S being complete.

(3) $\Rightarrow$ (1). By the Parseval equality, we know

$$\left\| x - \sum_{\alpha\in\mathcal{A}} \langle x, e_\alpha\rangle e_\alpha \right\|^2 = \|x\|^2 - \sum_{\alpha\in\mathcal{A}} |\langle x, e_\alpha\rangle|^2 = 0.$$

Therefore $x = \sum_{\alpha\in\mathcal{A}} \langle x, e_\alpha\rangle e_\alpha$ and hence S is a basis. ☐

Example 2.75. Consider $L^2([0, 2\pi])$, then $\{e_k(t) := e^{\iota kt}/\sqrt{2\pi} : k \in \mathbb{Z}\}$ is a basis. The Fourier coefficients are

$$\langle u, e_k\rangle = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} u(t) e^{-\iota kt}\, \mathrm{d}t, \quad \forall k \in \mathbb{Z}.$$

Example 2.76. Consider l^2, then $\{e_k := (0, \dots, 0, 1, 0, \dots) : k \in \mathbb{N}\}$ is a basis.

Now we consider orthogonalization in and homeomorphism between Hilbert spaces. Let $\{x_1, x_2, \dots\}$ be a linear independent set of H, then we can apply Gram-Schmidt orthogonalization: start from $y_1 = x_1$ and $e_1 = y_1/\|y_1\|$, we obtain e_k recursively by

$$y_k = x_k - \sum_{i=1}^{k-1} \langle x, e_i\rangle e_i, \qquad e_k = \frac{y_k}{\|y_k\|}$$

for all $k > 1$. Then it is easy to show that $\{e_k : k \in \mathbb{N}\}$ is an orthonormal set.

Definition 2.77 (Isometric inner product spaces)**.** Let $(X_1, \langle\cdot,\cdot\rangle_1)$ and $(X_2, \langle\cdot,\cdot\rangle_2)$ be two inner product spaces. If there exists an isomorphism $T : X_1 \to X_2$ such that

$$\langle Tx, Ty\rangle_2 = \langle x, y\rangle_1, \quad \forall\, x, y \in X_1,$$

then we say X_1 and X_2 are *isometric.*

Theorem 2.78. *Let $(X, \langle\cdot,\cdot\rangle)$ be a Hilbert space. Then the following statements hold:*

1. *X is separable iff any basis S of X is at most countable.*
2. *Let S be a basis of a separable Hilbert space X. If $|S| < \infty$, then X is isomorphic to $\mathbb{R}^n$. If $|S| = \infty$, then X is isomorphic to l^2.*

Proof. 1. ($\Rightarrow$) Let $\{x_k\}$ be a countable dense set of X. Then we can choose an at most countable subset $\{y_k\}$ (by screening $\{x_k\}$ in order, and skipping x_k if it can be represented by a linear combination of the previous $x_{k'}$'s for $k' < k$, resulting in the at most countable subset $\{y_k\}$. Let $\{e_k\}$ be the orthonormal set obtained by applying Gram-Schmidt process to $\{y_k\}$. Then

$$\text{span}(x_k) = \text{span}(y_k) = \text{span}(e_k),$$

and hence $\overline{\text{span}(e_k)} = \overline{\text{span}(x_k)} = X$. Therefore $\overline{\text{span}(e_k)}^{\perp} = \{0\}$, which implies that $\{e_k\}$ is complete, i.e., $\{e_k\}$ is a basis.

($\Leftarrow$) Consider the set $E := \{x = \sum_{k=1}^{\infty} \alpha_k e_k : \Re(\alpha_k), \Im(\alpha_k) \in \mathbb{Q}\}$ where $S = \{e_k : k \in \mathbb{N}\}$. Then E is countable and dense in X since S is a basis.

2. We only consider the case $|S| = \infty$. Suppose $S = \{e_k : k \in \mathbb{N}\}$. Consider the mapping $T : X \to l^2$ by $T(x) = (a_1, \dots, a_k, \dots)$ where $a_k = \langle x, e_k \rangle$ for all $k \in \mathbb{N}$. Then T is well defined (a is unique), linear, bijective. Moreover

$$\langle x, y \rangle = \left\langle \sum_{k=1}^{\infty} \langle x, e_k \rangle e_k, \sum_{k=1}^{\infty} \langle y, e_k \rangle e_k \right\rangle = \sum_{k=1}^{\infty} \langle x, e_k \rangle \overline{\langle y, e_k \rangle} = \langle T(x), T(y) \rangle_{l^2}$$

for all $x, y \in X$. Hence T is an isomorphism. $\square$

Remark. There exist non-separable Hilbert spaces. For example, let $l^2(\mathbb{R})$ denote the set of functions that are nonzero for at most countably many points in $\mathbb{R}$. Define the inner product $\langle f, g \rangle := \sum_{x \in \mathbb{R}} f(x) g(x)$. Then $(l^2(\mathbb{R}), \langle \cdot, \cdot \rangle)$ is a Hilbert space. Define $e_x \in l^2(\mathbb{R})$ for every $x \in \mathbb{R}$ such that $e_x(y) = 1$ if $y = x$ and 0 if $y \neq x$. Then $S := \{e_x : x \in \mathbb{R}\}$ is an uncountable basis of $l^2(\mathbb{R})$.

Now we revisit the approximation problem but this time in Hilbert space X. Given $x \in X$ and a closed convex set C of X, does there exist $y \in C$ such that $\|y - x\| = \inf_{z \in C} \|z - x\|$? If such y exists, is it unique?

Theorem 2.79. *Let C be a closed convex set of X, then there exists a unique $x_0 \in C$ such that it is the point in C closest to 0, i.e., $\|x_0\| = \inf_{z \in C} \|z\|$.*

Proof. We first show that such x_0 exists. Let $d = \inf_{z \in C} \|z\|$, then for any $k \in \mathbb{N}$, there exists $x_k \in C$, such that $d \le \|x_k\| < d + 1/k$. We can show that $\{x_k\}$ is Cauchy: for $j > k$, there is

$$\begin{aligned} \|x_k - x_j\|^2 &= 2(\|x_k\|^2 + \|x_j\|^2) - \|x_k + x_j\|^2 \\ &= 2(\|x_k\|^2 + \|x_j\|^2) - 4\|\frac{x_k + x_j}{2}\|^2 \\ &\le 2\left((d + \frac{1}{k})^2 + (d + \frac{1}{j})^2\right) - 4d^2 \\ &\le 4\left(d + \frac{1}{k}\right)^2 - 4d^2 = \frac{4}{k}\left(d + \frac{1}{k}\right) \to 0 \end{aligned}$$

as $k \to \infty$. As $\{x_k\}$ is Cauchy, there exists $x_0 \in C$ such that $x_k \to x_0$. Therefore $\|x_k\| \to \|x_0\|$, which implies that $\|x_0\| = d$.

Let x_0, x_0' be such that $\|x_0\| = \|x_0'\| = d$, then

$$\|x_0 - x_0'\|^2 = 2(\|x_0\|^2 + \|x_0'\|^2) - 4\left\|\frac{x_0 + x_0'}{2}\right\|^2 \le 4d^2 - 4d^2 = 0.$$

Hence $x_0 = x_0'$. $\square$

Corollary 2.80. *If C is a closed convex subset of a Hilbert space X, then for any $y \in X$, there exists a unique $x_0 \in C$ such that $\|y - x_0\| = \inf_{w \in C} \|y - w\|$.*

Proof. Consider $C - \{y\} := \{z = x - y : x \in C\}$ which is still convex. Apply the theorem above yields the claim. □

Remark. If M is a linear subspace, then for any $y \in X$, there exists a unique $x_0 \in M$ such that $\|y - x_0\| = \inf_{w \in M} \|y - w\|$.

Theorem 2.81. *Let C be a closed convex set of a Hilbert space X and $y \in X$. Then $x_0 \in X$ is such that $\|y - x_0\| = \inf_{w \in C} \|y - w\|$ (we call x_0 a projection of y onto C) iff $\Re(\langle y - x_0, x_0 - x\rangle) \geq 0$ for all $x \in C$.*

Proof. For any $x \in C$, consider $\phi_x : [0, 1] \to \mathbb{R}$ by

$$\phi_x(t) = \|y - (tx + (1 - t)x_0)\|^2.$$

Then x_0 is the projection of y onto C iff $\phi_x(t) \geq \phi_x(0)$ for any $x \in C$ and $t \in [0, 1]$.

Note that $\phi'_x(0) = 2\Re(\langle y - x_0, x - x_0\rangle)$ and $\phi''_x(t) \geq 0$ (hence ϕ_x is convex). Therefore $\phi_x(t) \geq \phi_x(0)$ for all $t \in [0, 1]$ and $x \in C$ iff $\phi'_x(0) \geq 0$ for all $x \in C$ iff $\Re(\langle y - x_0, x_0 - x\rangle) \geq 0$ for all $x \in C$. □

Corollary 2.82. *Let M be a linear manifold in the Hilbert space X. For any $y \in X \setminus M$, x_0 is the projection of y onto M iff $\langle y - x_0, x - x_0\rangle = 0$ for all $x \in M$.*

Proof. Since M is closed and convex, we know by Theorem 2.81 that $\Re(\langle y - x_0, x_0 - x\rangle) \geq 0$ for all $x \in M$. Let x' be such that $x' - x_0 = -(x - x_0)$, then we know $\Re(\langle y - x_0, x' - x_0\rangle) = -\Re(\langle y - x_0, x - x_0\rangle) \geq 0$. Hence $\Re(\langle y - x_0, x - x_0\rangle) = 0$. Similarly, by choosing x' such that $x' - x_0 = \pm\iota(x - x_0)$, we have $\Im(\langle y - x_0, x - x_0\rangle) = 0$ for all $x \in M$. Therefore $\langle y - x_0, x - x_0\rangle = 0$ for all $x \in M$. □

Corollary 2.83. *Let M be a closed linear subspace of a Hilbert space X, then for any $x \in X$, there exist unique $y \in M$ and $z \in M^\perp$ such that $x = y + z$.*

Proof. Let y be the projection of x onto M, then $z = x - y \in M^\perp$. So $x = y + z$, where $y \in M$ and $z \in M^\perp$. If there exists $y' \in M$ and $z' \in M^\perp$, then $x = y + z = y' + z'$ implies that $y - y' = z' - z \in M \cap M^\perp = \{0\}$. So $y = y'$ and $z = z'$. □

3 Linear Operator and Linear Functional

3.1 Linear operator

The concept of linear operators was motivated by a number of operations, such as linear transformations from $\mathbb{R}^m$ to $\mathbb{R}^n$, realized by matrices in $\mathbb{R}^{m\times n}$. Other linear mappings include $P(\partial_x) : C^\infty(\Omega) \to C^\infty(\Omega)$ where P is a polynomial and integral operator $I_K : C(\bar{\Omega}) \to C(\bar{\Omega})$ defined by

$$I_K(u) := \int_\Omega K(x,y)u(y)\,\mathrm{d}y.$$

In this section, we exploit the properties of linear operators defined as follows.

Definition 3.1 (Linear operator)**.** Let X and Y be two linear spaces and D a linear subspace of X. Then $T : D \to Y$ is called a *linear operator* if for any $x, x' \in D$ and $\alpha, \alpha' \in \mathbb{K}$, there is

$$T(\alpha x + \alpha' x') = \alpha T(x) + \alpha' T(x').$$

We call D the *domain of* T and $T(D)$ the *range of* T.

Example 3.2. Let $X = \mathbb{R}^n$ and $Y = \mathbb{R}^m$. Consider the mapping $T = [t_{ij}] \in \mathbb{R}^{m\times n}$. Then $T : x \mapsto Tx$ is a linear operator from X to Y.

Example 3.3. Let $\Omega \in \mathbb{R}^n$ be open and bounded, and $X = Y = C^\infty(\Omega)\cap C(\bar{\Omega})$, then $T := \sum_{|\alpha|\le k} \sigma_\alpha \partial^\alpha$, where $\sigma_\alpha \in C^\infty(\bar{\Omega})$ for all $|\alpha| \le k$, is a linear operator.

Remark. If $X = Y = L^2(\Omega)$, then $D = C^k(\Omega) \cap C(\bar{\Omega})$ is the domain of T, and $T : D \to Y$ is a linear operator.

Example 3.4. Let $X = L^1(\mathbb{C})$ and $T : X \to X$ be defined by

$$(Tu)(\zeta) := \int_{\mathbb{C}} e^{\iota\zeta z} u(z)\,\mathrm{d}z$$

for all $u \in X$ and $\zeta \in \mathbb{C}$. Then T is a linear operator.

Definition 3.5 (Linear functional)**.** A linear operator T from X to $\mathbb{K}$ is called a *linear functional*.

Example 3.6. Let $X \in C(\bar{\Omega})$ where $f : X \to \mathbb{R}$ is defined by

$$f(x) := \int_\Omega x(\xi)\,\mathrm{d}\xi \in \mathbb{R}.$$

Then f is a linear functional. Note that, however, $x \mapsto \int_\Omega x(\xi)^2\,\mathrm{d}\xi$ is not a linear functional.

Example 3.7. Let $X \in C(\bar{\Omega}) \cap C^\infty(\Omega)$, α a given multi-index, and $\xi_0 \in \Omega$ fixed. Then $f : X \to \mathbb{R}$ is defined by

$$f(u) := \partial^\alpha u(\xi_0), \quad \forall\, u \in C^\infty(\Omega)$$

is a linear functional on X.

Definition 3.8. Let X and Y be B* spaces, $D \subset X$ a linear subspace, then $T : D \to Y$ is called *continuous at* $x \in D$ if $Tx_k \to Tx$ whenever $x_k \to x$.

Proposition 3.9. *Let* $T : X \to Y$ *be a linear operator. Then* T *is continuous in* X *iff* T *is continuous at* 0.

Proof. We only need to show necessity. For any $x_k \to x$, we know $x_k - x \to 0$ and hence

$$T(x_k) - T(x) = T(x_k - x) \to 0.$$

which completes the proof. □

Definition 3.10 (Bounded operator)**.** Let X and Y be B* spaces, $D \subset X$ a linear subspace, then $T : D \to Y$ is called *bounded* if there exists $M > 0$ such that $\|Tx\| \leq M\|x\|$ for all $x \in X$.

Proposition 3.11. *Let* X *and* Y *be B* spaces,* $T : X \to Y$ *is a linear mapping. Then* T *is continuous iff* T *is bounded.*

Proof. ($\Leftarrow$) If T is bounded, then T is continuous at 0, and thus continuous in X.

($\Rightarrow$) If not, then for any $k \in \mathbb{N}$, there exists $x_k \in X$, such that $\|Tx_k\| > k\|x_k\|$. Let $y_k := x_k/(k\|x_k\|)$, then $\|y_k\| = 1/k \to 0$ but

$$\|Ty_k\| = \|Tx_k\|/(k\|x_k\|) > 1,$$

which contradicts to T being continuous. □

Definition 3.12. Let X and Y be B* spaces. The set of bounded linear operators is denoted by $L(X, Y)$. For any $T \in L(X, Y)$, the norm of T is defined by

$$\|T\| := \sup_{x \in X \setminus \{0\}} \frac{\|Tx\|}{\|x\|} = \sup_{\|x\|=1} \|Tx\|.$$

We denote $L(X, X)$ by $L(X)$ and $L(X, \mathbb{K})$ by X^* for short.

Theorem 3.13. *Let* X *be a B* space and* Y *a Banach space. Define summation and scalar multiplication in* $L(X, Y)$ *as follows:*

$$(\alpha_1 T_1 + \alpha_2 T_2)(x) = \alpha_1 T_1(x) + \alpha_2 T_2(x)$$

for all $x \in X$*,* $\alpha_1, \alpha_2 \in \mathbb{K}$*, and* $T_1, T_2 \in L(X, Y)$*. Then* $(L(X, Y), \|\cdot\|)$ *is a Banach space.*

Proof. 1. It is easy to show that $L(X, Y)$ is a linear space.

2. We need to show that $\|\cdot\|$ is a norm. It is easy to show that $\|\cdot\|$ is positive definite and homogeneous. To show the triangle inequality, we observe that

$$\begin{aligned}\|T_1 + T_2\| &= \sup_{\|x\|=1} \|(T_1 + T_2)(x)\| \leq \sup_{\|x\|=1} \|T_1(x)\| + \|T_2(x)\| \\ &\leq \sup_{\|x\|=1} \|T_1(x)\| + \sup_{\|x\|=1} \|T_2(x)\| = \|T_1\| + \|T_2\|.\end{aligned}$$

3. We need to show that $(L(X,Y), \|\cdot\|)$ is complete. Let $\{T_k\}$ be a Cauchy sequence in $L(X,Y)$, i.e., $\|T_{k+p} - T_k\| \to 0$ as $k \to 0$ for all $p \in \mathbb{N}$. In other words, for any $\epsilon > 0$, there exists $K = K(\epsilon) \in \mathbb{N}$, such that

$$\|T_{k+p}(x) - T_k(x)\| < \epsilon \|x\|, \quad \forall\, k \geq K,\ p \in \mathbb{N},\ x \in X. \tag{1}$$

Hence for each $x \in X$, $\{T_k(x)\}$ is a Cauchy sequence in Y. As Y is complete, there exists $y \in Y$ such that $T_k(x) \to y$ as $k \to \infty$.

Now consider the mapping $T : x \mapsto Tx := y$. Then it is easy to verify that T is linear. We also claim that T is bounded: for any $x \in X$, there is

$$\begin{aligned}\|Tx\| &\leq \|T_k x - Tx\| + \|T_k x\| = \lim_{p\to\infty} \|T_k x - T_{k+p} x\| + \|T_k x\| \\ &\leq \epsilon\|x\| + \|T_k\|\|x\| = (\|T_k\| + \epsilon)\|x\|.\end{aligned}$$

Hence $T \in L(X,Y)$. Moreover, for any $x \in X$, we also have

$$\|T_k x - Tx\| = \lim_{p\to\infty} \|T_k x - T_{k+p} x\| \leq \epsilon \|x\|$$

which implies $\|T_k - T\| \leq \epsilon$. Hence $T_k \to T$ in $L(X,Y)$. ☐

Example 3.14. If $T : X \to Y$ where X and Y are finite dimensional Banach spaces, then T is continuous.

Proof. In this case X is isomorphic to $\mathbb{R}^n$ and Y to $\mathbb{R}^m$, and T can be fully characterized by a matrix in $\mathbb{R}^{m\times n}$. Thus $|Tx| \leq |T|_F |x|$ where $|\cdot|_F$ is the Frobenius norm. ☐

Example 3.15. Let X be a Hilbert space and Y a nonzero closed linear subspace of X. Then for any $x \in X$, there exists a unique $y \in Y$ such that $z = x - y \in Y^\perp$. Define the projection operator $P : X \to Y$ by $P(x) = y$. Then P is linear, continuous, and $\|P\| = 1$.

Proof. It is easy to verify that P is linear. Moreover,

$$\|Px\|^2 = \|x - z\|^2 = \|x\|^2 - \|z\|^2 \leq \|x\|^2.$$

Hence $\|P\| \leq 1$, and thus P is continuous. For a nonzero $x \in Y$, there is $Px = x$ and hence $\|P\| = 1$. ☐

3.2 Riesz theorem and its applications

Let X be a Hilbert space. For any fixed $y \in X$, define $f_y(x) := \langle x, y\rangle$ for all $x \in X$. Then clearly $f_y : X \to \mathbb{K}$ is linear. Moreover, $\|f_y\| \leq \|y\|$ since $|f_y(x)| \leq \|y\|\|x\|$ for all $x \in X$. Taking $x = y$ also yields $f_y(y) = \|y\|^2$ and hence $\|f_y\| = \|y\|$. The converse of the statements above are also true, as shown by the following theorem.

Theorem 3.16 (Riesz representation theorem)**.** *Let X be a Hilbert space and $f \in X^*$, then there exists a unique $y_f \in X$ such that $f(x) = \langle x, y_f\rangle$ for all $x \in X$.*

Proof. If $f = 0$ then $y_f = 0$. If $f \neq 0$, then $Z := \{x \in X : f(x) = 0\}$ is a proper closed linear subspace of X (check yourself). Let $x_0 \in Z^{\perp}$ such that $\|x_0\| = 1$, then we can show that for any $x \in X$, $x = ax_0 + z$ for some $z \in Z$ and $a := f(x)/f(x_0)$:

$$f(z) = f(x - ax_0) = f(x) - af(x_0) = 0, \tag{2}$$

by the definition of a. Let $y_f := \overline{f(x_0)}x_0$, then

$$\langle x, y_f \rangle = \langle ax_0 + z, \overline{f(x_0)}x_0 \rangle = f(x)\|x_0\|^2 = f(x).$$

If there exist $y, y' \in X$ such that $f(x) = \langle x, y \rangle = \langle x, y' \rangle$ for all $x \in X$, then $\langle x, y - y' \rangle = 0$ for all $x \in X$. Taking $x = y - y'$ yields $\|y - y'\|^2 = 0$ which implies $y = y'$. □

Theorem 3.17. *Let X be a Hilbert space and $a(\cdot, \cdot) : X \times X \to \mathbb{R}$ a sesquilinear functional. If there exists $M > 0$ such that*

$$|a(x, y)| \leq M\|x\|\|y\|, \quad \forall\, x, y \in X$$

then there exists a unique $A \in L(X)$ such that

$$a(x, y) = \langle x, Ay \rangle, \quad \forall\, x, y \in X$$

and

$$\|A\| = \sup_{x,y \neq 0} \frac{a(x, y)}{\|x\|\|y\|} = \sup_{\|x\|=\|y\|=1} a(x, y).$$

Proof. For any fixed $y \in X$, $a(\cdot, y)$ is a linear functional on X. Hence by Theorem 3.16 (Riesz representation), there exists a unique $z_y \in X$ such that $a(x, y) = \langle x, z_y \rangle$ for all $x \in X$. Define $A : X \to X$ by $Ay = z_y$ (well defined due to the uniqueness of z_y).

We first show that A is linear: for any $\alpha_1, \alpha_2 \in \mathbb{K}$ and $y_1, y_2 \in X$, we have

$$\begin{aligned} \langle x, A(\alpha_1 y_1 + \alpha_2 y_2) \rangle &= a(x, \alpha_1 y_1 + \alpha_2 y_2) = \overline{\alpha_1} a(x, y_1) + \overline{\alpha_2} a(x, y_2) \\ &= \overline{\alpha_1} \langle x, Ay_1 \rangle + \overline{\alpha_2} \langle x, Ay_2 \rangle = \langle x, \alpha_1 Ay_1 + \alpha_2 Ay_2 \rangle, \end{aligned}$$

for any $x \in X$. Therefore $A(\alpha_1 y_1 + \alpha_2 y_2) = \alpha_1 Ay_1 + \alpha_2 Ay_2$, i.e., A is linear.

Since $|a(x, y)| \leq M\|x\|\|y\|$, we know

$$\|Ay\| = \sup_{x \neq 0} \frac{|\langle x, Ay \rangle|}{\|x\|} = \sup_{x \neq 0} \frac{|a(x, y)|}{\|x\|} \leq M\|y\|$$

where the second one due to the definition of A, and the inequality due to the bound on a. Hence $A \in L(X)$. □

Example 3.18 (Weak solution of PDE)**.** Let Ω be an open bounded set in $\mathbb{R}^n$, and $f \in L^2(\Omega)$ be given. The Dirichlet boundary value problem of the Poisson equation reads

$$\text{(PDE)} \qquad \begin{cases} -\Delta u = f, & \text{in } \Omega, \\ u = 0, & \text{on } \partial\Omega. \end{cases}$$

A function $u \in H_0^1(\Omega)$ is called a *weak solution* (or *generalized solution*) of (PDE) if

$$\text{(Weak form)} \qquad \int_\Omega \nabla u \cdot \nabla v \, \mathrm{d}x = \int_\Omega f v \, \mathrm{d}x, \quad \forall \, v \in H_0^1(\Omega).$$

The weak form is obtained by multiplying $v \in C_0^2(\Omega)$ on both sides of the PDE and applying integration by parts, noting the boundary condition of u and v and that $C_0^2(\Omega)$ is dense in $H_0^1(\Omega)$.

Note that a classical solution (also called strong solution) of the PDE requires u to be twice differentiable, which may be too strong to hold for certain f. However, if a classical solution does exist, then it must be a weak solution. Hence weak solution is a generalization of classical solution. The typical strategy in modern PDE theory is to first show existence and uniqueness of weak solutions of a given PDE (this can be carried out much more easily in the Sobolev space with proper compactness properties), and then study the regularity of the weak solutions and potentially show that a weak solution is in fact a strong solution.

Here we show an application of Theorem 3.16 (Riesz representation) to prove the existence and uniqueness of the PDE above, giving a hint on how powerful functional analysis is in modern PDE theory. Specifically, we show that for any given $f \in L^2(\Omega)$, the PDE has a unique weak solution: By Theorem 2.64 (Poincaré inequality), we know

$$\langle u, v \rangle_1 := \int_\Omega \nabla u(x) \cdot \nabla v(x) \, \mathrm{d}x$$

for any $u, v \in H_0^1(\Omega)$ is an inner product on $H_0^1(\Omega)$ (check yourself). Moreover, $H_0^1(\Omega)$ is complete under the norm $\| \cdot \|_1$ induced by this inner product. Define $T_f : H_0^1(\Omega) \to \mathbb{R}$ by $T_f(v) := \int_\Omega f v \, \mathrm{d}x$, then it can be shown that T_f is linear and bounded:

$$|T_f(v)| = \Big| \int_\Omega f v \Big| \le \Big(\int_\Omega |f|^2 \Big)^{1/2} \Big(\int_\Omega |v|^2 \Big)^{1/2} \le C \|f\|_{L^2} \|v\|_1,$$

where the constant C depends on Ω only (see Theorem 2.64). By Theorem 3.16 (Riesz representation), there exists a unique $u \in H_0^1(\Omega)$, such that $T_f(v) = \langle u, v \rangle_1$ for any $v \in H_0^1(\Omega)$, i.e., u satisfies the claimed (Weak form) above and is a weak solution of (PDE).

Remark. If the boundary condition is $u = g$ on $\partial\Omega$ for some $g \in C(\partial\Omega)$, then we can first try to find a function $u_0 \in C(\bar{\Omega}) \cap C^2(\Omega)$ such that $u_0 = g$ on $\partial\Omega$, and denote $f_0 = f - \Delta u_0$ and $v = u - u_0$. Then the problem can be converted to (PDE) above of v with f_0. There are results on the condition under which such u_0 exists.

Theorem 3.19. *Let C be a closed convex subset of $H_0^1(\Omega)$. If $f \in L^2(\Omega)$, then the following variational inequality (VI) has a unique solution $u^* \in C$:*

$$\text{(VI)} \qquad \int_\Omega \nabla u^* \cdot \nabla (v - u^*) \, \mathrm{d}x \ge \int_\Omega f (v - u^*) \, \mathrm{d}x, \quad \forall \, v \in C.$$

Proof. By Theorem 3.16 (Riesz representation), there exists a unique $w \in H_0^1(\Omega)$ such that

$$\int_\Omega \nabla w \cdot \nabla u \, \mathrm{d}x = \int_\Omega f u \, \mathrm{d}x, \quad \forall\, u \in H_0^1(\Omega).$$

Therefore (VI) is equivalent to

$$\text{(VI')} \quad \int_\Omega \nabla u^* \cdot \nabla(v - u^*) \, \mathrm{d}x \geq \int_\Omega \nabla w \cdot \nabla(v - u^*) \, \mathrm{d}x, \quad \forall\, v \in C,$$

which is $\langle u^* - w, v - u^* \rangle_1 \geq 0$ for all $v \in C$. Hence (VI') holds iff u^* is the projection of w onto C. From Corollary 2.82, such u^* exists and is unique. $\square$

Remark. This result can be applied to more general setting. Let Ω be an open bounded set in $\mathbb{R}^n$, $A : \bar{\Omega} \to \mathbb{R}^{n\times n}$ where $A(x) := [a_{ij}(x)]_{i,j}$ is a symmetric positive definite matrix for every $x \in \Omega$. Let $M := \max_{1\leq i,j\leq n} \max_{x\in\bar{\Omega}} |a_{ij}(x)|$. Moreover, there exists $\delta > 0$ such that

$$\sum_{i,j=1}^n a_{ij}(x)\xi_i\xi_j \geq \delta \sum_{i=1}^n |\xi_i|^2, \quad \forall\, x \in \Omega,\ \forall\, \xi = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n.$$

Namely, $A(x) \succeq \delta I$ for every $x \in \Omega$. Now consider a generalization of (PDE) where the Poisson equation is replaced with $-\nabla \cdot (A\nabla u) = f$ in Ω, i.e.,

$$-\sum_{i,j} \partial_{x_i}(a_{ij}(x)\partial_{x_j} u(x)) = f(x), \quad \forall\, x \in \Omega.$$

Then we can also show that a weak solution of (PDE) exists and is unique.

To see this, define a mapping $b : H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$ as follows: for any $u, v \in H_0^1(\Omega)$,

$$b(u, v) := \int_\Omega \nabla u(x) \cdot A(x)\nabla v(x) \, \mathrm{d}x = \sum_{i,j=1}^n \int_\Omega a_{ij}(x)\partial_{x_i} u(x)\partial_{x_j} v(x) \, \mathrm{d}x.$$

Then it is clear that b is bilinear and symmetric. Moreover,

$$b(u, u) \geq \delta \int_\Omega |\nabla u(x)|^2 \, \mathrm{d}x, \quad \forall\, u \in H_0^1(\Omega),$$

and $b(u, u) = 0$ iff $u = 0$ by Theorem 2.64 (Poincaré inequality). Hence b is an inner product on $H_0^1(\Omega)$. We denote it by $\langle u, v\rangle_b := b(u, v)$ for all $u, v \in H_0^1(\Omega)$. Let $\|u\|_b := \langle u, u\rangle_b^{1/2}$ be the norm induced by this inner product. Consider another inner product $\langle u, v\rangle_1 = \int_\Omega \nabla u \cdot \nabla v \, \mathrm{d}x$ on $H_0^1(\Omega)$ and its induced norm $\|\cdot\|_1$, then it is easy to verify that

$$\delta\|u\|_1^2 \leq \|u\|_b^2 \leq nM\|u\|_1^2,$$

which means that $\|\cdot\|_b$ and $\|\cdot\|_1$ are equivalent. Hence $(H_0^1(\Omega), \langle\cdot,\cdot\rangle_b)$ is complete and thus a Hilbert space.

For any $f \in L^2$, consider $T_f : H_0^1(\Omega) \to \mathbb{R}$ by $T_f(u) := \int_\Omega fu \, \mathrm{d}x$. Then it is easy to show that T_f is linear and bounded. Hence, by Theorem 3.16 (Riesz representation), there exists a unique $w_f \in H_0^1(\Omega)$ such that $T_f(v) = \langle v, w_f\rangle_b$ for all $v \in H_0^1(\Omega)$, which implies that w_f is the unique weak solution of (PDE).

We can also show that (VI) with the left-hand-side integral replaced with $\langle u^*, v - u^*\rangle_b$ has a unique solution: By Corollary 2.82, we know that the projection of w_f onto C, denoted by u^*, exists and is unique, and satisfies $\langle u^* - w_f, v - u^*\rangle_b \geq 0$ for all $v \in C$. This is equivalent to

$$\langle u^*, v - u^*\rangle_b \geq \langle w_f, v - u^*\rangle_b = T_f(v - u^*), \quad \forall\, v \in C.$$

Therefore, u^* is the unique solution of this generalized (VI).

3.3 Category and open mapping theorem

Definition 3.20 (Nowhere dense set)**.** Let (X, d) be a metric space. A subset E of X is called *nowhere dense* if $\operatorname{int}(\bar{E}) = \emptyset$, namely, $\bar{E}$ has no interior point.

Example 3.21. Finite set, $\mathbb{Z}$, Cantor set are nowhere dense sets in $\mathbb{R}$.

Proposition 3.22. *Let (X, d) be a metric space. Then a subset E is nowhere dense iff for any $B(x; r) \subset X$, there exists $B(x'; r') \subset B(x; r)$ such that $\bar{E} \cap \bar{B}(x'; r') = \emptyset$.*

Proof. ($\Rightarrow$) Since $\bar{E}$ has no interior point, $\bar{E}$ cannot contain any ball $B(x; r)$. That is, for any $B(x; r)$, there exists $x' \in B(x; r) \setminus \bar{E}$. Since $x \in \bar{E}^c$ which is open, we know there exists $r' > 0$ such that $B(x'; r') \subset B(x; r) \cap \bar{E}^c$. Thus $\bar{B}(x'; r') \cap \bar{E}^c = \emptyset$ (if necessary, choose $r' = r'/2$ to make this hold).

($\Leftarrow$) Assume E is not nowhere dense, then $\bar{E}$ has an interior point, and thus there exists an open ball $B(x; r) \subset \bar{E}$. Then any ball $B(x'; r')$ contained in $B(x; r)$ must satisfy $\bar{B}(x'; r') \cap B(x; r) \subset \bar{B}(x'; r') \cap \bar{E} \neq \emptyset$, a contradiction. ☐

Definition 3.23 (Category)**.** Let (X, d) be a metric space. Then E is said to be of *first category* if $E = \cup_{k=1}^\infty E_k$ where all E_k are nowhere dense. Otherwise, E is said to be of *second category.*

Remark. A set E is of first category (meager) if it can be written as a countable union of nowhere dense sets; otherwise it is of second category (fat).

Example 3.24. $\mathbb{Q}$ is of first category in $\mathbb{R}$. In fact, all countable sets are of first category, since they are countable union of singletons which are nowhere dense.

Theorem 3.25 (Baire)**.** *A complete metric space (X, d) is of second category.*

Proof. Assume not, then there exist nowhere dense sets E_k such that $X = \cup_{k=1}^\infty E_k$. For any $B(x_0; r_0)$, there exists $B(x_1; r_1) \subset B(x_0; r_0)$ (WLOG we assume $r_1 < 1$) such that $\bar{B}(x_1; r_1) \cap \bar{E}_1 = \emptyset$ since E_1 is nowhere dense. Then there exists $B(x_2; r_2) \subset B(x_1; r_1)$ (again assume $r_2 < 1/2$) such that $\bar{B}(x_2; r_2) \cap \bar{E}_2 = \emptyset$ since E_2 is nowhere dense, and so on. Thus we obtain a sequence of balls:

$$B(x_1; r_1) \supset B(x_2; r_2) \supset \cdots \supset B(x_k; r_k) \supset \cdots$$

where $r_k < 1/k$ and $\bar{B}(x_k; r_k) \cap (\cup_{i=1}^k \bar{E}_i) = \emptyset$ for all $k \in \mathbb{N}$.

Moreover, we know $x_{k+p} \in B(x_k; r_k)$ for all $p \in \mathbb{N}$, which implies that $d(x_{k+p}, x_k) < r_k < 1/k$. Hence $\{x_k\}$ is Cauchy. As X is complete, there exists $x \in X$ such that $x_k \to x$. Note that $d(x_k, x) = \lim_{p\to\infty} d(x_k, x_{k+p}) \leq r_k$, we know $x \in \bar{B}(x_k; r_k)$ for all $k \in \mathbb{N}$, and thus $x \notin (\cup_{k=1}^\infty E_k) = X$, which is a contradiction. □

Example 3.26. Weierstrass constructed a class of function which are continuous everywhere but nowhere differentiable, which is quite surprising and counter-intuitive. However, we can use Theorem 3.25 (Baire) to show that such functions are not rare. On the contrary, they dominate the space of continuous functions.

Theorem 3.27. *Let* $E := \{f \in C([0,1]) : f \text{ is nowhere differentiable}\}$. *Then* E^c *is of first category in* $(C([a,b]), d)$.

Proof. Note that E^c is the set of continuous functions which are differentiable at at least one point in $[0,1]$. For every $k \in \mathbb{N}$, we define

$$A_k := \{f \in C([0,1]) : \exists\, s \in (0,1),\ \text{s.t. } |f(s+h) - f(s)| \leq kh,\ \forall\, h \in (0, 1/k]\}.$$

(To simplify notation, we only consider h such that $h \leq 1/k$ and $s + h \leq 1$.) We can see that if f is differentiable at some s, then $f \in A_k$ for some $k \in \mathbb{N}$. Hence $E^c \subset \cup_{k=1}^\infty A_k$. Now we show that A_k is nowhere dense in $C([0,1])$. To this end, we show that A_k is closed (hence $A_k = \bar{A}_k$) but $\text{int}(A_k) = \emptyset$.

To show that A_k is closed, it suffices to show that A_k^c is open. Let $f \in A_k^c$, then for any $s \in [0,1]$, there exists $h_s \in (0, 1/k]$, such that $|f(s+h_s) - f(s)| > kh_s$. As f is continuous in $[0,1]$, we know there exists $\epsilon_s > 0$ (e.g., we can choose $\epsilon_s = (|f(s+h_s) - f(s)| - kh_s)/4 > 0$) and an open neighborhood J_s of s, such that for any $\sigma \in J_s$, there are $|f(s+h_s) - f(\sigma+h_s)|, |f(s) - f(\sigma)| < \epsilon_s$ and hence

$$\begin{aligned} |f(\sigma+h_s) - f(\sigma)| &\geq |f(s+h_s) - f(s)| - |f(s+h_s) - f(\sigma+h_s)| - |f(s) - f(\sigma)| \\ &> (kh_s + 4\epsilon_s) - \epsilon_s - \epsilon_s = kh_s + 2\epsilon_s, \end{aligned}$$

where the second inequality is due to the definition of ϵ_s. As $[0,1]$ is compact, there exists a finite subcover of $[0,1]$: $J_{s_1}, \ldots, J_{s_m}$, such that $[0,1] \subset \cup_{i=1}^m J_{s_k}$. Let $\epsilon := \min\{\epsilon_{s_1}, \ldots, \epsilon_{s_m}\} > 0$. For any $g \in C([0,1])$ such that $\|g - f\| < \epsilon$, we know for any $i = 1, \ldots, m$ and any $\sigma \in J_{s_i}$, there is

$$|g(\sigma + h_{s_i}) - g(\sigma)| \geq |f(\sigma + h_{s_i}) - f(\sigma)| - 2\epsilon > kh_{s_i},$$

which means that $B(f; \epsilon) \subset A_k^c$. Hence A_k^c is open and A_k is closed.

We now show $\text{int}(A_k) = \emptyset$. For any $f \in A_k$ and $\epsilon > 0$, there exists a polynomial p such that $\|f - p\| < \epsilon/2$. Since $p \in C^\infty([0,1])$, we know there exists $M > 0$ such that $|p'(x)| \leq M$ for all $x \in [0,1]$. Hence, for any $s \in [0,1]$ and $h \in (0, 1/k]$, there is $|p(s+h) - p(s)| \leq Mh$ by the mean value theorem. Now let $g \in C([0,1])$ be a piecewise linear function such that $\|g\| < \epsilon/2$ and the slope on each segment is larger than $M+k$. Then $\|p+g-f\| \leq \|p-f\| + \|g\| < \epsilon$, but

$$\begin{aligned}|(p(s+h) + g(s+h)) - (p(s) + g(s))| &\geq |g(s+h) - g(s)| - |p(s+h) - p(s)| \\ &> (M+k)h - Mh = kh.\end{aligned}$$

Therefore $p + g \in B(f;\epsilon)$ but $p + g \notin A_k$. Therefore f is not an interior point of A_k. As f is arbitrary, we know $\text{int}(A_k) = \emptyset$. ☐

Let X and Y be Banach spaces, $T \in L(X,Y)$. If T is bijective, we would like to know if T^{-1} is continuous. The answer is positive. In fact, this can be deduced from a more general result as follows.

Theorem 3.28 (Open mapping theorem)**.** *Let X and Y be Banach spaces and $T \in L(X,Y)$. If T is surjective, then T is an open mapping, namely, $T(W)$ is open in Y whenever W is open in X.*

Proof. We use $B(x;r)$ and $U(y;r)$ to denote the open balls in X and Y respectively.

1. We first show that T is an open mapping iff there exists $\delta > 0$ such that $U(0;\delta) \subset T(B(0;1))$, i.e., 0 is an interior point of $T(B(0;1))$.

($\Rightarrow$) Suppose T is an open mapping, then $T(B(0;1))$ is open in Y and $0 \in T(B(0;1))$. Hence there exists $\delta > 0$ such that $U(0;\delta) \subset T(B(0;1))$.

($\Leftarrow$) Suppose there exists $\delta > 0$ such that $U(0;\delta) \subset T(B(0;1))$. As T is linear, we know for any $r > 0$ and $x \in X$, there is $U(Tx; r\delta) \subset T(B(x;r))$ (by homogeneity of norm). Now for any $y \in T(W)$, where W is open, we know there exists $x \in W$ such that $y = Tx$. As W is open, there exists $B(x;r) \subset W$ (and hence $T(B(x;r)) \subset T(W)$), so $T(W)$ is open.

2. We now show that there exists $\delta > 0$, such that $U(0;3\delta) \subset \overline{T(B(0;1))}$. To this end, since T is surjective, we know $Y = T(X) = \cup_{k=1}^\infty T(B(0;k))$. Since Y is of second category, then there exists $k \in \mathbb{N}$ such that $T(B(0;k))$ is not nowhere dense, i.e., $\overline{T(B(0;k))}$ contains at least one interior point. Hence, there exists $y_0 \in Y$ and $r > 0$, such that $U(y_0;r) \subset \overline{T(B(0;k))}$. Since $\overline{T(B(0;k))}$ is symmetric, we know $U(-y_0;r) \subset \overline{T(B(0;k))}$. Thus,

$$U(0;r) \subset \frac{1}{2}U(y_0;r) + \frac{1}{2}U(-y_0;r) \subset \overline{T(B(0;k))}.$$

As T is homogeneous, we know by choosing $\delta = r/(3k)$ there is $U(0;3\delta) \subset \overline{T(B(0;1))}$.

3. We need to show that $U(0;\delta) \subset T(B(0;1))$ (without closure). For any $y_0 \in U(0;\delta) \subset \overline{T(B(0;1/3))}$, there exists $x_1 \in B(0;1/3)$ such that $\|y_0 - Tx_1\|_Y < \delta/3$ (since $y_0 \in T(B(0;1/3))$ or y_0 is a limit point of $T(B(0;1/3))$).

For $y_1 := y_0 - Tx_1 \in U(0; \delta/3) \subset \overline{T(B(0; 1/3^2))}$, there exists $x_2 \in B(0; 1/3^2)$ such that $\|y_1 - Tx_2\| < \delta/3^2$, and inductively, for $y_k := y_{k-1} - Tx_k \in U(0; \delta/3^k) \subset \overline{T(B(0; 1/3^{k+1}))}$, there exists $x_{k+1} \in B(0; 1/3^{k+1})$ such that $\|y_k - Tx_{k+1}\| < \delta/3^{k+1}$ for all $k \in \mathbb{N}$. Hence $\sum_{k=1}^{\infty} \|x_k\| < \sum_{k=1}^{\infty} 1/3^k = 1/2$, which means that $\sum_{k=1}^{\infty} x_k$ is absolutely convergent. Let $x_0 := \sum_{k=1}^{\infty} x_k$, then $x_0 \in B(0; 1)$. Denote $s_k = \sum_{i=1}^{k} x_i$, then we know that $s_k \to x_0$ and $y_k = y_0 - Ts_k \to 0$, i.e., $Ts_k \to y_0$. As T is continuous, we know $Tx_0 = y_0$. As $y_0 \in U(0; \delta)$ is arbitrary and $x_0 \in B(0; 1)$, we obtain $U(0; \delta) \subset T(B(0; 1))$. □

Theorem 3.29 (Banach)**.** *Let X and Y be Banach spaces and $T \in L(X, Y)$. If T is bijective, then T^{-1} is continuous.*

Proof. From the proof above, we know $U(0; 1) \subset T(B(0; 1/\delta))$. As T is bijective, $T^{-1}(U(0; 1)) \subset B(0; 1/\delta)$, that is, $\|T^{-1}y\| < 1/\delta$ for all $y \in Y$ where $\|y\| < 1$. For any $\epsilon > 0$ and nonzero $y \in Y$, we know $\|\frac{1}{1+\epsilon} \frac{y}{\|y\|}\| = \frac{1}{1+\epsilon} < 1$, and hence

$$\left\|T^{-1}\left(\frac{1}{1+\epsilon}\frac{y}{\|y\|}\right)\right\| = \frac{1}{1+\epsilon}\frac{1}{\|y\|}\|T^{-1}y\| < \frac{1}{\delta},$$

which means $\|T^{-1}y\| \leq \frac{1+\epsilon}{\delta}\|y\|$. Letting $\epsilon \to 0$, we have $\|T^{-1}y\| \leq \|y\|/\delta$. Hence T^{-1} is bounded and therefore continuous. □

Remark. In Theorem 3.28 (Open mapping), it is necessary that $T(X)$ is of second category (this is guaranteed by $Y = T(X)$ being complete). If this condition is missing, then the conclusion may not hold, as shown in the following example.

Example 3.30. Let $X = Y = C([0, 1])$ with the standard norm $\|\cdot\|$. Define $T : X \to Y$ such that $(Tx)(t) = \int_0^t x(s)\,\mathrm{d}s$. Then $T \in L(X, Y)$ and $T(X) = Y_0 := \{y \in C^1([0, 1]) : y(0) = 0\}$ which is not of second category in Y. In this case $T^{-1} = \frac{\mathrm{d}}{\mathrm{d}t}$, which is not continuous in Y_0. For example, let $x_k(t) = \sin(k\pi t)$, then $\|x_k\| = 1$ but $\|T^{-1}x_k\| = \|k\pi\cos(k\pi t)\| = k\pi \to \infty$ as $k \to \infty$. Hence, by letting $y_k = x_k/k$, we know $\|y_k\| \to 0$ but $\|T^{-1}y_k\| = \pi$ for all k. Hence T^{-1} is not continuous.

However, if we use $\|x\|_1 := \max_{0 \leq x \leq 1} \max(|x(t)|, |x'(t)|)$, then $(Y_0, \|\cdot\|_1)$ is a closed subspace of $(C^1([0, 1]), \|\cdot\|_1)$ and hence $(Y_0, \|\cdot\|_1)$ is a Banach space. In this case T^{-1} is continuous: for any $y \in Y_0$, $\|T^{-1}y\| = \|y'(t)\| \leq \|y\|_1$, which means T is bounded and hence continuous.

Remark. In the proof of Theorem 3.28 (Open mapping), we only needed continuity in the last step: as $s_k \to x_0$ and $Ts_k \to y_0$, by the continuity of T we can show that $Tx_0 = y_0$. But T does not need to be continuous for this to hold. See the following definition.

Definition 3.31 (Closed operator)**.** Let X and Y be metric spaces, then an operator $T : X \to Y$ is called *closed* if $x_k \to x$ and $Tx_k \to y$ imply $Tx = y$.

Example 3.32. Let $X = C^1([0,1])$ and $Y = C([0,1])$. Let $T = \frac{\mathrm{d}}{\mathrm{d}t}$ (both use $\|\cdot\|$, so X is not complete). Suppose $x_k \to x$ and $Tx_k = x_k' \to y$. Then

$$x_k(t) - x_k(0) = \int_0^t x_k'(s)\,\mathrm{d}s \to \int_0^t y(s)\,\mathrm{d}s.$$

(Note that $x_k' \to y$ in the sense of $\|\cdot\|$ means x_k' converges to y uniformly.) On the other hand, $x_k(t) - x_k(0) \to x(t) - x(0)$, so $x(t) - x(0) = \int_0^t y(s)\,\mathrm{d}s$. Hence $x \in C^1([0,1])$ and $x' = y$. Therefore T is a closed (and linear) operator, but T is not continuous.

Remark. If we have a closed, rather than continuous, linear operator T in Theorem 3.28 (Open mapping), then we could start from $D(T)$, the domain of T, rather than X and consider $T : D(T) \to Y$. Then in the last step of the proof of Theorem 3.28 (Open mapping), we also have $s_k \to x_0$ and $Ts_k \to y_0$. Note that $D(T)$ is a B* space (not necessarily Banach space, as seen in the example above), but we still have $x_0 \in D(T)$ provided that T is a closed operator.

Theorem 3.33. *Let X and Y be Banach spaces, $T : X \to Y$ a closed linear operator, $T(X)$ is of second category in Y. Then $T(X) = Y$, and for any $\epsilon > 0$, there exists $\delta = \delta(\epsilon) > 0$, such that $U(0;\delta) \subset T(B(0;\epsilon))$.*

Proof. Following the the proof of Theorem 3.28 (Open mapping), we know that there exists $\delta > 0$, such that $U(0;\delta) \subset T(B(0;1))$. So now we only need to show $T(X) = Y$. Obviously $0 \in T(B(0;1))$. If $y \neq 0$, then let $\delta_1 \in (0,\delta)$, we have

$$\frac{\delta_1}{\|y\|} y \in U(0;\delta) \subset T(B(0;1)).$$

Hence there exists $x \in B(0;1)$ such that $Tx = \delta_1 y/\|y\|$. Thus $y = T(\|y\|x/\delta_1) \in T(X)$. Therefore $T(X) = Y$. □

We want to study the relation of continuity and closedness of operators.

Theorem 3.34 (Bounded linear transformation, or BLT)**.** *Let X be a B^* space and Y a Banach space, $T : X \to Y$ be linear. Then T can be extended to T_1 on $\overline{D(T)}$ such that $T_1|_{D(T)} = T$ and $\|T_1\| = \|T\|$.*

Proof. For any $x \in \overline{D(T)}$, there exists a sequence x_k in $D(T)$ such that $x_k \to x$. As T is continuous on $D(T)$, we know there exists $M > 0$ such that $\|Tx\| \leq M\|x\|$ for all $x \in D(T)$. Hence $\|Tx_{k+p} - Tx_k\| \leq M\|x_{k+p} - x_k\| \to 0$ as $k \to \infty$ for any $p \in \mathbb{N}$. So $\{Tx_k\}$ is Cauchy in Y. As Y is complete, there exists $y \in Y$ such that $Tx_k \to y$. Note that y only depends on x, not $\{x_k\}$. Hence define $T_1 : x \mapsto y$. Obviously $T_1|_{D(T)} = T$. For any $\alpha, \alpha' \in \mathbb{R}$ and $x, x' \in X$, we choose $\{x_k\}, \{x_k'\}$ such that $x_k \to x$ and $x_k' \to x'$, and $Tx_k \to y$ and $Tx_k' \to y'$. Then

$$T_1(\alpha x + \alpha' x') = \lim_{k\to\infty} T(\alpha x_k + \alpha' x_k') = \alpha y + \alpha' y' = \alpha T_1 x + \alpha' T_1 x'.$$

Hence T_1 is linear. Moreover,

$$\|T_1 x\| = \|y\| = \lim_{k\to\infty} \|Tx_k\| \leq \|T\| \lim_{k\to\infty} \|x_k\| = \|T\|\|x\|.$$

Hence $T_1 \in L(X, Y)$. Moreover $\|T_1\| = \|T\|$ (we know $\|T_1\| \geq \|T\|$ by the definition of norm). $\square$

Corollary 3.35. *Let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two norms of a linear space X. If X is complete with respect to both norms, and $\|\cdot\|_1$ is stronger than $\|\cdot\|_2$, then $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent.*

Proof. Consider the identity mapping $I : (X, \|\cdot\|_1) \to (X, \|\cdot\|_2)$. As $\|\cdot\|_1$ is stronger than $\|\cdot\|_2$, we know $\|Ix\|_2 = \|x\|_2 \leq C\|x\|_1$. Hence I is continuous. As I is bijective, we know by Theorem 3.29 (Banach) I^{-1} is continuous and hence bounded. So there exists $M > 0$ such that $\|x\|_1 = \|I^{-1}x\|_1 \leq M\|x\|_2$ for all $x \in X$. Therefore $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent. $\square$

Remark. This also proves that, in a finite dimensional B* space (which is in fact a Banach space), all norms are equivalent.

Theorem 3.36 (Closed graph theorem)**.** *Let X and Y be Banach spaces, $T : D(T) \subset X \to Y$ be closed linear operator. If $D(T)$ is closed, then T is continuous.*

Proof. 1. Since $D(T)$ is closed, $(D(T), \|\cdot\|)$ is a Banach space. Now consider $T : D(T) \to Y$ and a new norm $\|\cdot\|_G$ on $D(T)$ as follows:

$$\|x\|_G = \|x\| + \|Tx\|, \quad \forall\, x \in D(T).$$

It is easy to verify that $\|\cdot\|_G$ is a norm.

2. We claim that $(D(T), \|\cdot\|_G)$ is a Banach space. To this end, we only need to show that $D(T)$ is complete under $\|\cdot\|_G$. Let $\{x_k\}$ be a Cauchy sequence in $D(T)$ with respect to $\|\cdot\|_G$. Then

$$\|x_k - x_j\|_G = \|x_k - x_j\| + \|Tx_k - Tx_j\| \to 0$$

as $k, j \to \infty$. So $\{x_k\}$ is a Cauchy sequence in $D(T)$ with respect to $\|\cdot\|$, and $\{Tx_k\}$ is a Cauchy sequence in Y. As $D(T)$ and Y are both complete (as $D(T)$ is closed), we know there exists $x^* \in D(T)$ and $y^* \in Y$, such that $x_k \to x$ in $D(T)$ and $Tx_k \to y$ in Y. Since T is closed, we know $Tx^* = y^*$. Hence $Tx_k \to Tx$. Therefore

$$\|x_k - x^*\|_G = \|x_k - x^*\| + \|Tx_k - Tx^*\| \to 0$$

as $k \to \infty$. This implies that $(D(T), \|\cdot\|_G)$ is complete.

3. By the definition of $\|\cdot\|_G$, we know $\|\cdot\|_G$ is stronger than $\|\cdot\|$ in X. By Corollary 3.35, they are equivalent, namely, there exists $M > 0$ such that $\|x\| \leq \|x\|_G \leq M\|x\|$ for all $x \in D(T)$. Hence T is bounded and therefore continuous. $\square$

Remark. $G(T) := \{(x, Tx) : x \in D(T)\}$ is called the *graph of* T. So $\|x\|_G$ is the norm of (x, Tx) in the product space $X \times Y$. Moreover, T is closed means that $G(T)$ is closed under this norm on the graph.

Theorem 3.37 (Uniform boundedness theorem)**.** *Let X be a Banach space and Y a B^* space. Define $\mathcal{F} \subset L(X, Y)$ such that*

$$\sup_{T \in \mathcal{F}} \|Tx\| < \infty, \quad \forall\, x \in X.$$

Then there exists $M > 0$ such that $\|T\| \le M$ for all $T \in \mathcal{F}$.

Proof. For any $x \in X$, define

$$\|x\|_{\mathcal{F}} := \|x\| + \sup_{T \in \mathcal{F}} \|Tx\|.$$

It is easy to show that $\|\cdot\|_{\mathcal{F}}$ is a norm.

We now show that X is complete under $\|\cdot\|_{\mathcal{F}}$. Let $\{x_k\}$ be a Cauchy sequence under $\|\cdot\|_{\mathcal{F}}$, then

$$\|x_k - x_j\|_{\mathcal{F}} = \|x_k - x_j\| + \sup_{T \in \mathcal{F}} \|Tx_k - Tx_j\| \to 0$$

as $k, j \to \infty$. Since X is complete, we know $\{x_k\}$ is Cauchy in X, and thus there exists $x \in X$ such that $x_k \to x$. Moreover, for any $\epsilon > 0$, there exists $K = K(\epsilon) \in \mathbb{N}$, such that $\sup_{T \in \mathcal{F}} \|Tx_k - Tx_j\| < \epsilon$ for all $k, j \ge K$. Therefore, for any $T \in \mathcal{F}$, we know $\|Tx_k - Tx_j\| < \epsilon$. Hence $\|Tx_k - Tx\| = \lim_{j \to \infty} \|Tx_k - Tx_j\| \le \epsilon$. Hence $\sup_{T \in \mathcal{F}} \|Tx_k - Tx\| \le \epsilon$. So

$$\|x_k - x^*\|_{\mathcal{F}} = \|x_k - x^*\| + \sup_{T \in \mathcal{F}} \|Tx_k - Tx^*\| \to 0$$

as $k \to \infty$. Therefore X is complete under $\|\cdot\|_{\mathcal{F}}$.

Since $(X, \|\cdot\|)$ and $(X, \|\cdot\|_{\mathcal{F}})$ are both complete and $\|\cdot\|_{\mathcal{F}}$ is stronger than $\|\cdot\|$, we know by Corollary 3.35 the two norms are equivalent. Therefore there exists $M > 0$ such that

$$\sup_{T \in \mathcal{F}} \|Tx\| \le M\|x\|.$$

Hence we know $\|T\| \le M$ for all $T \in \mathcal{F}$. □

Remark. The conclusion of Theorem 3.37 is different and stronger than the condition: the condition says that for any $x \in X$, $\sup_{T \in \mathcal{F}} \|Tx\| < \infty$, which means that there exists $M_x > 0$ dependent on x, such that $\|Tx\| \le M_x\|x\|$ for all $T \in \mathcal{F}$. On the other hand, the conclusion says that $\|T\| \le M$ for all $T \in \mathcal{F}$ where M is independent of x. Therefore, the condition says that $\mathcal{F}$ is bounded at every point $x \in X$, whereas the conclusion says that $\mathcal{F}$ is uniformly bounded on X. Note that another way of stating Theorem 3.37 is

$$\sup_{T \in \mathcal{F}} \|Tx\| = \infty \quad \Longrightarrow \quad \exists\, x_0 \in X, \text{ s.t. } \sup_{T \in \mathcal{F}} \|Tx_0\| = \infty.$$

Theorem 3.38 (Banach-Steinhaus)**.** *Let X be a Banach space and Y a B^* space, M is dense in X. Suppose $T, T_k \in L(X,Y)$ for all $k \in \mathbb{N}$. Then $\lim_{k\to\infty} T_k x = Tx$ for all $x \in X$ iff the following two statements hold:*

1. *There exists $C > 0$ such that $\|T_k\| \leq C$ for all $k \in \mathbb{N}$.*
2. *$\lim_{k\to\infty} T_k x = Tx$ for all $x \in M$.*

Proof. ($\Rightarrow$) We only need to show Item 1. For any $x \in X$, $\{T_k x\}$ is a bounded set and hence $\sup_k \|T_k x\| < \infty$. By Theorem 3.37 (Uniform boundedness), there exists $C > 0$ such that $\|T_k\| \leq C$.

($\Leftarrow$) Notice that M is dense in X, for any $x \in X$ and $\epsilon > 0$, we know there exists $y \in M$ such that $\|x - y\| \leq \epsilon/(2(\|T\| + C))$. Then

$$\begin{aligned}\|T_k x - Tx\| &= \|T_k x - T_k y\| + \|T_k y - Ty\| + \|Ty - Tx\| \\ &= C\|x - y\| + \|T_k y - Ty\| + \|T\|\|y - x\| \\ &= \frac{\epsilon}{2} + \|T_k y - Ty\|.\end{aligned}$$

Let $K = K(\epsilon) \in \mathbb{N}$ be such that $\|T_k y - Ty\| < \epsilon/2$ for all $k \geq K$. Then $\|T_k x - Tx\| < \epsilon/2 + \epsilon/2 = \epsilon$. Hence $\lim_{k\to\infty} T_k x = Tx$. ☐

Now we consider a few applications of the results in above.

Theorem 3.39 (Lax-Milgram)**.** *Let X be a Hilbert space and $a(\cdot,\cdot) : X \times X \to \mathbb{K}$ a sesquilinear functional which satisfies the following two conditions:*

1. *There exists $C > 0$ such that $|a(x, y)| \leq C\|x\|\|y\|$ for all $x, y \in X$.*
2. *There exists $\delta > 0$ such that $|a(x, x)| \geq \delta\|x\|^2$ for all $x \in X$.*

Then there exists a unique $A \in L(X)$ such that $a(x, y) = \langle x, Ay\rangle$ for all $x, y \in X$, A^{-1} exists and $\|A^{-1}\| \leq 1/\delta$.

Proof. By Theorem 3.17, we know such $A \in L(X)$ exists and is unique. We now show that A is bijective.

1. If $Ay_1 = Ay_2$, then $a(x, y) = \langle x, Ay_1\rangle = \langle x, Ay_2\rangle = a(x, y_2)$ for all $x \in X$. That is, $a(x, y_1 - y_2) = 0$ for all $x \in X$. Taking $x = y_1 - y_2$ yields $\delta\|y_1 - y_2\|^2 \leq a(y_1 - y_2, y_1 - y_2) = 0$, which implies that $y_1 = y_2$. Hence A is injective.

2. To show that A is surjective, i.e., $A(X) = X$, we first show that $A(X)$ is closed. To this end, let $w \in \overline{A(X)}$, then there exists a sequence $\{y_k\}$ in X such that $Ay_k \to w$. Now we know $\{Ay_k\}$ is Cauchy, so $\|Ay_{k+p} - Ay_k\| \to 0$ as $k \to \infty$ for all $p \in \mathbb{N}$. Hence

$$\begin{aligned}\delta\|y_{k+p} - y_k\|^2 &\leq |a(y_{k+p} - y_k, y_{k+p} - y_k)| = |\langle y_{k+p} - y_k, A(y_{k+p} - y_k)\rangle| \\ &\leq \|y_{k+p} - y_k\|\|Ay_{k+p} - Ay_k\|.\end{aligned}$$

So $\|y_{k+p} - y_k\| \leq \|Ay_{k+p} - Ay_k\|/\delta \to 0$ as $k \to 0$ and $p \in \mathbb{N}$.

Therefore, $\{y_k\}$ is Cauchy in X, which implies that there exists $y \in X$ such that $y_k \to y$. As A is continuous, we know $Ay_k \to Ay$. So $Ay = w$ and hence $w \in A(X)$. Therefore $A(X)$ is closed.

We then show that $A(X)^\perp = \{0\}$. Let $w \in A(X)^\perp$, then $\langle w, Ay\rangle = 0$ for all $y \in X$. Taking $y = w$, we have

$$\delta\|w\|^2 \leq a(w,w) = \langle w, Aw\rangle = 0.$$

So $w = 0$. Hence $A(X)^\perp = \{0\}$. Combining with $A(X)$ being closed, we know $A(X) = X$.

3. As $A : X \to X$ is bijective and X is complete, we know by Theorem 3.29 (Banach) that A^{-1} is continuous. Hence $A^{-1} \in L(X)$. For any $x \in X$, we know

$$\delta\|A^{-1}x\|^2 \leq |a(A^{-1}x, A^{-1}x)| = |\langle A^{-1}x, AA^{-1}x\rangle| \leq \|A^{-1}x\|\|x\|.$$

Hence $\|A^{-1}x\| \leq \|x\|/\delta$ for all $x \in X$, and thus $\|A^{-1}\| \leq 1/\delta$. □

Consider $L(X,Y)$ and suppose $T, T_k \in L(X,Y)$ are all bijective. Then for any $y \in Y$ there exist unique $x, x_k \in X$ such that $Tx = y$ and $T_k x_k = y$ for all $k \in \mathbb{N}$. In Numerical Analysis, T_k often corresponds to a discretization of the problem with mesh grid or step size h_k such that $h_k \to 0$, which is used to approximate the original T. Then we hope to have $x_k \to x$. We have the following definitions regarding $\{T_k\}$:

1. $\{T_k\}$ is said to be *convergent* if $T_k^{-1}y \to T^{-1}y$ for any $y \in Y$.
2. $\{T_k\}$ is said to be *consistent* if $T_k x \to Tx$ for any $x \in X$.
3. $\{T_k\}$ is said to be *stable* if there exists $C > 0$ such that $\|T_k^{-1}\| \leq C$ for all $k \in \mathbb{N}$.

Theorem 3.40 (Lax equivalence theorem)**.** *Let X, Y be Banach space and $T, T_k \in L(X,Y)$ be bijective for all $k \in \mathbb{N}$. Suppose $\{T_k\}$ is consistent. Then $\{T_k\}$ is convergent iff $\{T_k\}$ is stable.*

Proof. ($\Leftarrow$) For any $y \in X$, we have

$$\|T_k^{-1}y - T^{-1}y\| = \|T_k^{-1}y - x\| = \|T_k^{-1}Tx - T_k^{-1}T_k x\| \leq C\|Tx - T_k x\| \to 0$$

as $\{T_k\}$ is consistent.

($\Rightarrow$) For any $y \in Y$, $T_k^{-1}y \to T^{-1}y$. Therefore $\{T_k^{-1}y\}$ is bounded and thus $\sup_k \|T_k^{-1}y\| < \infty$. By Theorem 3.37 (Uniform boundedness), we know there exists $C > 0$ such that $\|T_k^{-1}\| \leq C$ for all $k \in \mathbb{N}$. □

4 Hahn-Banach Theorem

4.1 Hahn-Banach theorem

Let X be a linear space. Suppose we have a linear functional $f : X_0 \to \mathbb{R}$, where X_0 is a linear subspace of X. If we have a sublinear functional $p : X \to \mathbb{R}$, and f_0 is upper bounded by p on X_0, can we extend f_0 to a linear functional $f : X \to \mathbb{R}$, such that $f(x) = f_0(x)$ for all $x \in X_0$ and $f(x) \le p(x)$ for all $x \in X$? The answer is yes, as given by the Hahn-Banach Theorem.

Theorem 4.1 (Hahn-Banach theorem)**.** *Let X be a real linear space and $p : X \to \mathbb{R}$ a sublinear functional. Let X_0 be a linear subspace of X_0 and $f_0 : X_0 \to \mathbb{R}$ a linear functional such that $f_0(x) \le p(x)$ for all $x \in X_0$. Then there exists a linear functional $f : X \to \mathbb{R}$ such that*

1. *(Controlled by p) $f(x) \le p(x)$ for all $x \in X$.*
2. *(Extending f_0) $f(x) = f_0(x)$ for all $x \in X_0$. (Also denoted by $f|_{X_0} = f_0$).*

Proof. 1. We first consider a simple extension of X_0. Let $y_0 \in X \setminus X_0$, and consider

$$X_1 := X_0 + \{ay_0 : a \in \mathbb{R}\} = \{x + ay_0 : x \in X_0, a \in \mathbb{R}\}.$$

Suppose a linear functional $f : X_1 \to \mathbb{R}$ is a desired extension of f_0 on X_1, then we will have

$$f(x + ay_0) = f(x) + af(y_0) = f_0(x) + af(y_0).$$

Therefore, to determine f, it suffices to determine the value of $f(y_0)$. Next we will show what the range of this value is.

Since f needs to be controlled by p, we know

$$f(x + ay_0) \le p(x + ay_0), \quad \forall\, x \in X.$$

Taking $a = 1$ and $x = -z \in X_0$, we have

$$f(y_0) - f_0(z) = f(y_0) - f(z) = f(y_0 - z) \le p(y_0 - z), \quad \forall\, z \in X_0.$$

Taking $a = -1$ and $x = y \in X_0$, we have

$$-f(y_0) + f_0(y) = -f(y_0) + f(y) = f(y - y_0) \le p(y - y_0), \quad \forall\, y \in X_0.$$

Thus, combining the two inequalities above, we can see that $f(y_0)$ must satisfy

$$f_0(y) - p(-y_0 + y) \le f(y_0) \le f_0(z) + p(y_0 - z), \quad \forall\, y, z \in X_0.$$

In other words, $f(y_0)$ must satisfy

$$\sup_{y \in X_0} f_0(y) - p(-y_0 + y) \le f(y_0) \le \inf_{z \in X_0} f_0(z) + p(y_0 - z).$$

It remains to show LHS $\le$ RHS above: to this end, for any $y, z \in X_0$, there is

$$f_0(y) - f_0(z) = f_0(y - z) \le p(y - z) \le p(y - y_0) + p(y_0 - z)$$

which implies that $f_0(y_0) - p(-y_0 + y) \le f_0(z) + p(y_0 - z)$ for all $y, z \in X_0$. Therefore indeed LHS $\le$ RHS. However, note that $<$ may hold in which case the choice of $f(y_0)$ is not unique.

2. Now we consider extending f_0 to X. Denote

$$\mathcal{F} := \{(X_\Delta, f_\Delta) : X_0 \subset X_\Delta \subset X,\ f_\Delta|_{X_0} = f_0,\ f_\Delta \le p|_{X_\Delta}\}.$$

We also define an order "$<$" on $\mathcal{F}$ by inclusion:

$$(X_{\Delta_1}, f_{\Delta_1}) < (X_{\Delta_2}, f_{\Delta_2}) \quad \Longleftrightarrow \quad X_{\Delta_1} \subset X_{\Delta_1} \text{ and } f_{\Delta_2}|_{X_{\Delta_1}} = f_{\Delta_1}.$$

Then $\mathcal{F}$ is a partially ordered set. Moreover, for any totally ordered subset M

$$M := \{(X_{\Delta_k}, f_{\Delta_k}) : (X_{\Delta_k}, f_{\Delta_k}) < (X_{\Delta_{k+1}}, f_{\Delta_{k+1}}),\ \forall\, k \in \mathbb{N}\} \subset \mathcal{F},$$

we let

$$X_M := \bigcup_{(X_\Delta, f_\Delta) \in M} X_\Delta, \quad f_M|_{X_\Delta} = f_\Delta, \quad \forall\, (X_\Delta, f_\Delta) \in M.$$

Hence (X_M, f_M) is an upper bound of M. By Lemma 2.69 (Zorn), we know that there exists $(X_\Lambda, f_\Lambda) \in \mathcal{F}$ which is maximal in $\mathcal{F}$.

We claim that $X_\Lambda = X$: If not, then there exists $y_0 \in X \setminus X_\Lambda$ and we can follow the proof in the first part to construct $(\tilde{X}_\Lambda, \tilde{f}_\Lambda)$ such that $X_\Lambda \subset \tilde{X}_\Lambda$ and $\tilde{f}_\Lambda|_{X_\Lambda} = f_\Lambda$. Hence $(X_\Lambda, f_\Lambda) < (\tilde{X}_\Lambda, \tilde{f}_\Lambda)$, which contradicts to $(\tilde{X}_\Lambda, \tilde{f}_\Lambda)$ being maximal in $\mathcal{F}$. □

Corollary 4.2. *Let $(X, \|\cdot\|)$ be a B^* space, X_0 a linear subspace of X, and $f_0 \in X_0^*$. Then there exists $f \in X^*$ such that*

1. $f|_{X_0} = f_0$.

2. $\|f\| = \|f_0\|_0$ where $\|\cdot\|_0$ is the norm in X_0^*.

Proof. Let $p(x) \le \|f_0\|_0 \|x\|$, then p is sublinear. By Theorem 4.1 (Hahn-Banach), there exists $f \in X^*$ such that $f|_{X_0} = f$ and $f(x) \le p(x) = \|f_0\|_0 \|x\|$ for all $x \in X$. Hence $\|f\| \le \|f_0\|_0$. Since $f(x) = f_0(x)$ for all $x \in X_0$, we also know that

$$\|f_0\|_0 = \sup_{x \in X_0 \setminus \{0\}} \frac{\|f_0(x)\|}{\|x\|} = \sup_{x \in X_0 \setminus \{0\}} \frac{\|f(x)\|}{\|x\|} \le \sup_{x \in X \setminus \{0\}} \frac{\|f(x)\|}{\|x\|} = \|f\|.$$

Hence $\|f\| = \|f_0\|_0$. □

Corollary 4.3. *For any $x_1, x_2 \in X$, $x_1 \ne x_2$, there exists $f \in X^*$ such that $f(x_1) \ne f(x_2)$.*

Proof. Let $x_0 = x_1 - x_2 \neq 0$. Consider $X_0 := \{ax_0 : a \in \mathbb{R}\}$ and $f_0 : X_0 \to \mathbb{R}$ by $f_0(ax_0) = a\|x_0\| \in \mathbb{R}$. Then we can show $f_0 \in X_0^*$: for any $ax_0, ax_0' \in X_0$, there is

$$\begin{aligned} f_0(ax_0 + ax_0') &= f_0((a+a')x_0) = (a+a')\|x_0\| = a\|x_0\| + a'\|x_0\| \\ &= f_0(ax_0) + f_0(a'x_0). \end{aligned}$$

It is easy to verify that $f_0(x_0) = \|x_0\|$ and $\|f_0\|_0 = 1$. By Corollary 4.2, we know there exists $f \in X^*$ such that $f(x_0) = f_0(x_0) = \|x_0\|$ and $\|f\| = \|f_0\|_0 = 1$. Then we know that $f(x_0) = f(x_1 - x_2) = f(x_1) - f(x_2) = \|x_0\| \neq 0$. □

Corollary 4.4. *Let X be a B^* space, then for any nonzero $x_0 \in X$, there exists $f \in X^*$ such that $f(x_0) = \|x_0\|$ and $\|f\| = 1$.*

Remark. This corollary says that $x = 0$ iff $f(x) = 0$ for all $f \in X^*$.

Let X be a B* space and $f \in X^*$, then consider $M := \{x \in X : f(x) = 0\}$. For any $x_0 \in X$, the distance from x_0 to M is defined by

$$d(x_0, M) := \inf_{x \in M} \|x_0 - x\|.$$

Hence, there exists a sequence $\{x_k\}$ in M such that

$$d(x_0, M) \leq \|x_0 - x_k\| < d(x_0, M) + \frac{1}{k}.$$

Then we have

$$|f(x_0)| = |f(x_k - x_0)| \leq \|f\|\|x_k - x_0\| \leq \|f\|\Big(d(x_0, M) + \frac{1}{k}\Big)$$

for all $k \in \mathbb{N}$. Therefore $f(x_0) \leq \|f\| d(x_0, M)$.

Now consider the case where X is a B* space and M is a linear subspace of X. For any given $x_0 \in X \setminus M$, does there exist $f \in X^*$ such that $f|_M = 0$ and $|f(x_0)| = \|f\| d(x_0, M)$? The answer is yes, as shown in the following theorem.

Theorem 4.5. *Let X be a B^* space and M a linear subspace of X. If $x_0 \in X$ and $\delta := d(x_0, M) > 0$, then there exists $f \in X^*$ such that $f(x) = 0$ for all $x \in M$, $f(x_0) = \delta$, and $\|f\| = 1$.*

Proof. Consider the linear subspace $X_0 := \{x = x' + ax_0 : x' \in M, a \in \mathbb{R}\}$. For any $x = x' + ax_0 \in X_0$, define $f_0(x) := a\delta$. Then it is easy to verify that f_0 is linear. Moreover, if $a \neq 0$, then

$$|f_0(x)| = |a\delta| = |a| d(x_0, M) \leq |a| \Big\|\frac{x'}{a} + x_0\Big\| = \|x' + ax_0\| = \|x\|.$$

Hence $\|f_0\|_0 \leq 1$ and $f_0 \in X_0^*$. Let $x_k \in M$ be such that $\|x_k - x\| \to \delta$ as $k \to \infty$. Then we have

$$\delta = |f_0(x_0 - x_k)| \leq \|f_0\|_0 \|x_0 - x_k\| \to \|f_0\|_0 \delta.$$

Hence $1 \leq \|f_0\|_0$, and therefore $\|f_0\|_0 = 1$. By Corollary 4.2, there exists $f \in X^*$ such that $f|_{X_0} = f_0$ (which implies that $f(x) = 0$ for all $x \in M$ and $f(x_0) = \delta$) and $\|f\| = \|f_0\|_0 = 1$. □

Corollary 4.6. *Let $(X, \|\cdot\|)$ be a B^* space and M a subset of X. Suppose $x_0 \in X$ is nonzero. Then $x_0 \in \overline{\text{span}(M)}$ iff $f(x_0) = 0$ whenever $f \in X^*$ and $f|_M = 0$.*

Proof. ($\Rightarrow$) If $x_0 \in \overline{\text{span}(M)}$, then there exists a sequence $\{x_k\}$ in M such that $x_k \to x_0$. For any $f \in X^*$, we know $f(x_k) \to f(x_0) = 0$ as f is continuous.

($\Leftarrow$) If not, then $d(x_0, M) > 0$. By Theorem 4.5, there exists $f \in X^*$ such that $f|_M = 0$ and $f(x_0) = d(x_0, M) > 0$, which is a contradiction. ☐

Example 4.7. Let $M = \{x_1, x_2, \cdots\} \subset X$ and $x_0 \in X$. Then $x_0 \in \overline{\text{span}(M)}$ iff $f(x_0) = 0$ whenever $f \in X^*$ and $f(x_k) = 0$ for all $k \in \mathbb{N}$.

4.2 Geometric Hahn-Banach theorem

Now we consider the geometric meaning of the Hahn-Banach theorem.

Definition 4.8. Let X be a linear space. Then a proper linear subspace M of X is called *maximal* if $Y = X$ whenever M is a proper linear subspace of Y.

Proposition 4.9. *Let X be a linear space and M a proper linear subspace. Then M is maximal iff for any $x_0 \in X \setminus M$ there is $X = \{ax_0 : a \in \mathbb{R}\} \oplus M$.*

Proof. ($\Rightarrow$) If $X_0 := \{ax_0 : a \in \mathbb{R}\} \oplus M \neq X$, then there exist $x_1 \in X \setminus X_0$ such that $M \subsetneq X_0 \subsetneq \{ax_1 : a \in \mathbb{R}\} \oplus X_0$ which contradicts to M being maximal.

($\Leftarrow$) Let M_1 be a linear subspace and $M \subset M_1$, then there exists $x_0 \in M_1 \setminus M$ and $\{ax_0 : a \in \mathbb{R}\} \oplus M \subset M_1$. Therefore $M_1 = X$. By the definition of maximal linear subspace, we know M is a maximal linear subspace. ☐

Definition 4.10 (Hyperplane)**.** Let X be a linear space and M a proper linear subspace of X. Then $L = x_0 + M$ is called a *maximal linear manifold*, or a *hyperplane*, if M is maximal in X.

Maximal linear manifold is an extension of hyperplane in $\mathbb{R}^n$ to a general linear space, as shown by the following theorem.

Theorem 4.11. *Let X be a B^* space. Then L is a hyperplane in X iff there exists a nonzero linear functional $f : X \to \mathbb{R}$ and $r \in \mathbb{R}$ such that $L = H_f^r$. Moreover, L is closed iff f is continuous.*

Proof. 1. ($\Leftarrow$) Let $H_f^0 := \{x \in X : f(x) = 0\}$. Hence H_f^0 is a linear subspace of X. Then for any $x_1 \in X \setminus H_f^0$ and any $x \in X$, we can show

$$x \in \frac{f(x)}{f(x_1)} x_1 + H_f^0.$$

Therefore $X = \{ax_1 : a \in \mathbb{R}\} \oplus H_f^0$, which implies that H_f^0 is maximal. For any $x_0 \in X \setminus H_f^0$, set $r := f(x_0) \neq 0$ and $H_f^r := \{x \in X : f(x) = r\}$, then we have

$$x \in H_f^r \;\Leftrightarrow\; f(x - x_0) = f(x) - r = 0 \;\Leftrightarrow\; x - x_0 \in H_f^0 \;\Leftrightarrow\; x \in x_0 + H_f^0.$$

Hence $L = H_f^r$ is a hyperplane.

($\Rightarrow$) Suppose L is a hyperplane in X, then there exists a maximal linear subspace M and $x_0 \in X \setminus M$ such that $L = x_0 + M$ (or $L = M$ if L coincides with M). Then for any $x \in X$, we know x can be written as $x = ax_0 + y$ for some $y \in M$ and $a \in \mathbb{R}$. Define $f : X \to \mathbb{R}$ by $f(x) = f(ax_0 + y) := a$. Then it easy to show that f is linear, $H_f^0 = M$, and $f(x_0) = 1$. Hence $L = H_f^1$ (or $L = H_f^0$ if $L = M$).

2. We notice that L is closed iff H_f^0 is closed iff f is continuous. $\square$

Definition 4.12 (Separation of two sets by a hyperplane)**.** Two sets E and F of X are said to be *separated by a hyperplane* H_f^r if there exist a linear functional $f : X \to \mathbb{R}$ and $r \in \mathbb{R}$ such that $f(x) \leq r$ for all $x \in E$ and $f(x) \geq r$ for all $x \in F$. If the strict inequalities hold, then we say the separation is strict.

Theorem 4.13 (Separating a convex set and a point using hyperplane)**.** *Let E be a proper convex subset of a B* space X, and $0 \in \operatorname{int}(E)$. If $x_0 \notin E$, then there exists a hyperplane H_f^r that separates x_0 and E.*

Proof. Let $p : X \to \mathbb{R}$ be the Minkowski functional of E:

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

Define $X_0 := \{ax_0 : a \in \mathbb{R}\}$ and $f_0 : X_0 \to \mathbb{R}$ by $f_0(ax_0) := ap(x_0) \in \mathbb{R}$ for all $a \in \mathbb{R}$. Then it is easy to show that $f_0 \in X_0^*$, and

$$f_0(x) = f_0(ax_0) = ap(x_0) \leq p(ax_0) = p(x)$$

for any $x = ax_0 \in X_0$ since $p : X \to \mathbb{R}_+$ is positive homogeneous. By Theorem 4.1 (Hahn-Banach), there exists $f \in X^*$ such that $f|_{X_0} = f_0$ and $f(x) \leq p(x)$ for all $x \in X$. In particular, $f(x_0) = f_0(x_0) = p(x_0) \geq 1$. On the other hand, $f(x) \leq p(x) \leq 1$ for all $x \in E$. Hence H_f^1 separates x_0 and E. $\square$

Remark. As long as E has an interior point (not necessarily the origin), we can apply translation such that $0 \in \operatorname{int}(E)$. But E must have an interior point in the case of infinite dimensional B* space.

Remark. We can show that H_f^r is closed. We only need to show that f is continuous. Note that

$$|f(x)| \leq \max(p(x), -p(x)), \quad \forall\, x \in X,$$

which is continuous at 0 since $p(x)$, $-p(x)$ and max are continuous functions. Hence f is continuous at 0disjoint, and therefore continuous in X since f is linear.

Theorem 4.14 (Separating two disjoint convex sets by a hyperplane)**.** *Let E_1 and E_2 be two disjoint convex sets in B* space X, and* $\operatorname{int}(E_1) \neq \emptyset$*. Then there exists $s \in \mathbb{R}$ and $f \in X^*$ such that H_f^s separates E_1 and E_2. That is, $f(x) \leq s$ for all $x \in E_1$ and $f(x) \geq s$ for all $x \in E_2$.*

Proof. We first convert this to the problem of separating a convex set and a point above. Denote

$$E := E_1 - E_2 = \{x_1 - x_2 \in X : x_1 \in E_1,\ x_2 \in E_2\}.$$

It is easy to verify that E is convex and contains an interior point. Moreover $0 \notin E$: if not, then there exist $x_1 \in E_1$ and $x_2 \in E_2$ such that $0 = x_1 - x_2$, which implies that $x_1 = x_2 \in E_1 \cap E_2$, contradicting to $E_1 \cap E_2 = \emptyset$.

By Theorem 4.13, we know there exist $f \in X^*$ and $r \in \mathbb{R}$ such that H_f^r separates 0 and E, i.e., $f(x) \leq r$ for all $x \in E$ and $f(0) \geq r$. This implies $r \leq 0$ since $f(0) = 0$ for $f \in X^*$. Hence $f(x_1 - x_2) = f(x_1) - f(x_2) \leq r \leq 0$ for all $x_1 \in E_1$ and $x_2 \in E_2$. Hence there exists $s \in \mathbb{R}$ in between $\sup_{x_1 \in E_1} f(x_1)$ and $\inf_{x_2 \in E_2} f(x_2)$, and hence H_f^s separates E_1 and E_2. $\square$

Remark. The condition $E_1 \cap E_2 = \emptyset$ can be relaxed to $\text{int}(E_1) \cap E_2 \neq \emptyset$. Since $\text{int}(E_1)$ is a nonempty convex set, we know there exists H_f^s separates $\text{int}(E_1)$ and E_2, such that $f(x) \leq s$ for all $x \in \text{int}(E_1)$ and $f(x) \geq s$ for all $x \in E_2$. As f is continuous, we know $f(x) \leq s$ for all $x \in \overline{\text{int}(E_1)} = \bar{E}_1$. Hence H_f^s still separates E_1 and E_2.

Theorem 4.15 (Ascoli)**.** *Let X be a B^* space and E a closed convex set, then for any $x_0 \in X \setminus E$, there exists $f \in X^*$ and $s \in \mathbb{R}$ such that $f(x) < s < f(x_0)$ for all $x \in E$.*

Proof. Since E^c is open, there exists $\delta > 0$ such that $B(x_0; \delta) \subset E^c$. Since $B(x_0; \delta)$ is open and convex, we know there exists $f \in X^*$ such that $\sup_{x \in E} f(x) \leq \inf_{x \in B(x_0;\delta)} f(x)$.

We claim that $f(x_0) > \inf_{x \in B(x_0;\delta)} f(x)$: If not, then $f(x_0) = \inf_{x \in B(x_0;\delta)} f(x)$. Since f is nonzero, there exists nonzero $y \in X$ such that $f(y) > 0$. Then $x_0 - \frac{\delta}{2\|y\|} y \in B(x_0; \delta)$ but

$$f\left(x_0 - \frac{\delta}{2\|y\|} y\right) = f(x_0) - \frac{\delta}{2\|y\|} f(y) < f(x_0),$$

which is a contradiction. Hence we can choose s to be strictly between $\sup_{x \in E} f(x)$ and $f(x_0)$, then $f(x) \leq \sup_{x \in E} f(x) < s < f(x_0)$ for all $x \in E$. $\square$

Theorem 4.16 (Mazur)**.** *Let X be a B^* space and E a convex subset such that $\text{int}(E) \neq \emptyset$. Let F be a linear manifold in X. If $\text{int}(E) \cap F = \emptyset$, then there exists a hyperplane L containing F, such that E is on one side of L only.*

Proof. Suppose $F = x_0 + X_0$ where $x_0 \in X$ and X_0 is a linear subspace of X. Hence there exists a linear functional $f : X \to \mathbb{R}$ and $r \in \mathbb{R}$, such that the hyperplane H_f^r separates E and F: $f(x) \leq r$ for all $x \in E$ and $f(x) \geq r$ for all $x \in F$. Denote $s := f(x_0)$, then for any $x' \in X_0$, there is

$$r \leq f(x) = f(x_0 + x') = f(x_0) + f(x') = s + f(x').$$

Hence $f(x') \geq r - s$ for all $x' \in X_0$. Since $0 \in X_0'$, we know $0 \geq r - s$.

As X_0 is a linear subspace, we claim $f(x') = 0$ for all $x' \in X_0$: If not, say $f(x') > 0$ for some $x' \in X_0$, then $f(-ax') < r - s$ for $a > 0$ sufficiently large, which results in a contradiction. Hence $X_0 \subset H_f^0$ and $F \subset x_0 + H_f^0 = H_f^s$. Therefore $s \geq r \geq f(x)$ for all $x \in E$, which implies that the hyperplane H_f^s is the claimed one. □

Remark. In other words, there exists a nonzero $f \in X^*$ and $s \in \mathbb{R}$ such that $f(x) \leq s$ for all $x \in E$ and $f(x) = s$ for all $x \in F$.

4.3 Applications

Example 4.17 (Convex program in linear space)**.** Let X be a linear space and C a convex subset. Then $f : C \to \mathbb{R}$ is called a *convex functional* if for all $x, y \in C$ and $\lambda \in [0, 1]$, there is

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

We define the *epigraph* of f as follows:

$$\operatorname{epi}(f) := \{(x, t) \in C \times \mathbb{R} : f(x) \leq t\}.$$

Then f is convex iff $\operatorname{epi}(f)$ is convex in $C \times \mathbb{R}$.

Now consider a convex program (CP) as follows,

$$\text{(CP)} \qquad \min_{x \in C} f(x) \quad \text{s.t.} \quad g_1(x), \dots, g_m(x) \leq 0,$$

where $f, g_1, \dots, g_m : C \to \mathbb{R}$ are convex functionals. The goal of (CP) is to find $x_0 \in C$ that minimizes $f(x)$ subject to the inequality constraints $g_i(x) \leq 0$ for all $i = 1, \dots, m$.

We want to characterize the solution x_0 of (CP), i.e., the necessary condition for x_0 to be a solution of (CP). In particular, we want to show that there exist $\hat{\lambda}_1, \dots, \hat{\lambda}_m \in \mathbb{R}$, called the *Lagrange multipliers*, such that x_0 satisfies the *optimality condition*

$$\text{(OC)} \qquad f(x_0) + \sum_{i=1}^{m} \hat{\lambda}_i g_i(x_0) = \min_{x \in C} \left\{ f(x) + \sum_{i=1}^{m} \hat{\lambda}_i g_i(x) \right\}.$$

Note that we can show $\hat{\lambda}_1, \dots, \hat{\lambda}_m$ exist, but do not know their values. Nevertheless, the structure of (OC) provides many useful information that may lead us to find x_0.

To show the existence of $\hat{\lambda}_1, \dots, \hat{\lambda}_m$, we introduce an additional $\hat{\lambda}_0$, and rewrite (OC) as

$$\text{(OC')} \qquad \hat{\lambda}_0 f(x_0) + \sum_{i=1}^{m} \hat{\lambda}_i g_i(x_0) \leq \hat{\lambda}_0 f(x) + \sum_{i=1}^{m} \hat{\lambda}_i g_i(x), \quad \forall\, x \in C.$$

If $\hat{\lambda}_0 > 0$, then (OC) and (OC') are equivalent. To show the existence of $\hat{\lambda}_0, \ldots, \hat{\lambda}_m$ in (OC'), we consider two sets in $\mathbb{R}^{m+1}$:

$$E := \{(t_0, \ldots, t_m) \in \mathbb{R}^{m+1} : t_0 \le f(x_0),\ t_i \le 0,\ i = 1, \ldots, m\},$$
$$F := \{(t_0, \ldots, t_m) \in \mathbb{R}^{m+1} : \exists\, x \in C,\ \text{s.t. } t_0 \ge f(x),\ t_i \ge g_i(x),\ i = 1, \ldots, m\}.$$

It is straightforward to show that both E and F are convex sets in $\mathbb{R}^{m+1}$ (check yourself). Also note that $\text{int}(E) = \{(t_0, \ldots, t_m) \in \mathbb{R}^{m+1} : t_0 < f(x_0),\ t_i < 0,\ i = 1, \ldots, m\}$ is open and has interior points. We also claim that $\text{int}(E) \cap F = \emptyset$: if not, then there exist $t_0, \ldots, t_m$ and $x \in C$ such that $f(x) \le t_0 < f(x_0)$ and $g_i(x) \le t_i < 0$ for all i, which contradicts to x_0 being optimal.

Now we know, by Hahn-Banach theorem, there exists a separating hyperplane, determined by a nonzero $(\hat{\lambda}_0, \ldots, \hat{\lambda}_m) \in \mathbb{R}^{m+1}$. Note that

$$(f(x_0), g_1(x_0), \ldots, g_m(x_0)) \in E$$
$$(f(x) + \xi_0, g_1(x) + \xi_1, \ldots, g_m(x) + \xi_m) \in F$$

for all $x \in C$ and $\xi_i \ge 0$. Hence

$$\hat{\lambda}_0 f(x_0) + \sum_{i=1}^m \hat{\lambda}_i g_i(x_0) \le \hat{\lambda}_0 (f(x) + \xi_0) + \sum_{i=1}^m \hat{\lambda}_i (g_i(x) + \xi_i), \quad \forall\, x \in C.$$

It is clear that $\hat{\lambda}_0, \ldots, \hat{\lambda}_m \ge 0$: if $\hat{\lambda}_i < 0$, then letting $\xi_i \to \infty$ makes the RHS tend to $-\infty$, contradiction.

Note that $(f(x_0), 0, \cdots, 0) \in E$ and $(f(x_0), g_1(x_0), \ldots, g_m(x_0)) \in F$, and therefore

$$\hat{\lambda}_0 f(x_0) \le \hat{\lambda}_0 f(x_0) + \sum_{i=1}^m \hat{\lambda}_i g_i(x_0).$$

Hence $\sum_{i=1}^m \hat{\lambda}_i g_i(x_0) \ge 0$. But $\hat{\lambda}_i \ge 0$ and $g_i(x_0) \le 0$, which implies that $\hat{\lambda}_i g_i(x_0) = 0$ for all i.

We can also show that $\hat{\lambda}_0 > 0$ as long as there exists $\hat{x} \in C$ such that $g_i(\hat{x}) < 0$ for all i. To this end, assume $\hat{\lambda}_0 = 0$, then we know

$$0 = \hat{\lambda}_0 f(x_0) \le \hat{\lambda}_0 f(\hat{x}_0) + \sum_{i=1}^m \hat{\lambda}_i g_i(\hat{x}_0) = \sum_{i=1}^m \hat{\lambda}_i g_i(\hat{x}_0).$$

Since $(\hat{\lambda}_0, \hat{\lambda}_1, \ldots, \hat{\lambda}_m)$ is nonzero, we know $(\hat{\lambda}_1, \ldots, \hat{\lambda}_m)$ is nonzero and hence $\hat{\lambda}_i g_i(x_0) < 0$, which is a contradiction. This proves that $\hat{\lambda}_0 > 0$. This result is summarized in the following theorem.

Theorem 4.18 (Kuhn-Tucker)**.** *Let X be a linear space and C a convex subset of X. Suppose $f, g_1, \ldots, g_m : C \to \mathbb{R}$ are convex functionals, and there exists $\hat{x} \in C$ such that $g_1(\hat{x}), \ldots, g_m(\hat{x}) < 0$. If $x_0 \in C$ solves (CP), then there exist $\lambda_1, \ldots, \lambda_m \ge 0$ such that*

$$f(x_0) = \min_{x \in C} \Big\{ f(x) + \sum_{i=1}^m \lambda_i g_i(x) \Big\}$$

and $\lambda_i g_i(x_0) = 0$ *for all* i.

Example 4.19. Let X be a Banach space and $f : X \to \mathbb{R}$ a convex functional (not necessarily differentiable), we can extend the definition of gradient to this case.

Definition 4.20 (Subdifferential and subgradient)**.** Let $f : X \to \mathbb{R}$ be convex. For any $x_0 \in X$, the set

$$\partial f(x_0) = \{x^* \in X^* : \langle x^*, x - x_0 \rangle + f(x_0) \leq f(x),\ \forall\, x \in X\}$$

is called the *subdifferential of* f *at* x_0. Each element of $\partial f(x_0)$ is called a *subgradient of* f *at* x_0. Note that here x^* is some bounded linear functional and not related to x.

Theorem 4.21. *Let* $f : X \to \mathbb{R}$ *be convex. If* f *is continuous at* $x_0 \in X$*, then* $\partial f(x_0) \neq \emptyset$*.*

Proof. Consider $\operatorname{epi}(f)$ and $\{(x_0, f(x_0))\}$. Note that $(x_0, f(x_0)+1)$ is an interior point of $\operatorname{epi}(f)$ (since f is continuous at x_0), and $(x_0, f(x_0)) \notin \operatorname{int}(\operatorname{epi}(f))$. Hence there exists $(x^*, \xi) \in X^* \times \mathbb{R}$ that separates $\operatorname{epi}(f)$ and $(x_0, f(x_0))$, i.e.,

$$\langle x^*, x_0 \rangle + \xi f(x_0) \leq \langle x^*, x \rangle + \xi t, \quad \forall\, (x, t) \in \operatorname{epi}(f).$$

Letting $x = x_0$ and $t = f(x_0) + s$ for all $s > 0$, we see $\xi \geq 0$, otherwise $s \to \infty$ makes RHS tend to $-\infty$.

We claim that $\xi > 0$. If not, then $\xi = 0$, and

$$\langle x^*, x_0 - x \rangle \leq 0, \quad \forall\, x \in X.$$

Then $x^* = 0$ in X^*, which contradicts to (x^*, ξ) being nonzero.

Denote $x_0^* = -x^*/\xi$, then there is

$$\langle x_0^*, x - x_0 \rangle + f(x_0) \leq f(x), \quad \forall\, x \in X.$$

Hence $x_0^* \in \partial f(x_0)$, and thus $\partial f(x_0) \neq \emptyset$. ∎

5 Weak Topologies

5.1 Dual spaces

Definition 5.1 (Dual space)**.** Let X be a B* space and define the norm $\|f\| := \sup_{\|x\|=1} |f(x)|$ for every linear functional $f \in X^*$. Then $(X^*, \|\cdot\|)$ is a Banach space, called the *dual space of* X.

Example 5.2 (Dual space of L^p)**.** For any $p \in [1, \infty)$, let q be the conjugate of p, i.e., $q = p/(p-1)$ if $p > 1$ and $q = \infty$ if $p = 1$. Then $L^p([0,1])^* = L^q([0,1])$.

Proof. By Hölder inequality, we know that for any $f \in L^p$ and $g \in L^q$, there is

$$\Big|\int_0^1 fg\Big| \le \int_0^1 |fg| \le \Big(\int_0^1 |f|^p\Big)^{1/p}\Big(\int_0^1 |g|^q\Big)^{1/q},$$

and the equality holds iff

$$g(x) = \begin{cases} \operatorname{sign}(f(x))\frac{|f(x)|^{p-1}}{\|f\|_p^{p-1}}, & \text{if } p \in (1,\infty), \\ \operatorname{sign}(f(x)), & \text{if } p = 1, \end{cases} \qquad \text{a.e. } [0,1].$$

Hence, for any $g \in L^q$, we define $F_g : L^p \to \mathbb{R}$ by $F_g(f) := \int_0^1 fg\,\mathrm{d}\mu$. Then F_g is a linear functional on L^p, and $|F_g(f)| \le \|g\|_q\|f\|_p$. Hence $\|F_g\| = \|g\|_q$ and $F_g \in (L^p)^*$.

Now we want to show $(L^p)^* = L^q$, i.e., for any $F \in (L^p)^*$, there exists $g \in L^q$, such that $F(f) = \int_0^1 fg$ for all $f \in L^p$. We proceed this in three steps as follows.

1. For any $F \in (L^p)^*$, we define $\nu : \mathcal{M} \to \mathbb{R}$ as follows: $\nu(E) := F(\chi_E)$, where $E \in \mathcal{M}$ is any Lebesgue measurable set and χ_E is the characteristic function of E, i.e., $\chi_E(x) = 1$ if $x \in E$ and 0 otherwise. Then we can show that ν is a signed measure: it is easy to show that $\nu(\emptyset) = 0$ and ν only takes finite values.

We shall show that ν is countably additive. Firstly, finite additivity is clear from the definition of ν. Let $\{E_k : k \in \mathbb{N}\}$ be mutually disjoint sets in $[0,1]$. Then

$$\nu(\cup_{i=1}^\infty E_i) = \sum_{i=1}^k \nu(E_i) + \nu(D_k) = \sum_{i=1}^k F(\chi_{E_i}) + F(\chi_{D_k}),$$

where $D_k := \cup_{i=k+1}^\infty E_k$. Then $\nu(D_k) < \infty$, $\lim_k D_k = \emptyset$, and

$$\nu(D_k) = F(\chi_{D_k}) \le \|F\|\|\chi_{D_k}\|_p = \|F\|(\mu(D_k))^{1/p} \to 0$$

as $k \to \infty$. Hence ν is a signed measure. Moreover, following the similar estimate for $\nu(D_k)$, it is easy to verify that $\nu(E) = 0$ whenever $\mu(E) = 0$. Hence $\nu \ll \mu$. Thus, by Radon-Nikodym theorem, there exists a measurable function g such that $\nu(E) = \int_E g\,\mathrm{d}\mu$ for all $E \in \mathcal{M}$. Hence

$$F(\chi_E) = \nu(E) = \int_E g\,\mathrm{d}\mu = \int_0^1 \chi_E g\,\mathrm{d}\mu$$

for all $E \in \mathcal{M}$. Hence $F(f) = \int fg \, \mathrm{d}\mu$ for all simple functions f. We hereafter drop $\mathrm{d}\mu$ for notation simplicity.

2. We now show that $\|g\|_q \le \|F\|$. If $p \in (1, \infty)$, then for any $t > 0$, we denote $E_t := \{x \in [0,1] : |g(x)| \le t\}$ and let $f := \chi_{E_t} |g|^{q-2} g$. Then $F(f) = \int_0^1 fg$ (see Remark below). It is easy to verify that

$$\int_0^1 fg = \int_{E_t} |g|^q \quad \text{and} \quad \int_0^1 |f|^p = \int_{E_t} |g|^q.$$

Therefore, we can show that

$$\int_{E_t} |g|^q = \int_0^1 fg = F(f) \le \|F\| \|f\|_p = \|F\| \Big(\int_{E_t} |g|^q \Big)^{1/p}.$$

Hence $(\int_{E_t} |g|^q)^{1/q} \le \|F\|$. Letting $t \to \infty$, we have $E_t \uparrow [0,1]$ and thus $\|g\|_q \le \|F\|$.

If $p = 1$, then for any $\epsilon > 0$, we denote $A_\epsilon := \{x \in [0,1] : |g(x)| \ge \|F\| + \epsilon\}$ and let $f = \chi_{E_t \cap A_\epsilon} \mathrm{sign}(g)$, then there is

$$\int_0^1 |f| = \int_0^1 \chi_{E_t \cap A_\epsilon} = \mu(E_t \cap A_\epsilon).$$

On the other hand, we know

$$F(f) = \int_0^1 fg = \int_0^1 \chi_{E_t \cap A_\epsilon} |g| = \int_{E_t \cap A_\epsilon} |g| \ge (\|F\| + \epsilon) \mu(E_t \cap A_\epsilon).$$

Hence there is

$$0 \le (\|F\| + \epsilon) \mu(E_t \cap A_\epsilon) \le F(f) \le \|F\| \|f\|_1 = \|F\| \mu(E_t \cap A_\epsilon).$$

Letting $t \to \infty$, we have $E_t \uparrow [0,1]$ and hence $(\|F\| + \epsilon)\mu(A_\epsilon) \le \|F\| \mu(A_\epsilon)$. Thus $\mu(A_\epsilon) = 0$. As $\epsilon > 0$ is arbitrary, we know $\|g\|_\infty \le \|F\|$.

3. Now we show $F(f) = \int_0^1 fg$ for any $f \in L^p$. Let f_k be simple functions such that $f_k \to f$ in L^p, then

$$\Big| \int_0^1 (f_k - f) g \Big| \le \|f_k - f\|_p \|g\|_q \le \|F\| \|f_k - f\|_p \to 0.$$

Hence $F(f_k) = \int_0^1 f_k g \to \int_0^1 fg$. As $F \in (L^p)^*$, we know $F(f_k) \to F(f)$, and hence $F(f) = \int_0^1 fg$. □

Remark. In Step 2 of the proof above, we used the fact that $F(f) = \int_0^1 fg$ where $f = \chi_{E_t} |g|^{q-2} g$ is not simple when $p \in (1, \infty)$. To see this, let f_k be a sequence of simple functions such that $f_k \in L^p$ and $f_k^{\pm} \uparrow f^{\pm}$ in L^p. Then, on the one hand, we know $F(f_k) \to F(f)$ since $F \in (L^p)^*$; and on the other hand, there is $f_k g \uparrow \chi_{E_t} |g|^q = fg$, which by Beppo Levi theorem we know $\int_0^1 f_k g \to \int_0^1 \chi_{E_t} |g|^q = \int_0^1 fg$. Since $F(f_k) = \int_0^1 f_k g$ for every simple function f_k, we know $F(f) = \int_0^1 fg$.

Remark. The result above can be extended from $[0,1]$ to any σ-finite measurable space.

Example 5.3 (Dual space of $C([0,1])$)**.** The dual space of $C([0,1])$ is the bounded variation space $\mathrm{BV}([0,1]) := \{f : [0,1] \to \mathbb{R} : f(0) = 0, \|f\|_{\mathrm{TV}} < \infty\}$, where $\|f\|_{\mathrm{TV}}$ is the total variation of f defined by

$$\|f\|_{\mathrm{TV}} := \sup_{\Delta} \sum_{i=1}^{n} |f(t_i) - f(t_{i-1})|,$$

and $\Delta : 0 = t_0 < t_1 < \cdots < t_n = 1$ is a partition of $[0,1]$ and $n \in \mathbb{N}$ is arbitrary.

Proof. 1. It is straightforward to show that $(\mathrm{BV}([0,1]), \|\cdot\|_{\mathrm{TV}})$ is a Banach space: $\mathrm{BV}([0,1])$ is a linear space, $\|\cdot\|_{\mathrm{TV}}$ is a norm, and $\mathrm{BV}([0,1])$ is complete under $\|\cdot\|_{\mathrm{TV}}$ (check yourself).

2. For any $x \in C([0,1])$ and $f \in \mathrm{BV}([0,1])$, consider the *Stieltjes integral*:

$$\int_0^1 x(t)\,\mathrm{d}f(t) := \lim_{|\Delta|\to 0} \sum_{i=1}^{n} x(t_i^*)(f(t_i) - f(t_{i-1})),$$

where $t_i^* \in [t_{i-1}, t_i]$ and $|\Delta| := \max_{1\le i\le n} |t_i - t_{i-1}|$. We can show $F_f : X \to \mathbb{R}$ defined by $F_f(x) := \int_0^1 x(t)\,\mathrm{d}f(t)$ is a bounded linear functional on $C([0,1])$: the linearity is obvious, and the boundedness is because for any $x \in C([0,1])$ and $\|x\| \le 1$, there is $|F_f(x)| = |\int_0^1 x(t)\,\mathrm{d}f(t) \le \|f\|_{\mathrm{TV}}$.

3. We shall show that, for any $F \in C([0,1])^*$, there exists a unique $f \in \mathrm{BV}([0,1])$, such that $F(x) = \int_0^1 x(t)\,\mathrm{d}f(t)$ for all $x \in C([0,1])$. The key is to determine f using the linear functional F, as shown below.

We first observe that $C([0,1])$ is a closed linear subspace of $(L^\infty([0,1]), \|\cdot\|_\infty)$. Then by Corollary 4.2, we know there exists $\tilde{F} \in L^\infty([0,1])^*$ such that $\tilde{F}|_{[0,1]} = F$, $\|\tilde{F}\| = \|F\|$. Define

$$f(s) := \begin{cases} \tilde{F}(\chi_{(0,s]}), & 0 < s \le 1, \\ 0, & s = 0. \end{cases}$$

Next we show that $f \in \mathrm{BV}([0,1])$, $\|f\|_{\mathrm{TV}} \le \|F\|$, and $F(x) = \int_0^1 x(t)\,\mathrm{d}f(t)$ for all $x \in C([0,1])$.

First of all, for any partition $\Delta : 0 = t_0 < t_1 < \cdots < t_n = 1$ of $[0,1]$, denote $\lambda_i = \mathrm{sign}(f(t_i) - f(t_{i-1}))$. Then

$$\begin{aligned}\sum_{i=1}^{n} |f(t_i) - f(t_{i-1})| &= \sum_{i=1}^{n} \lambda_i(f(t_i) - f(t_{i-1})) = \sum_{i=1}^{n} \lambda_i(\tilde{F}(\chi_{(0,t_i]}) - \tilde{F}(\chi_{(0,t_{i-1}]}) \\ &= \tilde{F}\left(\sum_{i=1}^{n} \lambda_i \chi_{(t_{i-1},t_i]}\right) \le \|\tilde{F}\| = \|F\|,\end{aligned}$$

since $\| \sum_{i=1}^n \lambda_i \chi_{(t_{i-1},t_i]}\|_\infty = 1$. Since Δ is arbitrary, we have $\|f\|_{\mathrm{TV}} \le \|F\|$. For any $x \in C([0,1])$ and $\epsilon > 0$, choose a partition Δ such that

$$|x(t) - x(t')| \le \frac{\epsilon}{2\|F\|}, \quad \forall\, t, t' \in [t_{i-1}, t_i],\ i = 1, \dots, n,$$
$$\Big|\int_0^1 x(t)\,\mathrm{d}f(t) - \sum_{i=1}^n x(t_i)(f(t_i) - f(t_{i-1}))\Big| < \frac{\epsilon}{2},$$

where the first inequality is because of the uniform continuity of x on $[0,1]$, and the second inequality is due to the definition of Stieltjes integral. We denote $x_\Delta := \sum_{i=1}^n x(t_i)\chi_{(t_{i-1},t_i]} + x(0)\chi_{\{0\}}$, then $\tilde{F}(x_\Delta) = \sum_{i=1}^n x(t_i)(f(t_i) - f(t_{i-1}))$, and

$$\begin{aligned}\Big|F(x) - \int_0^1 x(t)\,\mathrm{d}f(t)\Big| &\le |F(x) - \tilde{F}(x_\Delta)| + \Big|\tilde{F}(x_\Delta) - \int_0^1 x(t)\,\mathrm{d}f(t)\Big| \\ &\le \|F\|\|x - x_\Delta\|_\infty + \Big|\tilde{F}(x_\Delta) - \int_0^1 x(t)\,\mathrm{d}f(t)\Big| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.\end{aligned}$$

As $\epsilon > 0$ is arbitrary, we know $F(x) = \int_0^1 x(t)\,\mathrm{d}f(t)$. □

5.2 Bidual and reflexive space

We showed that X^* is a Banach space, hence we can consider X^{**}, called the *bidual space of* X. In particular, for any $x \in X$, we can define $J_x : X^* \to \mathbb{R}$ by $J_x(f) = f(x)$ for any $f \in X^*$. It is clear that J_x is linear, and $|J_x(f)| = |f(x)| \le \|x\|\|f\|$, which implies $\|J_x\| \le \|x\|$ and hence $J_x \in X^{**}$.

Now we can define $T : X \to X^{**}$ by $Tx := J_x$. Then T is called the *canonical mapping* (also called *evaluation mapping*). It is easy to show that T is linear: for any $x, y \in X$ and $\alpha, \beta \in \mathbb{R}$, there is

$$\begin{aligned}\big(T(\alpha x + \beta y)\big)(f) &= J_{\alpha x + \beta y}(f) = f(\alpha x + \beta y) \\ &= \alpha f(x) + \beta f(y) = \alpha J_x(f) + \beta J_y(f) \\ &= \alpha(Tx)(f) + \beta(Ty)(f) = (\alpha Tx + \beta Ty)(f),\end{aligned}$$

for all $f \in X^*$. Moreover, we have showed that $\|Tx\| = \|J_x\| \le \|x\|$ for all $x \in X$, so $T \in L(X, X^{**})$.

In fact, we can show $\|Tx\| = \|J_x\| = \|x\|$: for any $x \in X$, there exists $f \in X^*$ such that $\|f\| = 1$ and $\|f(x) = \|x\|$ (e.g., $f : X_0 \to \mathbb{R}$ defined by $f(ax) = a\|x\|$ where $X_0 := \{ax : a \in \mathbb{R}\}$, and then use Corollary 4.2 to extend f from X_0 to X). Then

$$\|x\| = f(x) = J_x(f) = (Tx)(f) \le \|Tx\|\|f\| = \|Tx\|.$$

Therefore $\|Tx\| = \|x\|$, and thus T is an isometric isomorphism.

If X is a Banach space, then we can show that X is isometrically embedded as a closed linear subspace of X^{**}. To see this, let $\{Tx_k\}$ be Cauchy in X^{**}, then $\{x_k\}$ is Cauchy in X due to the isometry between X and $T(X)$. Hence there exists $x \in X$ such that $x_k \to x$, and it is easy to verify that $Tx_k \to Tx$ in X^{**}. This result is summarized in the following theorem.

Theorem 5.4. *Let X be a Banach space. Then X is isometrically isomorphic to a closed linear subspace of X^{**}.*

Remark. We often just identify $Tx = J_x$ with $x \in X$.

Definition 5.5 (Reflexive space)**.** A B* space X is called *reflexive* if $X = X^{**}$. Namely, the canonical mapping $T : x \mapsto Tx$ is surjective.

Example 5.6. If $p \in (1, \infty)$, then $(L^p)^* = L^q$ and $(L^p)^{**} = (L^q)^* = L^p$. In addition, $(L^1)^* = L^\infty$, and

$$(L^\infty)^* = \text{ba}(\Omega) := \{\nu : \mathcal{M} \to \mathbb{R} : \nu \text{ is bounded, finitely additive, and } \nu \ll \mu\}.$$

We now consider adjoint operators, which are generalization of the conjugate transpose operator of matrices.

Definition 5.7 (Adjoint operator)**.** Let X and Y be B* spaces, $T \in L(X, Y)$. Then $T^* : Y^* \to X^*$ is called the *adjoint operator of* T if for any $f \in Y^*$ and $x \in X$, there is

$$(T^* f)(x) = f(Tx).$$

This is also written as $\langle T^* f, x\rangle = \langle f, Tx\rangle$.

Remark. T^* is well defined. To see this, for any $T \in L(X, Y)$ and $f \in Y^*$, define $g_f : X \to \mathbb{R}$ by $g_f(x) = f(Tx)$. Then g_f is linear and $|g_f(x)| = |f(Tx)| \leq \|f\|\|T\|\|x\|$. So $g_f \in X^*$. Then the adjoint operator $T^* : Y^* \to X^*$ by $T^* f := g_f$. It is easy to show that T^* is linear, and $\|T^* f\| = \|g_f\| \leq \|f\|\|T\|$ implies $\|T^*\| \leq \|T\|$, so $T^* \in L(Y^*, X^*)$. In fact we can show the following theorem.

Theorem 5.8. *Let the mapping $* : L(X, Y) \to L(Y^*, X^*)$ be defined by $*(T) = T^*$. Then $*$ is an isometric isomorphism.*

Proof. 1. We first show that * is linear. For any $\alpha_1, \alpha_2 \in \mathbb{R}$ and $T_1, T_2 \in L(X, Y)$, we know

$$\begin{aligned}
[*(\alpha_1 T_1 + \alpha_2 T_2)(f)](x) &= [(\alpha_1 T_1 + \alpha_2 T_2)^* f](x) = f(\alpha_1 T_1 x + \alpha_2 T_2 x) \\
&= \alpha_1 f(T_1 x) + \alpha_2 f(T_2 x) = \alpha_1 (T_1^* f)(x) + \alpha_2 (T_2^* f)(x) \\
&= \alpha_1 [(*T_1)(f)](x) + \alpha_2 [(*T)(f)](x) \\
&= [(\alpha_1 (*T_1) + \alpha_2 (*T_2)) f](x)
\end{aligned}$$

for all $x \in X$ and $f \in Y^*$.

2. We only need to show $\|T\| \leq \|T^*\|$. To this end, for any $x \in X$ such that $Tx \neq 0$, by Corollary 4.4, there exists $f \in Y^*$ such that $f(Tx) = \|Tx\|$ and $\|f\| = 1$. Then

$$\|Tx\| = f(Tx) = (T^* f)(x) \leq \|T^* f\|\|x\| \leq \|T^*\|\|f\|\|x\| = \|T^*\|\|x\|.$$

Therefore $\|T\| \leq \|T^*\|$. ☐

We can further consider the adjoint operator $T^{**} := (T^*)^*$ of T^*. Then $T^{**} \in L(X^{**}, Y^{**})$. Since $X \subset X^{**}$ and $Y \subset Y^{**}$, we denote U and V their canonical embedding respectively (i.e, $Ux := J_x = x^{**}$ and $Vy := J_y = y^{**}$). Then for any $x \in X$ and $f \in Y^*$, we have

$$\langle T^{**}Ux, f\rangle = \langle Ux, T^*f\rangle = \langle T^*f, x\rangle = \langle f, Tx\rangle = \langle VTx, f\rangle.$$

Hence $T^*Ux = VTx$, namely T^{**} is the extension of TX from X (as the embedded in X^{**}) to X^{**}. This is given in the following theorem.

Theorem 5.9. *Let X and Y be B^* space and $T \in L(X, Y)$. Then $T^{**} \in L(X^{**}, Y^{**})$ is the extension of T on X^{**} and $\|T^{**}\| = \|T\|$.*

Example 5.10. Let (Ω, μ) be a measure space, and $K : \Omega \times \Omega \to \mathbb{R}$ be a square integrable function on $\Omega \times \Omega$, i.e.,

$$\|K\| := \Big(\iint_{\Omega\times\Omega} |K(x,y)|^2 \,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y)\Big)^{1/2} < \infty.$$

Define the operator $T : L^2(\Omega, \mu) \to L^2(\Omega, \mu)$ by

$$(Tu)(x) = \int_\Omega K(x,y)u(y)\,\mathrm{d}\mu(y), \quad \forall\, u \in L^2(\Omega,\mu),\ \forall\, x \in \Omega.$$

Then it is easy to show T is a bounded linear mapping from $L^2(\Omega, \mu)$ to itself: the linearity is clear, and

$$\begin{aligned}
\|Tu\|^2 &= \int_\Omega \Big| \int_\Omega K(x,y)u(y)\,\mathrm{d}\mu(y)\Big|^2 \mathrm{d}\mu(x) \\
&\le \int_\Omega \Big[\int_\Omega |K(x,y)|^2\,\mathrm{d}\mu(y) \int_\Omega |u(y)|^2\,\mathrm{d}\mu(y)\Big]\,\mathrm{d}\mu(x) \\
&= \Big[\iint_{\Omega\times\Omega} |K(x,y)|^2\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y)\Big]\,\|u\|^2.
\end{aligned}$$

Hence $\|T\| \le \|K\|$.

We claim that $T^* : L^2(\Omega, \mu) \to L^2(\Omega, \mu)$ is given by

$$(T^*v)(x) = \int_\Omega K(y,x)v(y)\,\mathrm{d}\mu(y), \quad \forall v \in L^2(\Omega, \mu).$$

To see this, notice that

$$\begin{aligned}
\langle v, Tu\rangle &= \int_\Omega v(x)\Big[\int_\Omega K(x,y)u(y)\,\mathrm{d}\mu(y)\Big]\,\mathrm{d}\mu(x) \\
&= \int_\Omega u(y)\Big[\int_\Omega K(x,y)v(x)\,\mathrm{d}\mu(x)\Big]\,\mathrm{d}\mu(y) \\
&= \int_\Omega u(x)\Big[\int_\Omega K(y,x)v(y)\,\mathrm{d}\mu(y)\Big]\,\mathrm{d}\mu(x) \\
&= \langle T^*v, u\rangle,
\end{aligned}$$

where the second inequality is due to Fubini theorem, which holds here because

$$
\begin{aligned}
&\Big|\iint_{\Omega\times\Omega} K(x,y)u(x)v(y)\,\mathrm{d}\mu(x)\mu(y)\Big| \\
&\le \iint_{\Omega\times\Omega} |K(x,y)||u(x)||v(y)|\,\mathrm{d}\mu(x)\mu(y) \\
&\le \Big(\iint_{\Omega\times\Omega} |K(x,y)|^2\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y)\Big)^{1/2}\Big(\iint_{\Omega\times\Omega} |u(x)|^2|v(y)|^2\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y)\Big)^{1/2} \\
&= \Big(\iint_{\Omega\times\Omega} |K(x,y)|^2\,\mathrm{d}\mu(x)\,\mathrm{d}\mu(y)\Big)^{1/2}\|u\|\|v\| < \infty.
\end{aligned}
$$

Lemma 5.11 (Young's inequality)**.** *Let* $p \in [1,\infty]$, $f \in L^p(\mathbb{R})$ *and* $K \in L^1(\mathbb{R})$, *then* $\|K * f\|_p \le \|K\|_1\|f\|_p$.

Proof. If $p \in (1,\infty)$, then by Hölder inequality, we know for any $x \in \mathbb{R}$ there is

$$
\begin{aligned}
\Big|\int_{-\infty}^{\infty} K(x-y)f(y)\,\mathrm{d}y\Big| &\le \int_{-\infty}^{\infty} |K(x-y)|^{1/q}|K(x-y)|^{1/p}|f(y)|\,\mathrm{d}y \\
&\le \Big(\int_{-\infty}^{\infty} |K(x-y)|\,\mathrm{d}y\Big)^{1/q}\Big(\int_{-\infty}^{\infty} |K(x-y)||f(y)|^p\,\mathrm{d}y\Big)^{1/p} \\
&= \|K\|_1^{1/q}\Big(\int_{-\infty}^{\infty} |K(x-y)||f(y)|^p\,\mathrm{d}y\Big)^{1/p}.
\end{aligned}
$$

Hence, there is

$$
\begin{aligned}
\|K * f\|_p^p &= \int_{-\infty}^{\infty}\Big|\int_{-\infty}^{\infty} K(x-y)f(y)\,\mathrm{d}y\Big|^p\,\mathrm{d}x \\
&\le \|K\|_1^{p/q}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |K(x-y)||f(y)|^p\,\mathrm{d}y\,\mathrm{d}x \\
&= \|K\|_1^{p/q}\|K\|_1\|f\|_p^p,
\end{aligned}
$$

which implies that $\|K * f\|_p \le \|K\|_1\|f\|_p$.

If $p = \infty$, then for any $x \in \mathbb{R}$, we know

$$
\Big|\int_{-\infty}^{\infty} K(x-y)f(y)\,\mathrm{d}y\Big| \le \|f\|_\infty \int_{-\infty}^{\infty} |K(x-y)|\,\mathrm{d}y = \|f\|_\infty\|K\|_1.
$$

Hence $\|K * f\|_\infty \le \|K\|_1\|f\|_\infty$.

If $p = 1$, then

$$
\begin{aligned}
\int_{-\infty}^{\infty}\Big|\int_{-\infty}^{\infty} K(x-y)f(y)\,\mathrm{d}y\Big|\,\mathrm{d}x &\le \int_{-\infty}^{\infty}\Big(\int_{-\infty}^{\infty} |K(x-y)|\,\mathrm{d}x\Big)|f(y)|\,\mathrm{d}y \\
&\le \|K\|_1\|f\|_1.
\end{aligned}
$$

Hence $\|K * f\|_1 \le \|K\|_1\|f\|_1$. □

Example 5.12 (Adjoint of convolution)**.** Let $K \in L(\mathbb{R})$ and $p \in [1, \infty)$. Define the convolution $c_K : L^p(\mathbb{R}) \to L^p(\mathbb{R})$ as follows:

$$(c_K f)(x) = \int_{-\infty}^{\infty} K(x-y) f(y)\,\mathrm{d}y.$$

i.e., $c_K f = K * f \in L^p$ is the convolution of f using kernel K.

It is easy to show that c_K is linear, and by Young's inequality we know $\|c_K\| \le \|K\|_1$. Hence c_K is continuous.

We claim that $c_K^* = c_{\tilde{K}}$ where $\tilde{K}(x) = K(-x)$ for all $x \in \mathbb{R}$. To see this, we note that

$$\begin{aligned}
\langle g, c_K f\rangle = \langle g, K * f\rangle &= \int_{-\infty}^{\infty} \Big(\int_{-\infty}^{\infty} K(x-y) f(y)\,\mathrm{d}y \Big) g(x)\,\mathrm{d}x \\
&= \int_{-\infty}^{\infty} f(y) \Big(\int_{-\infty}^{\infty} K(x-y) g(x)\,\mathrm{d}x \Big)\,\mathrm{d}y \\
&= \int_{-\infty}^{\infty} f(y) \Big(\int_{-\infty}^{\infty} \tilde{K}(y-x) g(x)\,\mathrm{d}x \Big)\,\mathrm{d}y \\
&= \int_{-\infty}^{\infty} (\tilde{K} * g)(y) f(y)\,\mathrm{d}y \\
&= \langle \tilde{K} * g, f\rangle = \langle c_{\tilde{K}} g, f\rangle
\end{aligned}$$

for all $f \in L^p$ and $g \in L^q$.

5.3 Weak and weak* convergence

Definition 5.13 (Weak convergence)**.** Let X be a B* space, $x_k, x_0 \in X$ for all k. Then $\{x_k\}$ is said to *weakly converge* to x_0, denoted $x_k \rightharpoonup x_0$, if for any $f \in X^*$ there is $\lim_{k\to\infty} f(x_k) = f(x_0)$. In this case, x_0 is called the weak limit of $\{x_k\}$. To distinguish, the classical convergence $x_k \to x_0$ (i.e., $\|x_k - x_0\| \to 0$) is also called *strong convergence*.

Remark. If $\dim(X) < \infty$, then weak convergence reduces to strong convergence. To see this, we let $e_1, \dots, e_n$ be a basis (complete orthonormal set) of the n-dimensional Banach space X. Then

$$\begin{aligned}
x_k &= a_1^{(k)} e_1 + \cdots + a_n^{(k)} e_n, \quad \forall\, k \in \mathbb{N}, \\
x_0 &= a_1^{(0)} e_1 + \cdots + a_n^{(0)} e_n.
\end{aligned}$$

Now define $f_i \in X^*$ such that $f_i(e_j) = \delta_{ij}$ for all $j = 1, \dots, n$. Then there are $f_i(x_k) = a_i^{(k)}$ and $f_i(x_0) = a_i^{(0)}$ for all $i = 1, \dots, n$ and $k \in \mathbb{N}$. If $x_k \rightharpoonup x$, then $f_i(x_k) \to f_i(x_0)$, i.e., $a_i^{(k)} \to a_i^{(0)}$. Hence $\|x_k - x_0\| \le \sum_{i=1}^n |a_i^{(k)} - a_i^{(0)}| \to 0$, which means $x_k \to x$ strongly.

Proposition 5.14. *Weak limit is unique if exists.*

Proof. Suppose $x_k \rightharpoonup x$ and $x_k \rightharpoonup y$. Then for any $f \in X^*$, we know $f(x_k) \to f(x)$ and $f(x_k) \to f(y)$. Hence $f(x) = f(y)$, i.e., $f(x-y) = 0$ for all $f \in X^*$. Therefore $x = y$. □

Proposition 5.15. *If $x_k \to x_0$ then $x_k \rightharpoonup x_0$.*

Proof. For any $f \in X^*$, we know

$$|f(x_k) - f(x_0)| = |f(x_k - x_0)| \leq \|f\| \|x_k - x_0\| \to 0.$$

Therefore strong convergence implies weak convergence. □

Example 5.16. Consider $L^2([0,1])$. Let $x_k(t) := \sin(k\pi t)$. By Riemann-Lebesgue theorem, we know for any $f \in L^2([0,1])$, there is

$$\langle f, x_k \rangle = \int_0^1 f(t) \sin(k\pi t)\, \mathrm{d}t \to 0$$

as $k \to \infty$. Hence $x_k \rightharpoonup 0$. However, $\|x_k\|_2 = 1/\sqrt{2}$ for all $k \in \mathbb{N}$ and hence $\{x_k\}$ does not converge strongly.

Although in general weak convergence does not imply strong convergence, we can show that convex combinations of weakly convergence sequence may strongly converge to the weak limit, as shown in the following theorem.

Theorem 5.17 (Mazur)**.** *Let X be a B^* space and $x_k \rightharpoonup x_0$. Then for any $\epsilon > 0$, there exist $n \in \mathbb{N}$, $\lambda_i \geq 0$ and $\sum_{i=1}^n \lambda_i = 1$, such that $\|x_0 - \sum_{i=1}^n \lambda_i x_i\| \leq \epsilon$.*

Proof. Let $M := \overline{\mathrm{conv}(\{x_k\})}$, then M is a closed convex set of M. If $x_0 \notin M$, then by Theorem 4.15 (Ascoli), there exists $f \in X^*$ and $a \in \mathbb{R}$ such that $f(x) < a < f(x_0)$ for all $x \in M$. As $\{x_k\} \subset M$, we know $f(x_k) < a < f(x_0)$ for all $k \in \mathbb{N}$, which contradicts to $x_k \rightharpoonup x_0$. □

Since X^* is a Banach space, we can consider weak convergence in X^*: let $f_k, f \in X^*$, then $f_k \rightharpoonup f$ if $x^{**}(f_k) \to x^{**}(f)$ for any $x^{**} \in X^{**}$. However, we sometimes want to study convergence of $\{f_k\}$ without invoking X^{**}.

Definition 5.18 (Weak* convergence)**.** Let X be a B* space, $f_k, f \in X$. Then $\{f_k\}$ is said to *weak* converge* to f, denoted by $f_k \overset{*}{\rightharpoonup} f$ or $w^* - \lim_{k\to\infty} f_k = f$, if $\lim_{k\to\infty} f_k(x) = f(x)$ for any $x \in X$. In this case, f is called the *weak* limit* of $\{f_k\}$.

Remark. We know $X \subset X^{**}$, hence weak* convergence is weaker than weak convergence. If X is reflexive, then weak* convergence and weak convergence are equivalent. It is also easy to show that weak* limit is unique.

As a direct application of Theorem 3.38 (Banach-Steinhaus), we have the following result.

Theorem 5.19. *Let X be a B^* space, and $x_k, x \in X$. Then $x_k \rightharpoonup x$ iff $\|x_k\|$ is bounded and there is a dense subset M^* of X^* such that $f(x_k) \to f(x)$ for all $f \in M^*$.*

Proof. We treat $\{x_k\}$ as a sequence in X^{**}, then applying Theorem 3.38 (Banach-Steinhaus) yields the claim. □

Theorem 5.20. *Let X be a Banach space, $f_k, f \in X^*$. Then $f_k \overset{*}{\rightharpoonup} f$ iff $\|f_k\|$ is bounded and there exists a dense set M of X such that $f_k(x) \to f(x)$ for all $x \in M$.*

Similar to linear functional, we can consider various convergences of linear operators.

Definition 5.21 (Convergence of linear operators)**.** Let X and Y be B* spaces, $T_k, T \in L(X,Y)$ for all $k \in \mathbb{N}$. Then

- $\{T_k\}$ is said to *uniformly converge to* T, denoted by $T_k \rightrightarrows T$, if $\|T_k - T\| \to 0$.
- $\{T_k\}$ is said to *strongly converge to* T, denoted by $T_k \to T$, if $\|T_k x - Tx\| \to 0$ for any $x \in X$.
- $\{T_k\}$ is said to *weakly converge to* T, denoted by $T_k \rightharpoonup T$, if $f(T_k x) \to f(Tx)$ for all $x \in X$ and $f \in Y^*$.

Remark. It is easy to verify that

$$\text{Uniform convergence} \implies \text{Strong convergence} \implies \text{Weak convergence}$$

Moreover, all these limits are unique if exist.

Example 5.22 (Strongly convergent but not uniformly convergent)**.** Denote $T : l^2 \to l^2$ by $Tx = (x_2, x_3, \dots)$ for any $x = (x_1, x_2, \dots) \in l^2$. Namely, T is the left shift operator. Let $T_k := T^k$, i.e., $T_k x = (x_{k+1}, x_{k+2}, \dots)$. It is clear that $T_k \in L(l^2)$ for all $k \in \mathbb{N}$.

We first show that $T_k \to 0$: for any $x \in l^2$, we have $\|T_k x\| = (\sum_{i=k+1}^{\infty} x_i^2)^{1/2} \to 0$ as $k \to \infty$.

We then show that T_k is not uniformly converging to 0: denote $e_k = (0, \dots, 0, 1, 0, \dots)$, then $\|T_k e_{k+1}\| = \|e_1\| = 1$. Hence $\|T_k\| \geq 1$ for all $k \in \mathbb{N}$.

Example 5.23 (Weakly convergent but not strongly convergent)**.** Denote $S : l^2 \to l^2$ by $Sx = (0, x_1, x_2, \dots)$. Namely S is the right shift operator. Define $S_k := S^k$. It is clear that $S_k \in L(l^2)$ for all $k \in \mathbb{N}$.

We first show that $S_k \rightharpoonup 0$. To see this , we know for any $f \in (l^2)^* = l^2$, there exists $y_f = (y_1, y_2, \dots) \in l^2$ such that

$$\begin{aligned} f(S_k x) = \langle y_f, S_k x\rangle &= \sum_{i=1}^{\infty} y_{i+k} x_i \\ &\leq \Big(\sum_{i=1}^{\infty} |y_{i+k}|^2\Big)^{1/2} \|x\| \leq \Big(\sum_{i=k+1}^{\infty} |y_i|^2\Big)^{1/2} \|x\| \to 0 \end{aligned}$$

as $k \to \infty$. Hence $S_k \rightharpoonup 0$.

But it is easy to show that S_k does not converge to 0 strongly: for any $x \in l^2$ there is $\|S_k x\| = \|x\|$ for all $k \in \mathbb{N}$.

5.4 Weak and weak* compactness

One of the main goals to have weak and weak* convergence is to derive weak and weak* sequential compactness from boundedness. We say E is *weakly sequentially precompact* if any sequence in A has a weakly convergent subsequence, and E is *weak* sequentially precompact* if any sequence in A has a weak* convergent subsequence. If the weak or weak* limit is also in E respectively, then precompactness improves to compactness.

Theorem 5.24. *Let X be a separable B^* space. Then any bounded sequence $\{f_k\}$ in X^* has a weak* convergent subsequence.*

Proof. Since X is separable, there exists a countable dense set $\{x_m\}$ of X. Since $\{f_k\}$ is bounded, we know for any x_m, $\{f_k(x_m)\}$ is a bounded sequence in $\mathbb{R}$. Therefore, for $m = 1$, there exists a subsequence, denoted by $\{f_{k_j}\}$, such that $\{f_{k_j}(x_1)\}$ is convergent. Similarly, for $m = 2$, there exists a subsequence of $\{f_{k_j}\}$, still denoted by $\{f_{k_j}\}$, such that $\{f_{k_j}(x_2)\}$ is convergent, and so on. Continue doing so, we obtain a subsequence $\{f_{k_j}\}$, such that $\{f_{k_j}(x_m)\}$ is convergent for every $m \in \mathbb{N}$.

For any $x \in X$, there exists a sequence $\{x_{m_i}\}$ such that $\lim_{i\to\infty} x_{m_i} = x$ since $\{x_m\}$ is dense in X. Then we can show $\{f_{k_j}(x)\}$ is convergent (to see this, note that $|f_{k_j}(x) - f_{k_{j'}}(x)| \le |f_{k_j}(x) - f_{k_j}(x_{m_i})| + |f_{k_j}(x_{m_i}) - f_{k_{j'}}(x_{m_i})| + |f_{k_{j'}}(x_{m_i}) - f_{k_{j'}}(x)|$ and show $\{f_{k_j}(x)\}$ is a Cauchy sequence in $\mathbb{R}$). Denote $f : X \to \mathbb{R}$ by $f(x) = \lim_{j\to\infty} f_{k_j}(x)$. Then it is easy to show f is linear and $|f(x)| \le (\sup_{k\in\mathbb{N}} \|f_{k_j}\|)\|x\|$, i.e., $f \in X^*$. Hence $f_k \overset{*}{\rightharpoonup} f$ in X^*. ☐

If X is reflexive, we can show weak* convergence (which is equivalent to weak convergence now) without assuming X to be separable. The following theorems lead us to this result.

Theorem 5.25 (Banach)**.** *Let X be a B^* space. If X^* is separable, then X is separable.*

Proof. 1. We first show that $S_1^* := \{f \in X^* : \|f\| = 1\}$ is separable. To see this, let $\{f_k\}$ be a countable dense set of X^*, then for any $f \in S_1^*$, there exists a sequence $\{k_j\}$ in $\mathbb{N}$ such that $\lim_{j\to\infty} f_{k_j} = f$. Let $g_k := f_k/\|f_k\|$ for all $k \in \mathbb{N}$ (WLOG, we assume $f_k \neq 0$), then

$$\begin{aligned}\|f - g_{k_j}\| &\le \|f - f_{k_j}\| + \|f_{k_j} - g_{k_j}\| \\ &= \|f - f_{k_j}\| + \|f_{k_j} - f_{k_j}/\|f_{k_j}\|\| \\ &= \|f - f_{k_j}\| + |\ \|f_{k_j}\| - 1\ | \to 0,\end{aligned}$$

as $j \to \infty$, since $f_{k_j} \to f$ and $\|f_{k_j}\| \to \|f\| = 1$. Hence $\{g_k\}$ is a countable dense set of S_1^*.

2. Since $\|g_k\| = 1$, there exists $x_k \in S_1 := \{x \in X : \|x\| \leq 1\}$ such that $|g_k(x_k)| \geq 1/2$ for all $k \in \mathbb{N}$. Let $X_0 = \overline{\text{span}(\{x_k\})}$. Then X_0 is separable (to see this, consider the set of all linear combinations of $\{x_k\}$ with coefficients in $\mathbb{Q}$ and show that it is dense in X_0).

3. We now show $X_0 = X$. If not, then there exists $x_0 \in X \setminus X_0$ such that $d(x_0, X_0) > 0$. By Theorem 4.5, there exists $f_0 \in X^*$ such that $\|f_0\| = 1$ and $f_0(x) = 0$ for all $x \in X_0$. Note that $f_0 \in S_1^*$, but for any $k \in \mathbb{N}$ there is

$$\|g_k - f_0\| = \sup_{\|x\|=1} \|g_k(x) - f_0(x)\| \geq |g_k(x_k) - f_0(x_k)| = |g_k(x_k)| \geq \frac{1}{2},$$

which contradicts to $\{g_k\}$ being dense in S_1^*. Hence $X = X_0$ and is thus separable. □

Theorem 5.26 (Pettis)**.** *Let X be a Banach space. If X is reflexive and X_0 is a closed linear subspace of X, then X_0 is reflexive.*

Proof. We need to show for any $z_0 \in X_0^{**}$, there exists $x_0 \in X_0$ such that $\langle z_0, f_0 \rangle = \langle f_0, x_0 \rangle$ for all $f_0 \in X_0^*$.

For any $f \in X^*$, consider its restriction mapping T to X_0^*. Namely, $T : X^* \to X_0^*$ such that $Tf = f_0 := f|_{X_0} \in X_0^*$. Since $\|Tf\| = \|f_0\| \leq \|f\|$, we know $T \in L(X^*, X_0^*)$ and thus $T^* \in L(X_0^{**}, X^{**})$. Let $z := T^* z_0 \in X^{**}$. Since X is reflexive, we know there exists $x \in X$ such that $\langle z, f \rangle = \langle f, x \rangle$ for all $f \in X^*$.

We claim that $x \in X_0$: If not, then $\delta := d(x, X_0) > 0$. By Theorem 4.5, there exists $f \in X^*$ such that $f|_{X_0} = 0$ (thus $Tf = 0$) and $\langle f, x \rangle = f(x) = \delta$. However we also have

$$0 = \langle z_0, Tf \rangle = \langle T^* z_0, f \rangle = \langle z, f \rangle = \langle f, x \rangle = \delta > 0,$$

which is a contradiction. Hence $x \in X_0$.

Now we show $\langle z_0, f_0 \rangle = \langle f_0, x \rangle$ for all $f_0 \in X_0^*$. For any $f_0 \in X_0^*$, by Corollary 4.2, there exists $f \in X^*$ such that $Tf = f_0$. Hence $\langle z_0, f_0 \rangle = \langle z_0, Tf \rangle = \langle T^* z_0, f \rangle = \langle z, f \rangle$ and $\langle f_0, x \rangle = \langle f, x \rangle$ (since f extends f_0). So $\langle z_0, f_0 \rangle = \langle f_0, x \rangle$ for all $f_0 \in X_0^*$. □

Theorem 5.27 (Eberlein-Šmulian)**.** *If X is a reflexive Banach space, then the closed unit ball $\bar{B}(0; 1) := \{x \in X : \|x\| \leq 1\}$ is weakly sequentially compact.*

Proof. We first show that any bounded sequence $\{x_k\}$ in X has a weakly convergent subsequence. Let $X_0 := \overline{\text{span}(\{x_k\})}$. Since X_0 is a closed linear subspace of X, we know X_0 is reflexive, i.e., $X_0^{**} = X_0$, by Theorem 5.26 (Pettis). Since X_0 is separable, we know X_0^{**} is separable. By Theorem 5.25 (Banach), X_0^* is separable.

We consider $g_k \in X_0^{**}$, such that $\langle g_k, f_0 \rangle = \langle f_0, x_k \rangle$ for all $f_0 \in X_0^*$ (i.e., $x_k \mapsto g_k$ is the canonical mapping), then we know $\{g_k\}$ is bounded in X_0^{**} since $\{x_k\}$ is bounded in X_0. Thus, by Theorem 5.24, we know $\{g_k\}$ has a weak* convergent subsequence, i.e., there exist a subsequence $\{g_{k_j}\}$ and $g \in X_0^{**}$

such that $\lim_j \langle g_{k_j}, f_0\rangle = \langle g, f_0\rangle$ for all $f_0 \in X_0^*$. Let $x_0 \in X_0$ be such that $\langle g, f_0\rangle = \langle f_0, x_0\rangle$ for all $f_0 \in X_0^*$ (i.e., g is the canonical map of x_0). Thus we have

$$\lim_{j\to\infty} \langle f_0, x_{k_j}\rangle = \lim_{j\to\infty} \langle g_{k_j}, f_0\rangle = \langle g, f_0\rangle = \langle f_0, x_0\rangle$$

for all $f_0 \in X_0^*$.

We also need to show that $\langle f, x_{k_j}\rangle \to \langle f, x_0\rangle$ for all $f \in X^*$. To this end, for any $f \in X^*$, let $T : X^* \to X_0^*$ be defined by $Tf := f|_{X_0^*}$, i.e., Tf is the restriction of f onto X_0^*. Then $\langle f, x_{k_j}\rangle = \langle Tf, x_{k_j}\rangle \to \langle Tf, x_0\rangle = \langle f, x_0\rangle$. Therefore $x_{k_j} \rightharpoonup x$. To this point, we have showed that any bounded sequence in X has at least one weakly convergent subsequence.

Now we show that if $\{x_k\} \subset \bar{B}(0; 1) \subset X$ and x_0 is a weak accumulation point (the limit of a weakly convergent subsequence of $\{x_k\}$), then $x \in \bar{B}(0; 1)$ (we assume $x_0 \neq 0$ otherwise there is already $x_0 \in \bar{B}(0; 1)$). To this end, we know by Corollary 4.4 that there exists $f \in X^*$ such that $f(x_0) = \|x_0\|$ and $\|f\| = 1$. Hence

$$\|x_0\| = f(x_0) = \lim_{j\to\infty} f(x_{k_j}) \le \|f\| \lim_{j\to\infty} \|x_{k_j}\| \le \|f\| = 1.$$

Therefore $x_0 \in \bar{B}(0; 1)$, which completes the proof. □

Remark. Theorem 5.27 (Eberlein-Šmulian) implies that any bounded sequence has a weakly convergent subsequence. The converse of the theorem also holds.

Theorem 5.28 (Alaoglu)**.** *Let X be a B^* space. Then $\bar{B}^*(0; 1) := \{f \in X^* : \|f\| \le 1\}$ is weak* sequentially compact.*

Proof. First of all, note that, for any $x \in X$ and $f \in \bar{B}_1^* := \bar{B}^*(0; 1)$, there is $|f(x)| \le \|f\|\|x\| \le \|x\|$.

Denote $I_x := [-\|x\|, \|x\|] \subset \mathbb{R}$, and let $P := \prod_{x\in X} I_x$ be the product space equipped with the topology such that $p^{(k)} \to p$ in P if $p_x^{(k)} \to p_x$ for each $x \in X$, where $p^{(k)} = (p_x^{(k)})_{x\in X},\ p = (p_x)_{x\in X} \in P$. Moreover, by Tychonoff theorem, P is compact.

Let $\bar{B}_1^*$ be equipped with the weak* topology. Define $F : \bar{B}_1^* \to P$ by $F(f) := (f(x))_{x\in X} \in P$ for any $f \in \bar{B}_1^*$. Then it is clear that F is injective and continuous under the topologies specified above. Now we need to show $F(\bar{B}_1^*)$ is compact. Since P is compact, it suffices to show that $F(\bar{B}_1^*)$ is closed.

Let $p \in \overline{F(\bar{B}_1^*)} \subset P$ be arbitrary. Then let $f : X \to \mathbb{R}$ be defined by $f(x) := p_x$ for every $x \in X$ and $f_k \in \bar{B}_1^*$ be such that $F(f_k) \to p$ in P. Now we just need to show that $f \in \bar{B}_1^*$. To see this, for any $x, y \in X$, we have $f_k(x+y) = f_k(x) + f_k(y) \to p_x + p_y = f(x) + f(y)$ and $f_k(x + y) \to p_{x+y} = f(x + y)$ as $k \to \infty$. Hence $f(x + y) = f(x) + f(y)$. Similarly, we can show $f(cx) = cf(x)$ for all $c \in \mathbb{R}$ and $x \in X$. Therefore f is linear. Moreover, $|f(x)| = |p_x| \le \|x\|$. Hence $f \in \bar{B}_1^*$. □

Remark. Theorem 5.28 (Alaoglu) implies that any bounded sequence in X^* has a weak* convergent subsequence.

6 Spectral Theory of Linear Operators

6.1 Basics of Banach algebra

Definition 6.1 (Banach algebra)**.** A *normed algebra* $\mathcal{L}$ with norm $\|\cdot\|$ is a unital associative algebra over $\mathbb{C}$ with unit I such that $\|I\| = 1$ and $\|AB\| \le \|A\|\|B\|$ for all $A, B \in \mathcal{L}$. If $\mathcal{L}$ is complete under $\|\cdot\|$, then $\mathcal{L}$ is called a *Banach algebra*.

Example 6.2. Let X be a Banach space, then $\mathcal{L} := L(X)$ with the operator norm $\|\cdot\|$ is a Banach algebra.

Definition 6.3 (Inverse)**.** An element T of a Banach algebra $\mathcal{L}$ is said to be *invertible* if there exists $S \in \mathcal{L}$, often denoted by T^{-1}, such that $ST = TS = I$. We say T has *left inverse* (resp. *right inverse*) if there exists $A \in \mathcal{L}$ (resp. $B \in \mathcal{L}$) such that $AT = I$ (resp. $TB = I$). Note that if T has both left and right inverses, then T is invertible, and the left and right inverses are identical:

$$A = AI = A(TB) = (AT)B = IB = B,$$

which is just T^{-1}. It is also clear that, if T and U are both invertible, then TU is invertible and $(TU)^{-1} = U^{-1}T^{-1}$.

Theorem 6.4. *If T and S commute, i.e., $TS = ST$, and TS is invertible, then T and S are invertible.*

Proof. Let U be the inverse of TS, then $T(SU) = (TS)U = I$ and $(US)T = U(ST) = U(TS) = I$. Hence SU is the right inverse of T and US is the left inverse of T. Therefore T is invertible. Similar for S. □

Theorem 6.5. *Suppose $T \in \mathcal{L}$ is invertible, then $T - A$ is invertible for any $A \in B(0; 1/\|T^{-1}\|)$.*

Proof. 1. We first consider the case of $T = I$. We claim that $I - B$ is invertible for any B satisfying $\|B\| < 1$, and moreover $(I - B)^{-1} = \sum_{k=0}^{\infty} B^k$ (called the *Neumann series of* B). To see this, notice that

$$\Big\| \sum_{i=k+1}^{k+p} B^i \Big\| \le \sum_{i=k+1}^{k+p} \|B\|^i \le \sum_{i=k+1}^{\infty} \|B\|^i \to 0,$$

as $k \to \infty$ for any $p \in \mathbb{N}$. Hence $S := \sum_{k=0}^{\infty} B^k$ converges and $S \in \mathcal{L}$. Therefore

$$BS = B \sum_{k=0}^{\infty} B^k = \sum_{k=1}^{\infty} B^k = S - I,$$

which implies that $(I - B)S = I$. Similarly, we can show $S(I - B) = I$. Hence the claim holds.

2. Now we consider the general case where T is invertible. Consider $T - A = T(I - T^{-1}A)$, then $\|T^{-1}A\| \le \|T^{-1}\|\|A\| < 1$ for all A satisfying $\|A\| < 1/\|T^{-1}\|$. Hence we know $I - T^{-1}A$ is invertible from above and thus $T - A = T(I - T^{-1}A)$ is invertible by noticing that T is invertible. □

Remark. Theorem 6.5 implies that, if T is invertible, then so are all elements in the open neighborhood $B(T; 1/\|T^{-1}\|)$ of T.

6.2 Decomposition of spectrum

Definition 6.6 (Spectrum of linear operator)**.** Let X be a (complex) Banach space, $D \subseteq X$, and $T : D \to X$ a closed linear operator. Then we define

- *Point spectrum* $\sigma_p(T) := \{\lambda \in \mathbb{C} : \lambda I - T \text{ is not invertible}\}$. We call $\lambda \in \sigma_p(T)$ an *eigenvalue of* T.
- *Resolvent set* $\rho(T) := \{\lambda \in \mathbb{C} : \lambda I - T : D \to X \text{ is bijective}\}$. We call $\lambda \in \rho(T)$ a *regular value of* T.
- *Continuous spectrum* $\sigma_c(T) := \{\lambda \in \mathbb{C} : \lambda I - T \text{ is injective}, (\lambda I - T)(D) \subsetneq X, \overline{(\lambda I - T)(D)} = X\}$.
- *Residual spectrum* $\sigma_c(T) := \{\lambda \in \mathbb{C} : \lambda I - T \text{ is injective}, \overline{(\lambda I - T)(D)} \subsetneq X\}$.

Then we have the disjoint union

$$\mathbb{C} = \rho(T) \cup \underbrace{\sigma_p(T) \cup \sigma_c(T) \cup \sigma_r(T)}_{=:\sigma(T)},$$

where $\sigma(T)$ is called the *spectrum of* T.

Remark. Note that $\lambda \in \sigma_p(T)$ iff $(\lambda I - T)^{-1}$ does not exist iff $\lambda I - T$ is not injective iff there exists nonzero $x \in X$ such that $Tx = \lambda x$.

Remark. Note that if $\lambda \in \rho(T)$ then $(\lambda I - T)^{-1} : X \to D$ is bijective. Moreover, we can show that $(\lambda I - T)^{-1}$ is closed. To see this, let $y_k \to y$ and $x_k := (\lambda I - T)^{-1} y_k \to x$, then we have $\lambda x_k - Tx_k = y_k \to y$, which thus implies $Tx_k \to \lambda x - y$. Since T is closed, we know $Tx = \lambda x - y$, i.e., $(\lambda I - T)^{-1} y = x$, which proves our claim.

From the arguments above, we can also see that D is closed and hence is complete. Moreover, since X is closed, by Theorem 3.36 (Closed graph), we know $(\lambda I - T)^{-1} \in L(X, D)$. By Theorem 3.29, we know $\lambda I - T \in L(D, X)$.

Remark. If $\dim(X) < \infty$, then $\mathbb{C} = \rho(T) \cup \sigma_p(T)$. Otherwise, $\sigma_c(T)$ and $\sigma_r(T)$ may be nonempty.

Example 6.7 (Point spectrum)**.** Let $X = L^2([0,1])$, $D = \{x \in C^2([0,1]) : x(0) = x(1) = 0,\ x'(0) = x'(1) = 0\}$, and $T : D \to X$ be defined by $(Tx)(t) = -x''(t)$ for all $t \in [0,1]$ and $x \in D$. Then T is a closed linear operator, and $\sigma(T) = \sigma_p(T) = \Lambda := \{(2k\pi)^2 : k \in \mathbb{N} \cup \{0\}\}$.

Proof. For any $k \in \mathbb{N} \cap \{0\}$, let $x(t) = \cos(2\pi kt)$ which is nonzero. Then $(Tx)(t) = -x''(t) = (2k\pi)^2 x(t)$. Hence $\Lambda \subseteq \sigma_p(T)$.

If $\lambda \in \mathbb{C} \setminus \Lambda$, then for any $f \in X$, the equation $(T - \lambda I)x = f$ has a unique solution:

$$x(t) = \sum_{k \in \mathbb{Z}} \frac{c_k}{(2k\pi)^2 - \lambda^2} e^{2k\pi t \iota}, \quad \text{where} \quad c_k = \int_0^1 f(t) e^{-2k\pi t \iota} \, \mathrm{d}t.$$

(To see this, apply Fourier transform to both sides and solve for x, then c_k are the Fourier coefficients of f under the Fourier basis $\{e^{2k\pi t\imath} : k \in \mathbb{Z}\}$ of X.) Also note that

$$\|x\|^2 = \sum_{k\in\mathbb{Z}} \frac{|c_k|^2}{|(2k\pi)^2 - \lambda^2|^2} \le M_\lambda \sum_{k\in\mathbb{Z}} |c_k|^2 = M_\lambda \|f\|^2,$$

where $M_\lambda := \sup_{k\in\mathbb{Z}} |(2k\pi)^2 - \lambda^2|^{-2} < \infty$. Therefore $\lambda \in \rho(T)$. Hence $\Lambda = \sigma_p(T) = \sigma(T)$. □

Example 6.8 (Residual spectrum)**.** Let $X = C([0,1])$ and $T : X \to X$ be defined by $(Tx)(t) := tx(t)$ for any $t \in [0,1]$ and $x \in X$. Then $\sigma(T) = \sigma_r(T) = [0,1]$.

Proof. If $\lambda \in \mathbb{C} \setminus [0,1]$, then $[(\lambda I - T)^{-1}x](t) = \frac{x(t)}{\lambda - t}$, and $\|(\lambda I - T)^{-1}x\| \le (\sup_{0\le t\le 1} |\lambda - t|^{-1})\|x\|$. Hence $\lambda I - T$ is bijective and $\lambda \in \rho(T)$.

If $\lambda \in [0,1]$, then $(\lambda I - T)x = 0$ has a unique solution $x = 0$ and hence $\lambda I - T$ is injective. Moreover, for any $y \in (\lambda I - T)(X)$, there exists $x \in X$, such that $y(t) = (\lambda - t)x(t)$ for all $t \in [0,1]$. Hence $y(\lambda) = 0$. This implies that all nonzero functions in X must not be in $\overline{(\lambda I - T)(X)}$ (e.g., $1 \notin \overline{(\lambda I - T)(X)}$). Hence $\lambda \in \sigma_r(T) \subseteq \sigma(T) \subseteq [0,1]$. Therefore $\sigma(T) = \sigma_r(T) = [0,1]$. □

Example 6.9 (Continuous spectrum)**.** Let $X = L^2([0,1])$ and $T : X \to X$ be defined by $(Tx)(t) := tx(t)$ for any $t \in [0,1]$ and $x \in X$. Then $\sigma(T) = \sigma_c(T) = [0,1]$.

Proof. Similar as in the previous example, if $\lambda \in \mathbb{C}\setminus[0,1]$, then $[(\lambda I - T)^{-1}x](t) = \frac{x(t)}{\lambda - t}$, and $\|(\lambda I - T)^{-1}x\| \le (\sup_{0\le t\le 1} |\lambda - t|^{-1})\|x\|$. Hence $\lambda \in \rho(T)$.

If $\lambda \in [0,1]$, then we claim that $1 \notin (\lambda I - T)(X)$: if there exists $x \in X$ such that $(\lambda I - T)x = 1$, then $x(t) = \frac{1}{\lambda - t} \notin X$, which is a contradiction. On the other hand, for any $y \in X$, let $E_k = (\lambda - \frac{1}{k}, \lambda + \frac{1}{k}) \cap [0,1]$ and define $x_k(t) := \frac{1}{\lambda - t} y(t)\chi_{E_k^c}(t)$. Then

$$\|(\lambda I - T)(x_k) - y\|^2 = \int_0^1 |y(t)\chi_{E_k}(t)|^2 \,\mathrm{d}t = \int_{\lambda - \frac{1}{k}}^{\lambda + \frac{1}{k}} |y(t)|^2 \,\mathrm{d}t \to 0$$

as $k \to \infty$. Hence $y \in \overline{(\lambda I - T)(X)}$. Hence $\lambda \in \sigma_c(T) \subseteq \sigma(T) \subseteq [0,1]$. Therefore $\sigma(T) = \sigma_c(T) = [0,1]$. □

6.3 Gelfand theorem

Definition 6.10 (Resolvent)**.** Let X be a Banach space and $T : X \to X$ be a closed linear operator. Then the *resolvent of* T is the mapping $R_T : \rho(T) \to L(X)$ defined by $R_T(\lambda) := (\lambda I - T)^{-1}$ for every $\lambda \in \rho(T)$.

Lemma 6.11. *Suppose* $T \in L(X)$ *and* $\|T\| < 1$*, then* $(I - T)^{-1} \in L(X)$ *and* $\|(I - T)^{-1}\| \le 1/(1 - \|T\|)$.

Proof. By Theorem 6.5 we know $(I-T)^{-1} \in L(X)$. For any $y \in X$, let $x = (I-T)^{-1}y$, then

$$\|y\| = \|(I-T)x\| \geq \|x\| - \|Tx\| \geq \|x\| - \|T\|\|x\| = (1-\|T\|)\|x\|.$$

Hence $\|(I-T)^{-1}y\| = \|x\| \leq \frac{\|y\|}{1-\|T\|}$. Therefore $\|(I-T)^{-1}\| \leq \frac{1}{1-\|T\|}$. □

Remark. Recall that $(I-T)^{-1} = \sum_{k=0}^{\infty} T^k$, and hence $\|(I-T)^{-1}\| \leq \sum_{k=0}^{\infty} \|T\|^k = \frac{1}{1-\|T\|}$. We can also prove the lemma above using Banach contractive mapping theorem.

Corollary 6.12 ($\rho(T)$ is open in $\mathbb{C}$)**.** *Let X be a Banach space. If $T : X \to X$ is a closed linear operator, then $\rho(T)$ is open in $\mathbb{C}$.*

Proof. Let $\lambda_0 \in \rho(T)$ be arbitrary. Then $(\lambda_0 I - T)^{-1} \in L(X)$. Note also that

$$\lambda I - T = (\lambda - \lambda_0)I + (\lambda_0 I - T) = (\lambda_0 I - T)(I + (\lambda - \lambda_0)(\lambda_0 I - T)^{-1}).$$

For any $\lambda \in \mathbb{C}$ such that $|\lambda - \lambda_0| < 1/\|(\lambda_0 I - T)^{-1}\|$, we have $\|(\lambda - \lambda_0)(\lambda_0 I - T)^{-1}\| < 1$. Hence by Lemma 6.11 we know $U := I + (\lambda - \lambda_0)(\lambda_0 I - T)^{-1}$ is invertible. Thus $(\lambda I - T)^{-1} = U^{-1}(\lambda_0 I - T)^{-1} \in L(X)$. Therefore $\lambda \in \rho(T)$. This implies that $B(\lambda_0; 1/\|(\lambda_0 I - T)^{-1}\|) \subseteq \rho(T)$. Hence $\rho(T)$ is open. □

Lemma 6.13 (First resolvent identity)**.** *For any $\lambda, \mu \in \rho(T)$, there is $R_T(\lambda) - R_T(\mu) = (\mu - \lambda)R_T(\lambda)R_T(\mu)$.*

Proof. We have

$$\begin{aligned}(\lambda I - T)^{-1} &= (\lambda I - T)^{-1}(\mu I - T)(\mu I - T)^{-1} \\ &= (\lambda I - T)^{-1}[(\mu - \lambda)I + (\lambda I - T)](\lambda I - T)^{-1} \\ &= (\mu - \lambda)(\lambda I - T)^{-1}(\mu I - T)^{-1} + (\mu I - T)^{-1},\end{aligned}$$

which is the claimed identity. □

Theorem 6.14 (Analyticity of resolvent)**.** *The resolvent $R_T : \rho(T) \to L(X)$ is an analytic function.*

Proof. 1. We first show that R_T is continuous. Let $\lambda_0 \in \rho(T)$ be arbitrary. Then by Corollary 6.12, we know for any $\lambda \in C$ satisfying $|\lambda - \lambda_0| < (2\|R_T(\lambda_0)\|)^{-1}$, there is

$$\|R_T(\lambda)\| = \|R_T(\lambda_0)\|\|(I + (\lambda - \lambda_0)R_T(\lambda_0))^{-1}\| \leq 2\|R_T(\lambda_0)\|.$$

Then by Lemma 6.13, we know

$$\|R_T(\lambda) - R_T(\lambda_0)\| = \|R_T(\lambda_0)\|\|R_T(\lambda)\||\lambda - \lambda_0| \leq 2\|R_T(\lambda_0)\|^2|\lambda - \lambda_0|,$$

which implies that R_T is continuous at λ_0. As λ_0 is arbitrary, we know R_T is continuous in $\rho(T)$.

2. Now we show that R_T is analytic. By Lemma 6.13, we have

$$\lim_{\lambda\to\lambda_0}\frac{R_T(\lambda)-R_T(\lambda_0)}{\lambda-\lambda_0}=-\lim_{\lambda\to\lambda_0}R_T(\lambda_0)R_T(\lambda)=-R_T(\lambda_0)^2,$$

where the last equality is due to continuity of R_T at λ_0. Hence R_T is an analytic function in $\rho(T)$. $\square$

Remark. Differentiability in $\mathbb{C}$ implies analyticity. We can also prove the analyticity from another point of view: for any $\lambda_0\in\rho(T)$ and $h\in(0,1/\|R_T(\lambda_0)\|)$, we have

$$\begin{aligned}&R_T(\lambda_0-h)=((\lambda_0-h)I-T)^{-1}\\ &=\big((\lambda_0 I-T)[I-h(\lambda_0 I-T)^{-1}]\big)^{-1}=[I-h(\lambda_0 I-T)^{-1}]^{-1}(\lambda_0 I-T)^{-1}\\ &=\sum\nolimits_{k=0}^{\infty}[h(\lambda I-T)^{-1}]^k(\lambda I-T)^{-1}=\sum\nolimits_{k=0}^{\infty}R_T(\lambda_0)^{k+1}h^k.\end{aligned}$$

That is, R_T can be expanded as a convergent power series in an open neighborhood of λ_0, which means that R_T is an analytic function.

Theorem 6.15. *If $T\in L(X)$, then $\sigma(T)\neq\emptyset$.*

Proof. If not, then $\rho(T)=\mathbb{C}$ and thus $R_T:\mathbb{C}\to L(X)$ is analytic. Furthermore, for any positive number $\lambda_0>\|T\|$, we know $R_T(\lambda)=(\lambda I-T)^{-1}=\lambda^{-1}(I-\frac{T}{\lambda})^{-1}$ and therefore

$$\|R_T(\lambda)\|\le|\lambda|^{-1}\frac{1}{1-\|T/\lambda\|}=\frac{1}{|\lambda|-\|T\|}\le\frac{1}{\lambda_0-\|T\|}<\infty,$$

for any $|\lambda|\ge\lambda_0$. On the other hand, R_T is bounded in $\bar{B}(0;|\lambda_0|)$ since R_T is continuous. Therefore, R_T is bounded on $\mathbb{C}$. For any $f\in L(X)^*$, we define $w_f:\mathbb{C}\to\mathbb{C}$ by $w_f(\lambda):=f(R_T(\lambda))$. Then w_f is analytic and bounded on $\mathbb{C}$. By Liouville theorem, w_f must be a constant which depends on f but not λ. By Corollary 4.3, we know R_T is constant (if not, then $R_T(\lambda_1)=R_T(\lambda_2)$ for some $\lambda_1\neq\lambda_2$, and they can be distinguished by some $f\in L(X)^*$ due to Corollary 4.3). Therefore, by Lemma 6.13, we have $R_T(\lambda)R_T(\mu)=0\in L(X)$ for any $\lambda,\mu\in\mathbb{C}$, which contradicts to $R_T(\lambda)$ being invertible for any $\lambda\in\mathbb{C}$. $\square$

Remark. From the proof above, we can see that, if $|\lambda|\ge\|T\|$, then $R_T(\lambda)\in L(X)$ and hence $\lambda\in\rho(T)$. Therefore $\mathbb{C}\setminus B(0;\|T\|)\subseteq\rho(T)$ and hence $\sigma(T)\subseteq\bar{B}(0;\|T\|)$, which means that $\sigma(T)$ is bounded. Since $\rho(T)$ is open, we know $\sigma(T)$ is closed and thus compact in $\mathbb{C}$. Next we want to obtain a tight bound of $\sigma(T)$.

Definition 6.16 (Spectral radius)**.** Let X be a Banach space and $T\in L(X)$, then $r_\sigma(T):=\sup\{|\lambda|:\lambda\in\sigma(T)\}$ is called the *spectral radius of* T. Note that $r_\sigma(T)\le\|T\|$ from the remark above.

Theorem 6.17 (Gelfand)**.** *Let X be a Banach space and $T\in L(X)$, then $r_\sigma(T)=\lim_{k\to\infty}\|T^k\|^{1/k}$.*

Proof. For notation simplicity, we denote $r := r_\sigma(T)$, $l := \liminf_{k\to\infty} \|T^k\|^{1/k}$, and $u := \limsup_{k\to\infty} \|T^k\|^{1/k}$. Then it suffices to show $l = r = u$.

1. We first show that $r \le u$. To this end, recall that $R_T(\lambda) = (\lambda I - T)^{-1} = \sum_{k=0}^{\infty} \frac{T^k}{\lambda^{k+1}}$ for $|\lambda| > \|T\|$. By Cauchy-Hadamard theorem (the radius of convergence of the power series $\sum_{k=0}^{\infty} c_k z^k$ is $(\limsup_k |c_k|^{1/k})^{-1}$), we know $R_T(\lambda) \in L(X)$ if $|\lambda|^{-1} < u^{-1}$ or $|\lambda| > u$. In this case, $\lambda \in \rho(T)$, and hence $r \le u$.

2. Next we show $r \ge u$. For any $\epsilon > 0$ and $f \in L(X)^*$, we define the mapping $w_f : \rho(T) \to \mathbb{C}$ by $w_f(\lambda) := f(R_T(\lambda))$. Therefore $w_f(r+\epsilon) = \sum_{k=0}^{\infty} f(T^k)/(r+\epsilon)^{k+1} < \infty$, which implies that

$$\Big|\frac{f(T^k)}{(r+\epsilon)^{k+1}}\Big| = \Big|f\left(\frac{T^k}{(r+\epsilon)^{k+1}}\right)\Big| = \Big|\Big\langle \frac{T^k}{(r+\epsilon)^{k+1}}, f\Big\rangle\Big| < \infty,$$

where we identify $T^k/(r+\epsilon)^{k+1}$ with its canonical image in $L(X)^{**}$. Then by Theorem 3.37 (Uniform boundedness), we know there exists $M > 0$ such that

$$\Big\|\frac{T^k}{(r+\epsilon)^{k+1}}\Big\| = \frac{\|T^k\|}{(r+\epsilon)^{k+1}} \le M, \quad \forall\, k \in \mathbb{N}.$$

Therefore $r+\epsilon \ge \limsup_k \|T^k\|^{1/k} = u$. Since $\epsilon > 0$ is arbitrary, we know $r \ge u$.

3. Now we only need to show $r \le l$. Note that, for any $k \in \mathbb{N}$, there is

$$\lambda^k I - T^k = (\lambda I - T)P_k(\lambda; T) = P_k(\lambda; T)(\lambda I - T),$$

where $P_k(\lambda; T) := \sum_{j=1}^{k} \lambda^{j-1} T^{k-j}$. Hence, if $\lambda^k \in \rho(T^k)$, then $(\lambda^k I - T^k)^{-1} \in L(X)$. By Theorem 6.4, we know $(\lambda I - T)^{-1} \in L(X)$ and hence $\lambda \in \rho(T)$. Therefore, if $\lambda \in \sigma(T)$, then $\lambda^k \in \sigma(T^k)$ and hence $|\lambda^k| \le \|T^k\|$. This implies that $|\lambda| \le \|T^k\|^{1/k}$ for all $k \in \mathbb{N}$, and thus $|\lambda| \le \liminf_k \|T^k\|^{1/k} = l$. Therefore $r \le l$. □

Example 6.18 (Spectrum decomposition)**.** Let $X = l^2$ and $T \in L(X)$ be defined by $Tx := (0, x_1, x_2, \dots)$ for any $x = (x_1, x_2, \dots) \in X$. Namely, T is the right shift operator. Then $\sigma_p(T) = \emptyset$, $\sigma_c(T) = \{\lambda \in \mathbb{C} : |\lambda| = 1\}$, and $\sigma_r(T) = \{\lambda \in \mathbb{C} : |\lambda| < 1\}$.

Proof. Note that $\|T\| = 1$, and therefore $\sigma(T) \subseteq \bar{B}(0; 1)$. Therefore, we only need to check λ satisfying $|\lambda| \le 1$.

1. We first check the case where $|\lambda| < 1$. Specifically, we claim that, if $|\lambda| < 1$, then $(\lambda I - T)(X) = \operatorname{span}(z)^\perp$ where $z := (1, \bar{\lambda}, \bar{\lambda}^2, \dots) \in X$.

Suppose $y = (y_1, y_2, \dots) \in (\lambda I - T)(X)$, then there exists $x = (x_1, x_2, \dots) \in X$ such that $y = (\lambda I - T)x$, i.e., $y_k = \lambda x_k - x_{k-1}$ for all $k \in \mathbb{N}$ (we define $x_0 = 0$ for convenience). Notice that

$$\begin{aligned}\sum_{i=1}^{k} \lambda^{i-1} y_i &= y_1 + \lambda y_1 + \cdots + \lambda^{k-1} y_k \\ &= \lambda x_1 + \lambda(\lambda x_2 - x_1) + \cdots + \lambda^{k-1}(\lambda x_k - x_{k-1}) \\ &= \lambda^k x_k \to 0,\end{aligned}$$

as $k \to \infty$ (because $x \in X$ and hence x_k are bounded). Therefore we know

$$\langle y, z\rangle = \sum_{k=1}^{\infty} y_k \overline{z_k} = \sum_{k=1}^{\infty} \lambda^{k-1} y_k = 0.$$

Therefore $y \perp z$. Hence $(\lambda I - T)(X) \subset \text{span}(z)^{\perp}$.

Now suppose $y \perp z$. Then we need to show that $y \in (\lambda I - T)(X)$. To this end, let $x = (x_1, x_2, \dots)$ be such that $x_k = -\sum_{j=0}^{\infty} \lambda^j y_{k+j-1}$ for each $k \in \mathbb{N}$. If $\lambda = 0$, then obviously $x = -Ty = (\lambda I - T)y$. If $\lambda \neq 0$, then we have

$$|x_k|^2 = \Big| \sum_{j=0}^{\infty} \lambda^j y_{k+j+1} \Big|^2 \leq \Big(\sum_{j=0}^{\infty} |\lambda|^j \Big) \Big(\sum_{j=0}^{\infty} |\lambda|^j |y_{k+j+1}|^2 \Big)$$

where we used Cauchy-Schwarz inequality. Therefore, we obtain

$$\sum_{k=1}^{\infty} |x_k|^2 \leq \Big(\sum_{j=0}^{\infty} |\lambda|^j \Big) \Big(\sum_{k=1}^{\infty} \sum_{j=0}^{\infty} |\lambda|^j |y_{k+j+1}|^2 \Big) \leq \Big(\sum_{j=0}^{\infty} |\lambda|^j \Big)^2 \|y\|^2 < \infty$$

where we exchanged the summations in k and j to deduce the second inequality. Therefore $x \in X$. Similar as above, we know $\sum_{k=1}^{\infty} \lambda^{k-1} y_k = 0$ since $y \in X$. Hence

$$\begin{aligned} x_k &= -\sum_{j=0}^{\infty} \lambda^j y_{k+j+1} = -\lambda^{-k} \sum_{j=0}^{\infty} \lambda^{k+j} y_{k+j+1} = -\lambda^{-k} \sum_{j=k+1}^{\infty} \lambda^{j-1} y_j \\ &= -\lambda^{-k} \Big(0 - \sum_{j=1}^{k} \lambda^{j-1} y_j \Big) = \sum_{j=1}^{k} \lambda^{-k+j-1} y_j = \sum_{j=1}^{k} \lambda^{-j} y_{k-j+1}. \end{aligned}$$

Therefore $y_k = \lambda x_k - x_{k-1}$ for all $k \in \mathbb{N}$, which implies $y = (\lambda I - T)x$. Hence $\text{span}(z)^{\perp} \subseteq (\lambda I - T)(X)$.

In conclusion, we have $(\lambda I - T)(X) = \text{span}(z)^{\perp}$. Therefore $\overline{(\lambda I - T)(X)} \subsetneq X$. Thus, $\{\lambda \in \mathbb{C} : |\lambda| < 1\} \subset \sigma_r(T)$.

2. Now we check the case where $|\lambda| = 1$. We first consider the case $\lambda = 1$. If $y = (I - T)(X)$, then there exists $x \in X$ such that $y_k = x_k - x_{k-1}$ for all $k \in \mathbb{N}$. Therefore $x_k = \sum_{i=1}^{k} y_i$, and thus we know

$$(I - T)(X) = \Big\{ y \in X : \sum_{k=1}^{\infty} \Big| \sum_{i=1}^{k} y_i \Big|^2 < \infty \Big\} \subsetneq X.$$

Now we shall show $\overline{(I - T)(X)} = X$. For any $\xi = (\xi_1, \xi_2, \dots) \in X$ and any $\epsilon > 0$, we know there exists $K = K(\xi, \epsilon) \in \mathbb{N}$ such that $\sum_{k=K+1}^{\infty} |\xi_k|^2 < \epsilon^2/6$. Denote $c := \sum_{k=1}^{K} \xi_k$ and choose $m \in \mathbb{N}$ sufficiently large such that $\frac{|c|^2}{m} < \frac{\epsilon^2}{6}$. Define $y = (y_1, y_2, \dots)$ where

$$y_k = \begin{cases} \xi_k, & \text{if} \quad 1 \leq k \leq K, \\ -c/m, & \text{if} \quad K+1 \leq k \leq K+m, \\ 0, & \text{if} \quad K+m < k. \end{cases}$$

Then there is

$$\sum_{k=1}^{\infty}\Big|\sum_{i=1}^{k} y_i\Big|^2 = \sum_{k=1}^{K}\Big|\sum_{i=1}^{k}\xi_i\Big|^2 + \sum_{k=K+1}^{K+m}\Big|c-(k-K)\frac{c}{m}\Big|^2 < \infty,$$

which means that $y \in X$. Moreover,

$$\begin{aligned}\|\xi - y\|^2 &= \sum_{k=1}^{\infty}|\xi_k - y_k|^2 = \sum_{k=K+1}^{K+m}\Big|\xi_k - \frac{c}{m}\Big|^2 + \sum_{k=K+m+1}^{\infty}|\xi_k|^2 \\ &\le \sum_{k=K+1}^{K+m}\Big(2|\xi_k|^2 + 2\Big|\frac{c}{m}\Big|^2\Big) + \sum_{k=K+m+1}^{\infty}|\xi_k|^2 \\ &\le \frac{2c^2}{m} + 2\sum_{k=K+1}^{\infty}|\xi_k|^2 < \epsilon^2,\end{aligned}$$

where we used $|a-b|^2 \le 2|a|^2 + 2|b|^2$. Hence $\overline{(I-T)(X)} = X$.

For general λ with $|\lambda| = 1$, we can convert it to the case of $\lambda = 1$. To see this, notice that

$$\begin{aligned}\eta = (\lambda I - T)\xi \quad &\Longleftrightarrow \quad \eta_k = \lambda\xi_k - \xi_{k-1} \\ &\Longleftrightarrow \quad \lambda^{k-1}\eta_k = \lambda^k\xi_k - \lambda^{k-1}\xi_{k-1} \\ &\Longleftrightarrow \quad y_k = x_k - x_{k-1} \\ &\Longleftrightarrow \quad y = (I-T)x,\end{aligned}$$

where $x_k := \lambda^k\xi_k$ and $y_k := \lambda^{k-1}\eta_k$ for all $k \in \mathbb{N}$. Therefore the proof reduces to the case with $\lambda = 1$. In conclusion, we have $\{\lambda \in \mathbb{C} : |\lambda| = 1\} \subset \sigma_c(T)$.

3. Combining the results above and that $\sigma(T) = \sigma_p(T) \cup \sigma_c(T) \cup \sigma_r(T) \subset \bar{B}(0;1)$, we know $\sigma_c(T) = \{\lambda \in \mathbb{C} : |\lambda| = 1\}$, and $\sigma_r(T) = \{\lambda \in \mathbb{C} : |\lambda| < 1\}$, and $\sigma_p(T) = \emptyset$. The spectral radius is $r_\sigma(T) = 1$. Furthermore, $\rho(T) = \mathbb{C} \setminus \sigma(T) = \{\lambda \in \mathbb{C} : |\lambda| > 1\}$. □