Functional Analysis Notes
Fall 2004
Prof. Sylvia Serfaty

Yevgeny Vilensky
Courant Institute of Mathematical Sciences
New York University

March 14, 2006

Preface

These are notes from a one-semester graduate course in Functional Analysis given by Prof. Sylvia Serfaty at the Courant Institute of Mathematical Sciences, New York University, in the Fall of 2004. Thanks to Atilla Yilmaz, Caroline Müller, and Alexey Kuptsov for providing their course notes to help with the preparation of this typed version. The course was largely based on Haim Brezis' *Analyse fonctionnelle : thorie et applications*, and Michael Reed's and Barry Simon's *Methods of modern mathematical physics* vol. 1. The geometric versions of the Hahn-Banach Theorems were taken almost entirely out of Brezis and the section on Spectral Theory was based entirely on Reed and Simon.

Contents

Chapter 1

Hahn-Banach Theorems and Introduction to Convex Conjugation

1.1 Hahn-Banach Theorem - Analytic Form

1.1.1 Theorems on Extension of Linear Functionals

The Hahn-Banach Theorem concerns extensions of linear functionals from a subspace of a linear space to the entire space.

Theorem 1.1.1 (Real Version of the Hahn-Banach Theorem) *Let X be a* real *linear space and let $p : X \to \mathbb{R}$ be a function satisfying:*

$$p(tx) = t \cdot p(x), \quad p(x + y) \leq p(x) + p(y)$$

for all $t > 0, x, y \in X$. Let $f : Y \to \mathbb{R}$ be linear with $Y \subset X$ such that $f(x) \leq p(x)$ for all $x \in X$. Then, $\exists$ a linear map $\Lambda : X \to \mathbb{R}$ such that for $y \in Y, \Lambda(y) = f(y)$ and $\Lambda(x) \leq p(x)$ for all $x \in X$.

Before beginning the proof of the theorem, we need some definitions and a reminder of Zorn's Lemma.

Definition Let P be a set with a partial order relation "$\prec$". $Q \subset P$ is said to be *totally ordered* if $\forall a, b \in Q$ we have $a \prec b$ or $b \prec a$. c is an *upper bound* for Q if $a \in Q \Rightarrow a \prec c$. m is called a *maximal element* in Q if and only if $\forall a \in Q$ we have that if $m \prec a$ then $a = m$.

Lemma 1.1.2 (Zorn's Lemma) *Let P be a non-empty set with a partial ordering, such that every totally ordered subset of P admits an upper bound. Then, P has a maximal element.*

Proof of Real Version of the Hahn-Banach Theorem Let P be the collection of linear functions h, defined on their domain, $D(h) \supset Y$, that extend f and that satisfy:

$$\begin{aligned} h(y) &= f(y) \quad \forall y \in Y \\ h(x) &\leq p(x) \quad \forall x \in X. \end{aligned}$$

We now define a partial ordering on the set P, so that we can apply Zorn's Lemma. In P, we say $h_1 \prec h_2$ if and only if $D(h_1) \subset D(h_2)$ and $h_1 = h_2$ in $D(h_1)$.

Certainly, P is nonempty (because it at least contains f). Now, let $(h_\alpha)_{\alpha \in A}$ be a totally ordered subset of P. Let h be defined on $\bigcup_{\alpha \in A} D(h_\alpha)$ and let $h(x) = h_\alpha(x)$ if $x \in D(h_\alpha)$. This is well-defined because $(h_\alpha)_{\alpha \in A}$ is a totally-ordered set (and so, all h_α agree on the intersection). By our definition of $\prec$, it follows that h is an upper bound.

So, applying Zorn's Lemma to $(P, \prec)$, we see that P has a maximal element. Call this element Λ. We just need to check that $D(\Lambda) = X$.

Suppose that $D(\Lambda) \neq X$. Then, let $x_0 \notin D(\Lambda)$. Then, we claim that there is an a so that we can extend Λ to $h : D(\Lambda) \oplus \mathbb{R}x_0 \to \mathbb{R}$ by:

$$\begin{aligned} h(x + tx_0) &= \Lambda(x) + t \cdot a \\ &\text{and} \\ \Lambda(x) + t \cdot a &\leq p(x + tx_0) \end{aligned}$$

for all $x \in D(\Lambda)$ and $t \in \mathbb{R}$.

$$\Leftrightarrow \begin{cases} \Lambda(x) + a \leq p(x + x_0) \\ \Lambda(x) - a \leq p(x - x_0) \end{cases}$$

For all $x \in D(\Lambda)$ (just replace x by $\frac{x}{t}$ if $t > 0$ and $-\frac{x}{t}$ if $t < 0$). So, is there such an a? It is enough to check that:

$$\sup_{x \in D(\Lambda)} \Lambda(x) - p(x - x_0) \leq \inf_{y \in D(\Lambda)} p(y + x_0) - \Lambda(y)$$

To show this, note that by the linearity of Λ we have

$$\begin{aligned} \Lambda(x) + \Lambda(y) &= \Lambda(x + y) \\ &= \Lambda(x - x_0 + x_0 + y) \leq p(x - x_0 + x_0 + y) \\ &\leq p(x - x_0) + p(x_0 + y) \end{aligned}$$

The last inequality being true, by the subadditivity of p.

$$\Rightarrow \Lambda(x) - p(x - x_0) \leq p(x_0 + y) - \Lambda(y)$$

for all x, y. Hence, $\sup_x \text{LHS} \leq \inf_y \text{RHS}$. Hence, it is possible to choose an a so we can extend Λ to h such that $h(x+tx_0) = \Lambda(x)+ta$ and $\Lambda(x)+ta \leq p(x+tx_0)$. But this contradicts the fact that Λ was the maximal element. ∎

Theorem 1.1.3 (Complex Version of Hahn-Banach Theorem) *Let X be a complex linear space, $p : X \to \mathbb{R}$ a map such that:*

$$p(\alpha x + \beta y) \leq |\alpha| p(x) + |\beta| p(y) \quad \text{and} \quad p(tx) = t \cdot p(x)$$

for all $x, y \in X, t > 0$, and $\alpha, \beta \in \mathbb{C}$ satisfying $|\alpha| + |\beta| = 1$. Let $f : Y \subset X \to \mathbb{C}$ be linear such that $|f(y)| \leq p(y)$ for all $y \in Y$.

Then, there exists a linear $\Lambda : X \to \mathbb{C}$ such that $\Lambda(y) = f(y)$ for $y \in Y$ and $|\Lambda(x)| \leq p(x)$ for all $x \in X$.

Proof We want to reduce this to the real case. Let $l(x) = \Re f(x)$. Since $f(ix) = if(x)$, we have that $l(ix) = \Re f(ix) = \Re if(x) = -\Im f(x)$. So that, $f(x) = l(x) - il(x)$.

Then, since for any $z \in \mathbb{C}, |\Re z| \leq |z|$, we get that $l(x) \leq |f(x)| \leq p(x)$. So, we apply the Real Version of the Hahn-Banach Theorem to l, which is real linear and p satisfies $p((1-\alpha)x + \alpha y) \leq (1-\alpha)p(x) + \alpha p(y)\ \forall \alpha \in [0,1]$. Hence, there exists and L defined on all of X such that $L(x) \leq p(x)$ for all $x \in X$ and $l(y) = L(y)$ for all $y \in Y$. So, we take Λ to be given by $\Lambda(x) = L(x) - iL(ix)$. Λ is linear and $\Lambda(y) = L(y) - iL(y) = l(y) - il(y) = f(y)$ for $y \in Y$. Furthermore, since $|z|$ is real, for any $z \in \mathbb{C}$, we can write $|z| = e^{i\theta} z$ for some θ. So, $\mathbb{R} \ni |\Lambda(x)| = e^{i\theta}\Lambda(x) = \Lambda(e^{i\theta}x)$. Thus, since $\Lambda(e^{i\theta}x)$ is real, $\Lambda(e^{i\theta}x) = L(e^{i\theta}x) \leq p(e^{i\theta}x) \leq |e^{i\theta}|p(x) = p(x)$ (by setting $\beta = 0$ and $\alpha = e^{i\theta}$ and applying the assumptions of the theorem). ∎

1.1.2 Applications of the Hahn-Banach Theorem

Definition Let X be a normed linear space. The *dual* space, denoted X^*, is the space of all bounded linear functions on X :

$$f : X \to \mathbb{K} \text{ is linear, and } \|f\|_{X^*} = \sup_{\|x\|_X \leq 1} |f(x)| < \infty$$

$\|\cdot\|_{X^*}$ defines a norm on X^*, called the *dual norm.* For all $x \in X$, $|f(x)| \leq \|f\|_{X^*} \|x\|_X$.

Lemma 1.1.4 *Let $f : X \to \mathbb{R}$ be linear. The following are equivalent:*

1. *f is bounded*
2. *f is continuous*
3. *f is continuous at a point*

Proof It is clear that (2) $\Rightarrow$ (3).

To show (1) $\Rightarrow$ (2), suppose that f is bounded. Then, let $\|f\|_{X^*} = M$. Fix $\epsilon > 0$. So, letting $\delta = \epsilon/M$, if $\|x-y\| < \delta$, then $|f(x-y)| < \|f\|_{X^*}\delta = M\epsilon/M = \epsilon$.

Finally, to show (3) $\Rightarrow$ (1), assume that f is continuous at a point x_0. Then, $\forall \epsilon > 0,\ \exists \eta > 0$ such that

$$\|x - x_0\| < \eta \Rightarrow |f(x) - f(x_0)| < \epsilon \quad (\text{Hence, } |f(x - x_0)| < \epsilon).$$

Hence, for any y such that $\|y\|_X < \eta$, we have that $|f(y)| < \epsilon$. Now, for any $y \in X, y \neq 0$ let $x = \frac{y}{\|y\|_X}\frac{\eta}{2} \Rightarrow \|x\|_X = \eta/2 \Rightarrow |\frac{\eta}{2\|y\|_X}f(y)| = |f(x)| < \epsilon$. So, for all $y \neq 0$, $|f(y)| < \epsilon\frac{2}{\eta}\|y\|_X$. Hence, f is bounded. ∎

Remark Sometimes, we denote $f(x) =< f, x >$.

Corollary 1.1.5 *Let X be a normed linear space and f be a linear function defined on a subspace $Y \subset X$ with*

$$\|f\|_{Y^*} = \sup_{x\in Y,\ \|x\|_X\leq 1} |f(x)|.$$

Then, f can be extended to $g \in X^*$ such that $g = f$ on Y and $\|g\|_{X^*} = \|f\|_{Y^*}$.

Proof Apply the either the Real or Complex Version of the Hahn-Banach Theorem (depending on the field of scalars $\mathbb{K}$, for X) with $p(x) = \|f\|_{Y^*}\|x\|_X$. It is easy to check that it satisfies all of the semi-norm properties required for the assumptions in the Hahn-Banach Theorem and that $|f| \leq p$ in Y. So, we can extend f to g with $|g(x)| \leq p(x) = \|f\|_{Y^*}\|x\|_X$. Hence, $\|g\|_{X^*} \leq \|f\|_{Y^*}$. On the other hand, if we take any $y \in Y \subset X$, satisfying $\|y\|_Y \leq 1$, we see that $\|g\|_{X^*} \geq |g(y)| = |f(y)|$ (from the H-B Theorem). Hence, $\|g\|_{X^*} \geq \|f\|_{Y^*}$. ∎

Corollary 1.1.6 *For all $x_0 \in X$, there exists $f_0 \in X^*$ such that $f_0(x_0) = \|x_0\|_X^2$ and $\|f_0\|_{X^*} = \|x_0\|_X$.*

Proof Take $Y = \mathbb{K}x_0$, where $\mathbb{K}$ is the base field. Define $g : Y \to \mathbb{K}$ by:

$$g(tx_0) = t \cdot \|x_0\|_X^2.$$

So, $\|g\|_{Y^*} = \sup_{\|tx_0\|_X\leq 1} |g(tx_0)| = \sup_{\|tx_0\|_X\leq 1} |t|\|x_0\|_X^2 = \|x_0\|_X$, the last equality being true by considering the case of $t = \frac{1}{\|x_0\|_X}$. So, we can extend g to $f_0 \in X^*$ such that $\|f_0\|_{X^*} = \|x_0\|_X$ by applying the preceding corollary. ∎

Corollary 1.1.7 *For all $x \in X$,*

$$\begin{aligned} \|x\|_X &= \sup_{\|f\|_{X^*}\leq 1} |< f, x >| \\ &= \max_{\|f\|_{X^*}\leq 1} |< f, x >| \end{aligned}$$

Proof Fix $x_0 \neq 0$ and consider $g = \frac{f_0}{\|x_0\|}$ with f_0 as in the previous result. Then,

$$\sup_{\|f\|_{X^*}\leq 1} |< f, x >| \geq \left|\frac{f_0(x_0)}{\|x_0\|_X}\right| = \|x_0\|_X,$$

since $f_0(x_0) = \|x_0\|_X^2$ and $\|g\|_{X^*} = 1$.

But, $|< f, x >| \leq \|f\|_{X^*}\|x\|_X$. Hence, $\|x\|_X \geq \sup_{\|f\|_{X^*}\leq 1} |< f, x >|$. So, the first equality is proved. For the second one, we note that the sup is achieved for $g = f_0/\|x_0\|_X$. Since f_0 exists by the previous corollary, the sup becomes a max. ∎

Remark In light of this result, compare:

$$\|f\|_{X^*} = \sup_{\|x\|_X \leq 1} | < f, x > |$$
$$\|x\|_X = \sup_{\|f\|_{X^*} \leq 1} | < f, x > |$$

The first is the definition. The second is the previous result.

Corollary 1.1.8 $x = 0 \Leftrightarrow \forall f \in X^*, f(x) = 0$

1.2 Hahn-Banach Theorems - Geometric Versions

In this section, we will investigate a formulation of the Hahn-Banach theorem in terms of separating convex sets by hyperplanes. For the purposes of this section we assume that X is a normed linear space where the base field, $\mathbb{K}$, is $\mathbb{R}$.

1.2.1 Definitions and Preliminaries

Definition A *hyperplane* H is a set of solutions to the equation $f(x) = \alpha$ for some $\alpha \in \mathbb{R}$ and f is a non-zero linear function.

Proposition 1.2.1 *H is closed if and only if f is bounded.*

Definition Suppose $A, B \subset X$.

- The hyperplane $\{f = \alpha\}$ *separates A and B* if $\forall x \in A, f(x) \leq \alpha$, and $\forall x \in B, f(x) \geq \alpha$.
- The hyperplane $\{f = \alpha\}$ *separates A and B strictly* if $\exists \epsilon > 0$ such that $\forall x \in A, f(x) \leq \alpha - \epsilon$, and $\forall x \in B, f(x) \geq \alpha + \epsilon$.

Definition A set A is *convex* if for all $x, y \in A$ and for all $t \in [0, 1]$,

$$t \cdot x + (1 - t) \cdot y \in A.$$

Theorem 1.2.2 (Hahn-Banach Theorem - First Geometric Form) *Let $A, B \subseteq X$ be two non-empty disjoint convex sets, A open. Then, there exists a* closed *hyperplane separating A and B*

The primary tool to be used for proving such a theorem is the idea of a "gauge" of a convex set.

Definition Let C be an open convex subset of X, containing the origin. We define the *gauge* of C to be a map $p : X \to \mathbb{R}_+$ by:

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

Remark Some books refer to the gauge p as the *Minkowski Functional.*

Proposition 1.2.3 *Let p be the gauge of C. Then p has the following properties:*

1. $p(tx) = t \cdot p(x),\ \forall t > 0\ \forall x \in X$
2. $p(x+y) \leq p(x) + p(y),\ \forall x, y \in X$
3. $0 \leq p(x) \leq M\|x\|_X,\ \forall x \in X$
4. $p(x) < 1 \Leftrightarrow x \in C$

Proof

1. The proof is clear.
2. From the first property, if $\lambda > p(x)$ then, $\frac{x}{\lambda} \in C$. So, if $\epsilon > 0$, then, $\frac{x}{p(x)+\epsilon}, \frac{y}{p(y)+\epsilon} \in C$. So, $\forall t \in [0,1]$, since C is convex,
$$t\frac{x}{p(x)+\epsilon} + (1-t)\frac{y}{p(y)+\epsilon} \in C.$$
So, take $t = \frac{p(x)+\epsilon}{p(x)+p(y)+2\epsilon} \in [0,1]$. Hence, $\frac{x+y}{p(x)+p(y)+2\epsilon} \in C$. Since $p(x+y)$ is defined to be the smallest t such that $\frac{p(x+y)}{t} \in C$, it must be that $p(x+y) \leq p(x)+p(y)+2\epsilon$. Since, $\epsilon > 0$ was arbitrary, $p(x+y) \leq p(x)+p(y)$.
3. C is open. So, there is an $r > 0$ such that $C \supset B(0,r)$. So, for all $x \neq 0$ in X,
$$\frac{x}{\|x\|_X}\frac{r}{2} \in C \Rightarrow p(x) \leq \frac{2}{r}\|x\|_X$$
since $p(x)$ is the inf of all t such that $\frac{x}{t} \in C$.
4. Suppose $x \in C$. Then, $(1+\epsilon)x \in C$ for some $\epsilon > 0$ since C is open. So, reasoning as before regarding the minimality of $p(x)$, we have that $p(x) \leq \frac{1}{1+\epsilon} < 1$. Conversely, if $p(x) < 1$, then $\exists \alpha < 1$ such that $\frac{x}{\alpha} \in C$. So, $\alpha \cdot \frac{x}{\alpha} + (1-\alpha) \cdot 0 \in C$ since C is convex. Hence, $x \in C$. ∎

1.2.2 Separation of a Point and a Convex Set

The proof of the following lemma will allow us to prove the First Geometric Form of the Hahn-Banach Theorem

Lemma 1.2.4 *Let $C \subset X$ be an open, non-empty, convex set and x_0 a point such that $x_0 \notin C$. Then, there exists a bounded linear function f such that $f(x) < f(x_0),\ \forall x \in C$.*

Proof Up to translation, we can assume WLOG that $0 \in C$. Define the functional $g : \mathbb{R}x_0 \to \mathbb{R}$ by $g(tx_0) = t$. Then, we apply the Real Version of the Hahn-Banach Theorem to g and p, the gauge of C.

To do so, we just check that $g \leq p$ on $\mathbb{R}x_0$. If $t \geq 0$, then $p(tx_0) = tp(x_0) \geq t = g(tx_0)$ since $p(x) \geq 1$ for $x \notin C$, by Property 4 in

Proposition 1.2.3. On the other hand, if $t \leq 0$, $g(tx_0) = t \leq 0 \leq p(tx_0)$. In either case, $g \leq p$ on $\mathbb{R}x_0$.

So, from Hahn-Banach, we get a linear functional f such that $f = g$ on $\mathbb{R}x_0$ and $f \leq p$ on all of X. So, for x_0, $f(x_0) = g(x_0) = 1$. But, for $x \in C$, $f(x) \leq p(x) < 1$ (by property 2 of Proposition 1.2.3). By Property 2 of the same Proposition, p is bounded and since f is linear and bounded by p, it belongs to X^* Hence, f separates x_0 and C. ∎

Proof of Hahn-Banach Theorem - First Geometric Form Apply the preceding lemma to $C = A - B = \bigcup_{y \in B} A - y$. One can check by hand that C is convex. Also, since A is open, C is open as well (being the union of open sets). Finally, $0 \notin C$, for else, $A \cap B \neq \emptyset$. By the preceding lemma, $\exists f \in X^*$ such that $f(x) < 0$, for all $x \in C$ (since f is linear, $f(0) = 0$). So, for all $x \in A$ $y \in B$, $f(x - y) < 0$. So, by the linearity of f, $f(x) < f(y)$ for any $x \in A$, $y \in B$. Hence,

$$\sup_{x \in A} f(x) \leq \inf_{y \in B} f(y).$$

Therefore, $\exists \alpha \in \mathbb{R}$ such that $f(x) \leq \alpha \leq f(y)$ for all $x \in A$, $y \in B$. Hence, the hyperplane $\{f = \alpha\}$ separates A and B. ∎

Theorem 1.2.5 (Hahn-Banach Theorem - Second Geometric Form) *Let $A, B \neq \emptyset$ be disjoint convex sets with A closed, B compact. Then, there exists a closed hyperplane which strictly separates A and B.*

Proof Consider the sets $A_\epsilon = A + B(0, \epsilon)$, $B_\epsilon = B + B(0, \epsilon)$. For ϵ sufficiently small, they are disjoint. Indeed, suppose $\exists x_n \in A$, $y_n \in B$ such that $\|x_n - y_n\|_X \to 0$. Then, since B is compact, $\exists$ subsequence $\{y_{n_k}\}_{k \in \mathbb{Z}}$ such that $y_{n_k} \to l$. Hence, $x_{n_k} \to l$. Hence, since A is closed, $l \in A \cap B$, contradicting their disjointness. So, by the First Geometric Form of the Hahn-Banach Theorem, $\exists f \in X^*$, $f \neq 0$ and $\alpha \in \mathbb{R}$ such that $\forall x \in A_\epsilon$, $\forall y \in B_\epsilon$, $f(x) \leq \alpha \leq f(y)$. So, $\forall x \in A$, $y \in B$, and $z \in B(0,1)$, we have $f(x + \epsilon z) \leq \alpha \leq f(y + \epsilon z)$. Hence, by choosing z appropriately, we can get that $f(x) \leq \alpha - \epsilon \|f\|_{X^*}$, $f(y) \geq \alpha - \epsilon \|f\|_{X^*}$, $\forall x \in A$, $y \in B$. Hence, f separates A and B strictly. ∎

Corollary 1.2.6 *Let $Y \subseteq X$ be a subspace such that $\overline{Y} \neq X$. Then, $\exists f \in X^*$ such that $f \neq 0$ and $f(y) = 0$ $\forall y \in Y$.*

Remark Stated alternatively, $Y \subseteq X$ subspace is dense $\Leftrightarrow \forall f \in X^*$, $f = 0$ on Y implies $f = 0$.

Proof Assume $\overline{Y} \neq X$. Then, $\exists x_0 \in X \setminus \overline{Y}$. $\overline{Y}$ is closed and convex. $\{x_0\}$ is convex and compact. Hence, by the Second Geometric Form of the Hahn-Banach Theorem $\exists f \in X^*$, $f \neq 0$ such that $f(x) < f(x_0)$ for all $x \in \overline{Y}$. For all $t \in \mathbb{R}$, $tf(x) = f(tx) < f(x_0)$. Hence, for $x \in \overline{Y}$, $f(x) = 0$. Thus, $f = 0$ on Y, but $f \not\equiv 0$. ∎

1.2.3 Applications (Krein-Milman Theorem)

Definition Let K be a subset in a normed linear space.

- $S \subseteq K$ is an *extreme set* if:

$$tx + (1-t)y \in S \text{ for some t} \in (0,1), \quad x, y \in K \quad \Longrightarrow \quad x, y \in S.$$

- A point x_0 is an *extreme point* of K if and only if:

$$x_0 = tx + (1-t)y, \quad 0 < t < 1, \quad x, y \in K \quad \Longrightarrow \quad x = y = x_0.$$

Definition

- The *convex hull* of a set is the smallest convex set containing it.
- The *closed convex hull* of a set is the closure of the convex hull.

Theorem 1.2.7 (Krein-Milman) *Let K be a compact and convex set in X. Then, K is the closed convex hull of its extreme points.*

Remark If X is a normed linear space, then X^* separates points (i.e.: for all $x, y \in X$ such that $x \neq y$ $\exists f \in X^*$ such that $f(x) \neq f(y)$).

Proof of Krein-Milman Theorem Let $\mathcal{P}$ be the collection of all extreme sets in K. We will use the following two properties:

- The intersection of elements of $\mathcal{P}$ is in $\mathcal{P}$ or empty (check!)
- If $S \in \mathcal{P}$ and $f \in X^*$ then if we define $S_f = \{x \in S : f(x) = \max_S f\}$, $S_f \in \mathcal{P}$.

  To show this, let $tx + (1-t)y \in S_f \subseteq S$. Then, $f(tx + (1-t)y) = \max_S f$. Since S is extreme, $x, y \in S$. Thus,

$$tf(x) + (1-t)f(y) = \max_S f \tag{1.1}$$

  If $f(x) < \max_S f$ or $f(y) < \max_S f$, (i.e.: x or $y \notin S_f$) we would have:

$$tf(x) + (1-t)f(y) < \max_S f,$$

  contradicting Eq. (1.1). Hence, $f(x) = f(y) = \max_S f$. This means that, $x, y \in S_f$. Hence, S_f is extreme.

Now, let $S \in \mathcal{P}$. Let $\mathcal{P}'$ be the collection of all extreme sets in S. By the Hausdorff Maximality Theorem $\exists$ a maximal, totally ordered subcollection called $\Omega \subset \mathcal{P}'$. Let $M = \cap_{T \in \Omega} T$. M is an extreme set, ie: $M \in \mathcal{P}'$.

Then, $M_f = M$ by the definition of M.

$$\Rightarrow \ \forall f \in X^*, \ \forall x \in M, \quad f(x) = \max_{x \in M} f(x) = const$$

Hence, every f is constant on M. But, since X^* separates points, M is a singleton. Consequently, $\forall S \in \mathcal{P}$, S contains an extreme point.

Let H denote the convex hull of the set of extreme points in K. We want $\overline{H} = K$. For all $S \in \mathcal{P}$, $S \cap H \neq \emptyset$. Clearly, $\overline{H} \subseteq K$. On the other hand, suppose $\exists x_0 \in K \setminus \overline{H}$. Then, $\{x_0\}$ compact and $\overline{H}$ is closed. By the Second Geometric Form of Hahn-Banach, $\exists f \in X^*$ such that $f(x) < f(x_0)\ \forall x \in \overline{H}$. Hence, if we consider the set $K_f = \{x \in K : f(x) = \max_K f\}$, we see that $K_f \cap \overline{H} = \emptyset$ since $\forall x \in \overline{H},\ \ f(x) < f(x_0) \leq max_K f$. But, K_f is an extreme set and so must intersect $\overline{H}$. This is a contradiction. Hence, $K \subset \overline{H}$. ∎

1.3 Introduction to the Theory of Convex Conjugate Functions

Let X be a topological space. Consider functions $\varphi : X \to (-\infty, \infty]$. Let the domain of φ be defined as $D(\varphi) = \{x \in X : \varphi(x) < +\infty\}$ and the epigraph of φ be defined by $epi(\varphi) = \{(x, \lambda) \in X \times \mathbb{R} : \varphi(x) \leq \lambda\}$.

Definition We say a function $\varphi : X \to (-\infty, +\infty]$ is *lower semicontinuous* (or l.s.c.) if $\forall \lambda \in \mathbb{R}$, the set $\{x : \varphi(x) \leq \lambda\}$ is closed. Equivalently, if $epi(\varphi)$ is closed. Also, if $\forall x \in D(\varphi)$, $\forall \epsilon > 0$, $\exists$ a neighborhood V of x such that $y \in V \Rightarrow f(y) \geq f(x) - \epsilon$.

Remark This allows for possible downhill discontinuities. One can similarly define *upper semicontinuous*.

Proposition 1.3.1

- *If φ is l.s.c., and $x_n \to x$, then $\varphi(x) \leq \liminf_{n\to\infty} \varphi(x_n)$.*
- *A supremum of l.s.c. functions is l.s.c. (ie: $\varphi(x) = \sup_i \varphi_i(x)$ is l.s.c. if $\varphi_i(x)$ are l.s.c.).*

Definition We say f is *convex* if $\forall x, y \in X$, and $t \in [0, 1]$, $f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y)$. Or equivalently, if $epi(f)$ is convex.

Definition Let X be a normed linear space, and $\varphi : X \to (-\infty, +\infty]$. We define the *conjugate function* (or *Legendre-Fenchel transform*) of φ as $\varphi^* : X^* \to (-\infty, +\infty]$ by:

$$\varphi^*(f) = \sup_{x \in D(\varphi)} (< f, x > -\varphi(x)). \quad \text{(Requires that } \varphi \not\equiv +\infty)$$

Remark Observe that:

- $\forall x \in D(\varphi)$, $f \longmapsto < f, x > -\varphi(x)$ is an affine function (hence, continuous and convex).

- A supremum of affine functions is l.s.c. and convex $\longrightarrow \varphi^*$ is convex and l.s.c.

Proposition 1.3.2 *If φ is convex, l.s.c. and $\varphi \not\equiv +\infty$, then $\varphi^* \not\equiv +\infty$.*

Proof Look at *epi*(φ). It is closed and convex. So, let $(x_0, \tilde{\lambda}_0) \in epi(\varphi)$ (such a point exists since $\varphi \not\equiv +\infty$, so consider a point, (x_0, λ_0) below *epi*(φ). In other words, choose $\lambda_0 < \varphi(x_0)$. $\{(x_0, \lambda_0)\}$ is compact and convex. So, apply the Second Geometric Form of the Hahn-Banach Theorem in $X \times \mathbb{R}$ to this set and to *epi*(φ). So, $\exists$ a linear functional Λ on $X \times \mathbb{R}$ and an α such that $\forall (x, \lambda) \in epi(\varphi), \Lambda(x, \lambda) > \alpha > \Lambda(x_0, \lambda_0)$.

Now, we can write $\Lambda(x, \lambda) = f(x) + k\lambda$ for some $f \in X^*$ and $k \in \mathbb{R}$ since Λ is linear. So, $\forall x,\ \forall \lambda \geq \varphi(x),\quad f(x) + k\lambda > \alpha > f(x_0) + k\lambda_0$. In particular, for $\lambda = \varphi(x)$, and $\forall x \in D(\varphi)$,

$$f(x) + k\varphi(x) > \alpha > f(x_0) + k\lambda_0 \tag{1.2}$$

We consider the sign of k. At x_0, $f(x_0) + k\varphi(x_0) > f(x_0) + k\lambda_0 \quad \Rightarrow$ $k > 0$, since (x_0, λ_0) was chosen so that $\varphi(x_0) > \lambda_0$. So, we divide both sides of Equation (1.2) by k :

$$\begin{aligned} \frac{f(x)}{k} + \varphi(x) \quad &> \quad \frac{f(x_0)}{k} + \lambda_0 \\ \Rightarrow -\frac{f(x)}{k} - \varphi(x) \quad &< \quad -\frac{f(x_0)}{k} - \lambda_0 \ \ \forall x \in D(\varphi) \end{aligned}$$

The left hand side is linear in x. So, taking supremums in x, we get that:

$$\begin{aligned} \sup_x \left(-\frac{f(x)}{k} - \varphi(x) \right) \quad &\leq \quad -\frac{f(x_0)}{k} - \lambda_0 \\ \Longrightarrow \varphi^* \left(-\frac{f}{k} \right) \quad &\leq \quad -\frac{\alpha}{k} < \infty. \quad \blacksquare \end{aligned}$$

We can also define the *bi-conjugate* of φ in the following manner:

$$\varphi^{**} : X \to (-\infty, +\infty], \quad \varphi^{**}(x) = \sup_{f \in D(\varphi^*)} [< f, x > -\varphi^*(f)].$$

This function is convex and lower semicontinuous. So, the diagram looks like:

$$\begin{array}{ccccc} \varphi & \longrightarrow & \varphi^* & \longrightarrow & \varphi^{**} \\ \not\equiv & & \text{convex} & & \text{convex} \\ \infty & & \text{l.s.c.} & & \text{l.s.c.} \end{array}$$

Theorem 1.3.3 (Fenchel-Moreau) *If φ is convex and l.s.c. and $\not\equiv +\infty$, then $\varphi^{**} = \varphi$.*

Proof First, we show that $\varphi^{**} \leq \varphi$: By definition of φ^*, $\forall x \in X, \quad f \in X^*$:

$$< f, x > -\varphi(x) \quad \leq \quad \varphi^*(f) \tag{1.3}$$

$$(1.3) \implies \forall x \in X, \quad \sup_{f \in X^*} (< f, x > -\varphi^*(f)) \leq \varphi(x).$$

Hence, $\varphi^{**}(x) \leq \varphi(x)$.

Now, assume by contradiction that $\exists x_0$ such that $\varphi^{**}(x_0) < \varphi(x_0)$. Then, $epi(\varphi)$ lies "above" $(x_0, \varphi^{**}(x_0))$. In other words, we can use the Hahn-Banach Theorem to separate $epi(\varphi)$ and $\{(x_0, \varphi^{**}(x_0))\}$. So, there exists, $f \in X^*$, $k \in \mathbb{R}, \ \alpha \in \mathbb{R}$ such that $\forall x \in D(\varphi)$ and $\lambda \geq \varphi(x)$:

$$f(x) + k\lambda > \alpha > f(x_0) + k\varphi^{**}(x_0) \tag{1.4}$$

Note that we used a similar technique as in Proposition 1.3.2 to break down the operator given to us by the H-B Theorem into f and k. Again, we can conclude that $k \geq 0$, for else, we could send $\lambda \to +\infty$ and get a contradiction.

So, we first assume that $\varphi(x) \geq 0$. Applying the relation in Equation 1.4 to $\lambda = \varphi(x)$, we get that $f(x) + k\varphi(x) > \alpha$. Hence, for all $\epsilon > 0$:

$$f(x) + (k+\epsilon)\varphi(x) > \alpha \implies -\frac{f(x)}{k+\epsilon} - \varphi(x) \ < \ -\frac{\alpha}{k+\epsilon} \quad \forall x$$

$$\implies \sup_{x \in D(\varphi)} \left[-\frac{f(x)}{k+\epsilon} - \varphi(x) \right] \leq -\frac{\alpha}{k+\epsilon} \quad \implies \varphi^*\left(-\frac{f}{k+\epsilon} \right) \ \leq \ -\frac{\alpha}{k+\epsilon}$$

So, we see that:

$$\begin{aligned} \varphi^{**}(x_0) = \sup_{f \in X^*} (< f, x_0 > -\varphi^*(f)) &\geq \left\langle -\frac{f}{k+\epsilon}, x_0 \right\rangle - \varphi^*\left(-\frac{f}{k+\epsilon} \right) \\ &\geq \left\langle -\frac{f}{k+\epsilon}, x_0 \right\rangle + \frac{\alpha}{k+\epsilon} \\ \implies (k+\epsilon)\varphi^{**}(x_0) &\geq < -f, x_0 > +\alpha \end{aligned}$$

So, we take $\epsilon \to 0$, and get that $f(x_0) + k\varphi^{**}(x_0) \geq \alpha$, which contradicts $\alpha > f(x_0) + k\varphi^{**}(x_0)$ in Equation 1.4. Hence, $\varphi \geq 0 \Rightarrow \varphi^{**} = \varphi$.

Now, we consider any φ and $f_0 \in D(\varphi^*)$. Define a new function:

$$\overline{\varphi}(x) = \varphi(x) - < f_0, x > +\varphi^*(f_0).$$

Fix x. $\varphi^*(f_0) = \sup_y [< f, y > -\varphi(y)] \geq f_0(x) - \varphi(x)$. Since x was arbitrary, this shows that $\overline{\varphi} \geq 0$. So, we can apply the result we obtained above to see that $\overline{\varphi} = \overline{\varphi}^{**}$. But,

$$\begin{aligned} \overline{\varphi^*}(f) &= \sup_x [< f, x > -\overline{\varphi}(x)] = \sup_x [< f, x > + < f_0, x > -\varphi(x) - \varphi(f_0)] \\ &= \sup_x [< f + f_0, x > -\varphi(x)] - \varphi^*(f_0) = \varphi^*(f + f_0) - \varphi^*(f_0) \end{aligned}$$

Also,

$$\begin{aligned}
\overline{\varphi}^{**}(x) &= \sup_f [< f, x > -\overline{\varphi}^*(f)] = \sup_f [< f, x > -\varphi^*(f + f_0] + \varphi^*(f_0) \\
&= \sup_f [< f + f_0, x > -\varphi^*(f + f_0)] - < f_0, x > +\varphi^*(f_0) \\
&= \sup_g [< g, x > -\varphi^*(g)] - < f_0, x > +\varphi^* f_0 \\
&= \varphi^{**}(x) + \varphi^*(f_0) - < f_0, x >
\end{aligned}$$

Here, we have used the fact that f_0 is independent of the sup taken over all f. Hence, putting everything together, since $\overline{\varphi}^{**} = \overline{\varphi}$, we get that $\varphi^{**} = \varphi$. ∎

Example Say that $\varphi(x) = \|x\|$. Surely, φ is a convex and lower semicontinous function from X to $\mathbb{R}$. $\varphi^*(f) = \sup_{x \in X} [< f, x > -\|x\|]$.

- If $\|f\| \leq 1$, then $< f, x > \leq \|x\| \implies \varphi^*(f) \leq 0 \implies \varphi^*(f) = 0$ (since we can just take $x = 0$ and the sup will be at least 0).
- If $\|f\| > 1$, then $\exists x$ such that $f(x) > (1 + \epsilon)\|x\|$. So, $f(x) - \|x\| > \epsilon\|x\|$. If we consider the case of nx and then letting n go to $+\infty$ we see that $\varphi^*(f) = +\infty$.

This means that:

$$\varphi^{**}(x) = \sup_{f \in D(\varphi^*)} [< f, x > -\varphi^*(f)] = \sup_{\|f\| \leq 1} (< f, x >) = \|x\| = \varphi(x).$$

Theorem 1.3.4 (Fenchel-Rockafellar) *Assume φ, ψ are two convex functions and $\exists x_0 \in X$ such that $\varphi(x_0) < \infty$, $\psi(x_0) < \infty$ and φ is ontinuous at x_0. Then,*

$$\inf_{x \in X} (\varphi(x) + \psi(x)) = \sup_{f \in X^*} [-\varphi^*(-f) - \psi^*(f)] = \max_{f \in X^*} [-\varphi^*(-f) - \psi^*(f)]$$

Proof Exercise. The proof is similar to the previous result. ∎

Remark This theory has wide range of applications:

- Optimization (sometimes the dual problem is easier to deal with)
- PDE
- Convex Programming (see Ekeland - Teman, *Intro to Convex Analysis*).

Chapter 2

Baire Category Theorem and Its Applications

2.1 Review

2.1.1 Reminders on Banach Spaces

Definition A *Banach space* is a complete normed linear space (i.e. every Cauchy sequences converges in that space w.r.t. to its norm).

Example

- Hilbert spaces are Banach spaces
- $L^p(X, d\mu)$, $1 \leq p \leq \infty$ are Banach spaces.
- $l_p = \{(u_n)_{n\in\mathbb{N}} : \left(\sum_n |u_n|^p\right)^{1/p} < \infty\}$, $1 \leq p \leq \infty$ are Banach. (Note that $l_p = L^p(\mathbb{R}, d\mu)$ where $\mu = \sum_n \delta_n$ and the δ_n are the Dirac masses at the integers.)

2.1.2 Bounded Linear Transformations

Definition A *bounded linear transformation* (or bounded operator) T between two normed linear spaces $(X_1, \|\cdot\|_1)$ and $(X_2, \|\cdot\|_2)$ is a linear mapping such that $\exists C \geq 0$ such that:

$$\forall x \in X_1, \quad \|Tx\|_2 \leq C\|x\|_1.$$

As we had shown before, T bounded $\iff T$ continuous $\iff T$ continuous at a point.

We define the *operator norm* as:

$$\|T\| = \sup_{\|x\|_1 \leq 1} \|Tx\|_2$$

Lemma 2.1.1 *The operator norm is a norm.*

Proof First, we show the triangle inequality. This follows from the fact that $\|Tx+Sx\|_2 \leq \|Tx\|_2+\|Sx\|_2$ since $\|\cdot\|_2$ is a norm and that $\sup_{\|x\|_1\leq 1} \|Tx\|_2 + \|Sx\|_2 \leq \sup_{\|x\|_1\leq 1} \|Tx\|_2 + \sup_{\|x\|_1\leq 1} \|Sx\|_2$.

Next, we show that $\|aT\| = |a|\|T\|$. Again, this follows from the fact that $\|aTx\|_2 = |a|\|Tx\|_2$ since $\|\cdot\|_2$ is a norm and from the fact that for any positive number a, $\sup a\cdot = a \sup \cdot$.

Finally, we show that $\|T\| = 0 \Rightarrow T = 0$. In other words, we must show that $Tx = 0\ \forall x \in X_1$. Again, $\|T\| = 0\ \rightarrow \forall \|x\|_1 \leq 1, \|Tx\|_2 = 0$. Then, by linearity, this means that $\forall x \in X_1, \|Tx\|_2 = 0$. But then, since $\|\cdot\|_2$ is a norm, this means that $Tx = 0$. Hence, $T = 0$. ∎

We denote by $\mathcal{L}(X_1, X_2)$, the space of bounded linear operators from X_1 to X_2, with norm $\|T\|$ given by above.

Example $\mathcal{L}(X, \mathbb{K}) = X^*$

Theorem 2.1.2 *If Y is a Banach space, then $\mathcal{L}(X, Y)$ is as well.*

Proof It is clear that the space is normed, and linear. Remains to show completeness. Let A_n be a Cauchy sequence of functions in $\mathcal{L}(X, Y)$. i.e.:

$$\|A_n - A_m\| \underset{n,m\to\infty}{\longrightarrow} 0$$

$\implies \forall x \in X,\ A_n(x)$ is Cauchy. Hence, it has a limit, which we will denote by $A(x)$. The fact that the mapping $x \mapsto A(x)$ is linear is clear. Now, we seek to show boundedness:

$$\|A(x)\|_Y = \lim_{n\to\infty} \|A_n(x)\| \leq \limsup \|A_n\|\|x\|_X.$$

But, the $\{A_n\}_{n\in\mathbb{N}}$ form a Cauchy sequence. Hence, they are uniformly bounded by some constant C. Hence, $\|Ax\| \leq C\|x\|_X$ by above. Hence, $A \in \mathcal{L}(X, Y)$. Finally, we show that it is the limit of the A_n's:

$$\begin{aligned} \|(A_n - A)(x)\| &= \lim_{m\to\infty} \|(A_n - A_m)(x)\| \\ &\leq \limsup_{m\to\infty} \|A_n - A_m\|\|x\|_X \\ &\leq o(1)\|x\|_X \end{aligned}$$

since the A_n are Cauchy. Hence, $\|A_n - A\|_{\mathcal{L}(X,Y)} \to 0$. ∎

Definition We say two norms, $\|\cdot\|_1$ and $\|\cdot\|_2$ are *equivalent*, if $\exists$ constants, C_1, C_2 such that $C_1\|\cdot\|_1 \leq \|\cdot\|_2 \leq C_2\|\cdot\|_1$.

Proposition 2.1.3 *In a finite dimensional space, all norms are equivalent. Equivalent norms define the same topology.*

This result is not true in infinite dimensions.

Definition An *isomorphism* between two normed linear spaces is a bounded linear map which is bijective with bounded linear inverse.

$\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent $\Leftrightarrow$ the function $id : (X, \|\cdot\|_1) \to (X, \|\cdot\|_2)$ is an isomorphism (continuous with continuous inverse).

Example All separable (there exists a countable Hilbert basis) Hilbert spaces are isomorphic to l_2.

2.1.3 Duals and Double Duals

Theorem 2.1.4 (Riesz Representation Theorem) *Let H be a Hilbert space. Then, for every $F \in H^*$, then $\exists$ a unique $f \in H$ such that $\forall x \in H$, $F(x) =< f, x >$.*

So, $H \mapsto H^*$ is an isomorphism between H and H^*, through which we can identify H and its dual.

Example

- $(L^2)^* = L^2$
- For $1 < p < \infty$,$(L^p)^* = L^{p'}$ where $\frac{1}{p} + \frac{1}{p'} = 1$. In other words,

  Theorem 2.1.5 *Let $1 < p < \infty$. Then, $\forall F \in (L^p)^*$, there exists a unique $f \in L^{p'}$ where $\frac{1}{p} + \frac{1}{p'} = 1$ such that for all $\varphi \in L^p$,*

  $$F(\varphi) = \int f \varphi d\mu.$$

- While $(L^1)^* = L^\infty$, $(L^\infty)^* \supsetneq L^1$. In fact, $(L^\infty)^* = \{$ measures $\}$.

Definition The *double dual* is the dual of the dual. i.e.: $X^{**} = (X^*)^*$.

Consider $J : X \to X^{**}$ given by $J(x)(v) = v(x)$, for $v \in X^*$. This is the *canonical embedding* of X into X^{**}. It is an isometric embedding. To see this, note that:

$$\|J(x)\| = \sup_{\|f\| \leq 1} \|J(x)(f)\| = \sup_{\|f\| \leq 1} | < f, x > | \leq \sup_{\|f\| \leq 1} \|f\| \cdot \|x\| \leq \|x\|.$$

On the other hand, by a Corollary to the Hahn-Banach Theorem (Corollary 1.1.6), $\exists f \in X^*$ such that $\|f\| \leq 1$ and $< f, x >= \|x\|$. Hence, the sup is achieved and there is equality.

In general, $J(X) \subset X^{**}$. Formally, we say, "$X \subset X^{**}$" with the canonical identification. In finite dimensions, we have equality. But this is not necessarily true in infinite dimensions.

Definition If $J(X) = X^{**}$, we say that X is *reflexive*.

Examples of Reflexive Spaces

- Hilbert spaces, as seen above.
- L^p spaces for $1 < p < \infty$ (since $(L^p)^{**} = (L^{p'})^* = L^{p''}$ where $\frac{1}{p} + \frac{1}{p'} = 1$ and $\frac{1}{p'} + \frac{1}{p''} = 1 \Rightarrow p'' = p$).
- Not L^1 since $(L^1)^* = L^\infty$ but $(L^\infty)^* \supsetneq L^1$, as we saw above.

2.2 The Baire Category Theorem

This theorem is used to prove that some sets have non-empty interior.

Example of Usefulness If T is a linear map and $T^{-1}(B(0,1))$ has non-empty interior, then T is bounded.

To see this, note that if $T^{-1}(B(0,1))$ has non-empty interior, then $\exists x_0,\ \epsilon > 0$ such that $T^{-1}(B(0,1)) \supset B(x_0,\epsilon) \Rightarrow B(0,1) \supset T(B(x_0,\epsilon)) \Rightarrow \forall y$ such that $\|y\| < \epsilon, \|T(x_0 + y)\| \le 1 \Rightarrow \|T(y)\| \le 1 + \|T(x_0)\|$. For all x, $\|T\left(\frac{x}{\|x\|}\frac{\epsilon}{2}\right)\| \le C \Rightarrow \|T(x)\| \le \tilde{C}\|x\|$.

Theorem 2.2.1 (Baire Category Theorem) *Let X be a complete metric space and let F_n be a sequence of closed subsets of X with empty interior (i.e.: $\operatorname{int}(F_n) = \emptyset$) then* $\operatorname{int}(\cup(F_n)) = \emptyset$.

Complementary Form of Theorem If $\mathcal{O}_n$ is a sequence of dense open subsets of X, then $\cap\mathcal{O}_n$ is also dense.

Remark A subset whose closure has empty interior is called "nowhere" dense.

Proof of the Baire Category Theorem Let $\mathcal{O}_n$ be a sequence of dense open subsets. Then, $\cap\mathcal{O}_n$ is dense if we can prove that it intersects every open set.

Let W be an arbitrary open set. Let $x_0 \in W$. Then, $\exists r_0 > 0$ such that $B(x_0, r_0) \subset W$. Since $\mathcal{O}_1$ is dense, its intersection with $B(x_0, r_0)$ is non-empty. Hence, $\exists x_1 \in B(x_0, r_0) \cap \mathcal{O}_1$. Since $B(x_0, r_0) \cap \mathcal{O}_1 = \mathcal{O}_1'$ is an intersection of open sets, it is itself open. Hence, $\exists r_1 > 0$ such that $\overline{B(x_1, r_1)} \subset \mathcal{O}_1'$ and $r_1 < \frac{r_0}{2}$.

So, by induction we build a sequence of x_n's such that

$$\overline{B(x_n, r_n)} \subset B(x_{n-1}, r_{n-1}) \cap \mathcal{O}_n \quad \text{and} \quad r_n < r_{n-1}/2.$$

Hence, $d(x_n, x_{n-1}) \le r_{n-1} \le \frac{r_0}{2^{n-1}}$. This is a Cauchy sequence in $\cap\mathcal{O}_n$. Since X is complete, the x_n have a limit, $l \in X$. Now, since the $\overline{B(x_n, r_n)}$ are closed, and form a decreasing sequence of sets, for each n, $l \in \overline{B(x_n, r_n)} \subset \mathcal{O}_n$. Hence, $l \in \cap\mathcal{O}_n$ and $l \in \cap\overline{B(x_n, r_n)}$. Hence, $l \in B(x_0, r_0) \subset W$. Thus, $l \in W \cap (\cap\mathcal{O}_n)$. Thus, $\cap\mathcal{O}_n$ intersects every open set and is therefore dense. ∎

2.3 The Uniform Boundedness Principle

Theorem 2.3.1 (Uniform Boundedness Principle (Banach-Steinhaus)) *Let X be a Banach Space and Y a normed linear space. Let $(T_i)_{i\in I}$ be an arbitrary family of elements of $\mathcal{L}(X,Y)$ such that:*

$$\forall x \in X, \quad \sup_{i\in I} \|T_i(x)\| < \infty$$

Then:

$$\sup_{i\in I} \|T_i\| < \infty$$

Proof Consider the sets:

$$F_n = \{x \in X : \|T_i(x)\| \le n \ \ \forall i \in I\}.$$

$\cup_{n\in\mathbb{N}} F_n = X$ because each x is in some F_n by assumption. Moreover, each F_n is closed by the continuity of the T_i. The Baire Category Theorem says that if the F_n have empty interior, then their union must have empty interior as well. But, $\cup_{n\in\mathbb{N}} F_n = X$ which certainly doesn't have empty interior. Hence, at least one of the F_n cannot have an empty interior. Suppose F_{n_0} is such an F_n. Then, it follows that $\exists x_0,\ \epsilon > 0$ such that $B(x_0,\epsilon) \subset \ int(F_{n_0})$. So,

$$\forall y,\ \ \|T_i(x_0+\frac{y}{\|y\|}\frac{\epsilon}{2})\|_Y \le n_0 \implies \|T_i(\frac{y}{\|y\|}\frac{\epsilon}{2})\| \le n_0+\|T_i(x_0)\| < n_0+C_0 \ \ \forall i \in I$$

The last inequality is true since by assumption, for each y, the $T_i(y)$ are bounded uniformly in i. So, for all y, $\|T_i(y)\| < \frac{2(C_0+n_0)}{\epsilon}\|y\|$. Since each quantity on the right is independent of i, we can take the sup over all i on both sides and get that $\sup_i \|T_i\| < \frac{2(C_0+n_0)}{\epsilon}$. ∎

Corollary 2.3.2 *Let $(T_n)_{n\in\mathbb{N}}$ be a sequence of bounded linear functions between two Banach spaces X and Y such that $\forall x \in X$, $T_n(x)$ converges to a limit denoted by Tx. Then,*

- $T \in \mathcal{L}(X,Y)$.
- $\sup_{n\in\mathbb{N}} \|T_n\| < \infty$.
- $\|T\|_{\mathcal{L}(X,Y)} \le \liminf_{n\to\infty} \|T_n\|_{\mathcal{L}(X,Y)}$.

Proof That T is linear ought be clear. $\forall x \in X$, such that $\|x\| = 1$,

$$\sup_n \|T_n(x)\|_Y < \infty$$

since $T_n(x)$ converges by assumption. Hence, by the Uniform Boundedness Principle, $\sup_{n\in\mathbb{N}} \|T_n\| < \infty$. This proves the second claim.

To show the first, $\forall x \in X$, $\|T_n(x)\|_Y \le C\|x\|_X \ \Rightarrow \ \|Tx\|_Y \le C\|x\|$ by passing to the limit (since the RHS in the first inequality is independent of n). Hence, $T \in \mathcal{L}(X,Y)$.

Finally, for the third claim, notice that $\forall x \in X$, $n \in \mathbb{N}$, $\|T_n(x)\|_Y \leq \|T_n\|\|x\|_X$. Passing to the limit again, we see that since $T_n(x) \longrightarrow T(x)$,

$$\begin{aligned} \|T(x)\|_Y &\leq \liminf_{n\to\infty} \|T_n\|\|x\|_Y \\ \Longrightarrow \|T\| &\leq \liminf_{n\to\infty} \|T_n\| \quad \blacksquare \end{aligned}$$

Example Limits in distributions. It is sufficient to just have pointwise convergence.

Corollary 2.3.3 *Let B be a subset in a Banach space X. If $\forall f \in X^*$, $f(B) = \cup_{x\in B} f(x)$ is bounded, then B is bounded.*

Proof The idea is to apply the Uniform Boundedness Principle to the family $\{T_b\}_{b\in B}$ given by:

$$T_b : X^* \longrightarrow \mathbb{K}, \qquad T_b(f) = < f, b >$$

for each $b \in B$. But, we have the following:

$$\begin{aligned} \forall f \in X^* \quad &, \ \sup_{b\in B} \|T_b(f)\| < \infty \\ \Longleftrightarrow \quad & \forall f \in X^*, \ \sup_{b\in B} | < f, b > | < \infty \\ \Longleftrightarrow \quad & \forall f \in X^*, \ f(B) \text{ is bounded.} \end{aligned}$$

Hence, $\{T_b\}_{b\in B}$ satisfies the hypothesis of the Uniform Boundedness Principle. So, $\exists C$ such that $\forall f \in X^*$, $\|T_b(f)\| \leq C\|f\|$ $\forall b \in B$. $\Longleftrightarrow$ $\forall f \in X^*, \forall b \in B$, $| < f, b > | \leq C\|f\| \iff \|b\|_X \leq C$. $\blacksquare$

Corollary 2.3.4 *Let X be a Banach space and B' a subset of X^*. If $\forall x \in X$, $B'(x) = \cup_{f\in B'} f(x)$ is bounded, then B' is bounded.*

Proof $T_f(x) = f(x)$ where $T_f : X \to \mathbb{K}$. So,

$$\sup_{f\in B'} \|T_f(x)\| < \infty \quad \forall x$$

So, by the Uniform Boundedness Principle,

$$\sup_{f\in B'} \|T_f\| < \infty$$

and we can finish as above $\blacksquare$.

2.4 The Open Mapping Theorem and Closed Graph Theorem

Theorem 2.4.1 (The (Banach) Open Mapping Theorem) *Let T be a linear map from the Banach space X, **onto** another Banach space Y. Then, T is open: The image of any open set is open.*

Proof By translation and linearity, for any $r > 0$, it is enough to prove that

$$T(B_X(0,r)) \supseteq B_Y(0,r') \quad \text{for some } r'.$$

Define $F_n = \overline{T(B(0,n))}$. Since T is onto, we have $\cup_{n\in\mathbb{N}} F_n = Y$. So, by Baire Category Theorem, some F_{n_0} has nonempty interior. By rescaling, $\text{int}(\overline{T(B(0,1))}) \neq \emptyset$. Hence, we can assume that for some $\epsilon > 0$, $B(0,\epsilon) \subseteq \overline{T(B(0,1))}$ (since it has non-empty interior).

We are going to show that $\overline{T(B(0,1))} \subseteq T(B(0,3))$ and therefore, that $B(0,\epsilon) \subseteq T(B(0,3))$. Since we're in a linear space and T is linear, we can rescale so that $B(0,\epsilon r/3) \subseteq T(B(0,r))$ and so we can choose $r' = \epsilon r/3$.

So, let y be in $\overline{T(B(0,1))}$ and we will find $x \in B(0,3)$ such that $y = Tx$. By the definition of closure there exists $x_1 \in B(0,1)$ such that

$$\begin{aligned} \|y - Tx_1\| &\leq \frac{\epsilon}{2} \\ \Rightarrow y - Tx_1 &\in B\left(0, \frac{\epsilon}{2}\right) \subset \overline{T\left(B\left(0, \frac{1}{2}\right)\right)}. \end{aligned}$$

By the definition of closure, $\exists x_2 \in B(0,1/2)$ such that $\|y - Tx_1 - Tx_2\| \leq \epsilon/4$. So, we iterate in this manner to get that:

$$\forall n, \quad \exists x_n \text{ such that } \; \|y - Tx_1 - \ldots - Tx_n\| < \frac{\epsilon}{2^n} \; \text{ and } \; \|x_n\| < \frac{1}{2^n}.$$

So, we take $x = \sum_{i=1}^{\infty} x_i$, which converges since the sequence is Cauchy. So,

$$\begin{aligned} &\Rightarrow \quad \|x\| \leq \sum_{i=1}^{\infty} \|x_i\| \leq 2 < 3 \\ &\Rightarrow \quad x \in B(0,3) \;\text{ and }\; Tx = y \\ &\Rightarrow \quad \overline{T(B(0,1))} \subset T(B(0,3)) \quad \blacksquare \end{aligned}$$

Corollary 2.4.2 *If T is a bounded linear map between two Banach spaces which is also bijective, then its inverse is also continuous. Hence, T is an isomorphism.*

Theorem 2.4.3 (Closed Graph Theorem) *Let X, Y be two Banach spaces and $T : X \to Y$ linear. Then, T is bounded if and only if its graph, $\Gamma(T) = \{(x,Tx) : x \in X\} \subset X \times Y$ is closed.*

Remark If X is a Banach space for both $\|\cdot\|_1$ and $\|\cdot\|_2$, and $\exists C > 0$ such that $\|x\|_1 \leq C\|x\|_2$, then $\exists C_1$ such that $\|x\|_2 \leq C_1\|x\|_1$.

Indeed, consider $Id : (X, \|\cdot\|_2) \to (X, \|\cdot\|_1)$. By assumption, it is a bounded map that is also a bijection. So, by the corollary to the Open Mapping Theorem, it has bounded inverse.

Proof of Closed Graph Theorem Apply the above remark to the norms: $\|\cdot\|_1, \|\cdot\|_2$ given by:

$$\|x\|_1 = \|x\|_X \quad \|x\|_2 = \|x\|_X + \|Tx\|_Y$$

Certainly, $\|x\|_1 \leq \|x\|_2$.

($\Longrightarrow$) So, assume that $\Gamma(T)$ is closed. Is $(X, \|\cdot\|_2)$ a Banach space? Well, take a Cauchy sequence $\{x_n\}_{n\in\mathbb{N}}$ in $(X, \|\cdot\|_2)$. Then, $x_n \to x$ for $\|\cdot\|_1$, since $\|\cdot\|_1$ is always bounded from above by $\|\cdot\|_2$. Similarly, $\{T(x_n)\}_n$ is Cauchy in Y. So, since Y is Banach, it converges to some $y \in Y$. So, $(x_n, T(x_n)) \to (x, y)$. Hence, $Tx = y$ since the graph of T is closed (and thus, the graph contains all limit points). Therefore,

$$\|x_n - x\|_2 = \|x_n - x\|_1 + \|T(x_n - x)\|_Y \to 0 + 0 = 0.$$

This proves that $(X, \|\cdot\|_2)$ is a Banach space. So, by the remark above, $\|T(x)\|_Y = \|x\|_2 - \|x\|_1 \leq \|x\|_2 \leq C_1\|x\|_1 = C_1\|x\|_X$, as desired.

($\Longleftarrow$) Now, assume that T is bounded. Hence, it's continuous. So, let $\{(x_n, T(x_n))\}_n$ be convergent in $X \times Y$ such that $(x_n, T(x_n)) \to (x, y) \in X \times Y$. Then, $x_n \to x$, so $T(x_n) \to T(x)$ by continuity of T. Hence, $y = Tx$ and $(x, y) = (x, Tx) \in \Gamma(T)$. This proves the graph is closed. ∎

Chapter 3

Weak Topology

3.1 General Topology

Definition A *topological space* is a set S with a distinguished family of subset τ called *the topology* (a.k.a. all open sets) satisfying:

- $\emptyset$ and S are in τ.
- A finite intersection of elements of τ is in τ.
- An arbitrary union of elements of τ is in τ.

Definition A set S is *closed* if its complement is open.

Definition A family $\mathcal{B} \subseteq \tau$ is called a *base* if every element of τ can be written as a union of elements of $\mathcal{B}$.

Definition A set N is a *neighborhood* of $x \in S$ if there exists $U \in \tau$ such that $x \in U \subset N$ (the neighborhood does not have to be open).

Definition A family $\mathcal{N}$ is a *neighborhood base* of x if it is a family of neighborhoods of x s.t. for every neighborhood M of x, $\exists N \in \mathcal{N}$ s.t. $N \subset M$.

Definition A function between two topological spaces is *continuous* if the inverse image of any open set is open.

Definition A topological space is *Hausdorff* if $\forall x, y \in S$, there exists $O_x, O_y \in \tau$ such that $x \in O_x$, $y \in O_y$ and $O_x \cap O_y = \emptyset$.

Example Metric spaces are Hausdorff.

Definition A set K is *compact* if every open cover of K has a finite subcover.

Remark

- The image of a compact set by a continuous function is compact.
- Two extreme cases: $\tau = \cup_{x \in S}\{x\}$, the discrete topology, where the only sequences that converge are constant sequences and $\tau = \{\emptyset, S\}$, the indescrete topology, where all sequences converge.
- More generally, the more open sets there are, the harder it is to converge.

Now, let $\varphi_i : X \to Y_i,\ i \in I$ be mappings from X to topological spaces Y_i. What is the weakest topology on X that makes all the φ_i *continuous*?

Obviously, it must contain $\varphi_i^{-1}(O_i)$ where O_i is any open set in Y_i, along with arbitrary unions and finite intersections. So, the answer is:

$$\tau = \{ \bigcup_{\text{arbitrary}} \bigcap_{\text{finite}} \varphi_i^{-1}(O_i) \}$$

where the O_i are open in Y_i.

3.2 Frechet Spaces

Definition A *seminorm* ρ on a linear space X is a map from X to $[0, +\infty)$ that satisfies the following:

1. $\rho(x+y) \leq \rho(x) + \rho(y)$
2. $\rho(\lambda x) = |\lambda|\rho(x) \quad \forall \lambda \in \mathbb{K}$

A family of seminorms, $\{\rho_\alpha\}_{\alpha \in A}$ is said to *separate points* if and only if $\rho_\alpha(x) = 0\ \forall \alpha \implies x = 0$.

Definition A *locally convex space* is a linear space endowed with a family of seminorms, $\{\rho_\alpha\}_{\alpha \in A}$, which separate points. The natural topology is the one that makes all of the ρ_α continuous, and makes addition in the space continuous.

In a locally convex space, a basis of neighborhoods of 0 is given by sets of the form:

$$\mathcal{N}_{\alpha_1,\ldots,\alpha_N;\epsilon} = \{x \in X : \rho_{\alpha_i}(x) < \epsilon, \quad \forall i = 1, \ldots, N\}.$$

A basis of neighborhoods of any point $x_0 \in X$ is given by sets of the form:

$$\mathcal{N}_{\alpha_1,\ldots,\alpha_N;\epsilon} = \{x \in X : \rho_{\alpha_i}(x - x_0) < \epsilon, \quad \forall i = 1, \ldots, N\}.$$

Characterization A linear mapping T is continuous if and only if $\exists C > 0$ such that $\|T(x)\| \leq C(\rho_{\alpha_1}(x) + \ldots + \rho_{\alpha_N}(x))$.

Proposition 3.2.1 *A locally convex space is Hausdorff*

Proof Take $x \neq y$. Then, $\exists \alpha$ such that $\rho_\alpha(x-y) \neq 0$ (otherwise, we'd have $x-y=0$ since the family of seminorms separates points). Now, let $\eta = \rho_\alpha(x-y)$ and let:

$$\begin{aligned} O_x &= \{z \in X : \rho_\alpha(z-x) < \eta/4\} \\ O_y &= \{z \in X : \rho_\alpha(z-y) < \eta/4\} \end{aligned}$$

By the definition of the locally convex topology, these sets are open. Furthermore, if $z \in O_x \cap O_y$, then:

$$\eta = \rho_\alpha(x-y) \leq \rho_\alpha(x-z)+\rho_\alpha(z-y) = \rho_\alpha(z-x)+\rho_\alpha(z-y) < \eta/4+\eta/4 = \eta/2,$$

which yields an obvious contradiction. Hence, $O_x \cap O_y = \emptyset$. ∎

Convergence of Sequences In this topology, $x_n \to x$ if and only if $\forall \alpha \in A$, $\rho_\alpha(x_n - x) \to 0$.

Definition

- A convex set in a linear space is called *balanced* or *circled* if $x \in C \;\Rightarrow\; \lambda x \in C \;\; \forall \lambda,\ |\lambda| = 1$.
- It is called *absorbing* if
$$\bigcup_{t>0} tC = X.$$

Remark If ρ_α is a family of seminorms on X then the sets

$$\mathcal{N}_{\alpha_1,\ldots,\alpha_N;\epsilon} = \{x \in X : \rho_{\alpha_i}(x) < \epsilon, \quad \forall i = 1,\ldots,N\}$$

are convex, balanced, absorbing sets.

Theorem 3.2.2 *Let X be a linear space with a Hausdorff topology in which addition and scalar multiplication are continuous. Then, X is a locally convex space if and only if 0 has a basis of neighborhoods which are convex, balanced (circled) absorbing sets.*

Proof ($\Longrightarrow$) This follows from the preceding remark.

($\Longleftarrow$) What we want to do here is to build the family of seminorms. Take C to be a convex neighborhood of 0 and let ρ_C be its gauge:

$$\rho_C(x) = \inf\{t > 0 : \frac{x}{t} \in C\}.$$

It is easy to check that ρ_C is a seminorm. Also,

$$\{\rho_C(x) < 1\} \subseteq C \subseteq \{\rho_C(x) \leq 1\}.$$

But that means that the neighborhood basis given by the seminorms is the same as the original neighborhood basis given by the C's. Hence, the two topologies are the same, i.e. the original topology is induced by seminorms. So, the space is locally convex. ∎

Proposition 3.2.3 *Let X be a locally convex vector space. The following are equivalent:*

1. *X is metrizable (the topology is induced by a distance).*
2. *0 has a countable basis of neighborhoods that are convex, balanced, absorbing.*
3. *The topology is generated by a countable family of seminorms.*

Proof
(1) $\Longrightarrow$ (2). Take balls of countable radius (say the rationals).
(2) $\Longrightarrow$ (3). Do the previous construction using gauges.
(3) $\Longrightarrow$ (1). The distance can be given by:

$$d(x,y)=\sum_{n=0}^{\infty}\frac{1}{2^n}\frac{\rho_n(x-y)}{1+\rho_n(x-y)}. \qquad \blacksquare$$

Definition A *Frechet space* is a complete, metrizable locally convex space.

In particular, the Baire Category Theorem applies to Frechet spaces.

Example The Schwartz Class, $\mathcal{S}$ of functions of rapid decrease:

$\mathcal{S}=\{f:\mathbb{R}^n\to\mathbb{C}\ :\ \sup_{x\in\mathbb{R}^n}|x|^{\alpha}|\partial^{\beta}f(x)|<C\ \forall\alpha\in\mathbb{Z},\ \forall\beta$ multi-index of integers $\}$.

For $f\in\mathcal{S}$, define: $\|f\|_{\alpha,\beta}=\sup_x|x|^{\alpha}|\partial^{\beta}f(x)|$

The set $\mathcal{S}^*$ (the dual of $\mathcal{S}$ = the space of all continuous linear functions on $\mathcal{S}$) is called the space of all tempered distributions.

$\mathcal{S}$ is a Frechet space.

Example Let $D(\Omega)=C^{\infty}(\Omega)$ with seminorms given by $\|f\|_{\beta}=\sup_{x\in\Omega}|\partial^{\beta}f(x)|$.

Let $D'(\Omega)$ be $D(\Omega)^*$ = dual of $D(\Omega)$ = space of distributions.

$$\begin{aligned} T\in D'(\Omega) &\iff T \text{ is continuous, linear}\\ &\iff \exists C,n \text{ such that } T(f)\le C\sum_{|\beta|\le n}\|f\|_{\beta}\end{aligned}$$

n is called the *order* of the distribution.

3.3 Weak Topology in Banach Spaces

Definition Let X be a Banach space. The *weak topology on X* is defined as the weakest topology which makes all of the $f\in X^*$ continuous. In other words, it is:

$$=\bigcup_{\text{arbitrary}}\ \bigcap_{\text{finite}} f^{-1}(\mathcal{O}),$$

where $\mathcal{O}$ is open. It is denoted by $\sigma(X,X^*)$.

Note:

- A weakly open set is always strongly open.
- In infinite dimensions, the weak topology is not metrizable.
- A basis of neighborhoods for x_0 is given by sets of the form:

$$\mathcal{N}_{f_1,\ldots,f_N;\epsilon} = \{x \in X : |f_i(x - x_0)| < \epsilon, \quad \forall i = 1, \ldots, N\}$$

Proposition 3.3.1 *The weak topology is Hausdorff*

Proof Let $x \neq y$. Apply the Geometric Version of the Hahn-Banach Theorem to x, y. Then, $\exists f \in X^*$ such that $f(x) < \alpha < f(y)$. So, define:

$$O_1 = f^{-1}((-\infty, \alpha)), \quad O_2 = f^{-1}((\alpha, +\infty))$$

O_1, O_2 are weakly open, they separate x and y and are certainly disjoint. ∎

Remark Given a sequence $\{x_n\}_n$, we distinguish between:

1. $x_n \to x$ strongly means convergence in the X norm.
   i.e. $\|x_n - x\|_X \to 0$.
2. $x_n \rightharpoonup$ means that $x_n \to x$ in the weak topology.
   i.e. $\forall f \in X^*$, $f(x_n) \to f(x)$.

Proposition 3.3.2 *Let $\{x_n\}_n$ be a sequence in X. Then, the following are true:*

1. *$x_n \rightharpoonup x$ if and only if $f(x_n) \to f(x)$ $\forall f \in X^*$.*
2. *If $x_n \to x$, then $x_n \rightharpoonup x$ (The converse is not true, however).*
3. *If $x_n \rightharpoonup x$ then, $\{\|x_n\|_X\}_n$ is bounded and*

$$\|x\|_X \leq \liminf_{n\to\infty} \|x_n\|_X.$$

4. *If $x_n \rightharpoonup x$ and $f_n \to f$ in X^*, then $f_n(x_n) \to f(x)$.*

Proof
(1) This is the definition of weak convergence.

(2) If $x_n \to x$, then:

$$|f(x_n) - f(x)| \leq \|f\|_{X^*}\|x_n - x\|_X \to 0$$

since $\|x_n - x\|_X \to 0$, independent of f. Hence, $f(x_n) \to f(x)$ $\forall f \in X^*$. So, $x_n \rightharpoonup x$.

(3) $\forall f \in X^*$, $\{f(x_n)\}_n$ is bounded. By a corollary of the Uniform Boundedness Principle (Corollary 2.3.3), we deduce that $\{x_n\}_n$ is bounded. So,

$$\begin{aligned} |f(x_n)| &\leq \|f\|_{X^*}\|x_n\|_X \\ |f(x)| &\leq \liminf_{n\to\infty} \|f\|_{X^*}\|x_n\|_X \\ &= \|f\|_{X^*}(\liminf_{n\to\infty} \|x_n\|_X). \end{aligned}$$

But, $\|x\|_X = \sup_{\|f\|_{X^*}\leq 1} |f(x)|$. So,

$$\|x\|_X = \sup_{\|f\|_{X^*}\leq 1} |f(x)| = \sup_{\|f\|_{X^*}\leq 1} \|f\|_X^*(\liminf_{n\to\infty} \|x_n\|_X) \leq \liminf_{n\to\infty} \|x_n\|_X$$

(4)

$$\begin{aligned} |f_n(x_n) - f(x)| &\leq |f_n(x_n) - f(x_n)| + |f(x_n) - f(x)| \\ &\leq \|f_n - f\|_{X^*}\|x_n\|_X \; + \; |f(x_n) - f(x)| \;\longrightarrow\; 0 \end{aligned}$$

since $f_n \to f$ and $f(x_n) - f(x) \to 0$ for all f, by the weak convergence of $\{x_n\}_n$ to x, and since $\{x_n\}_n$ is bounded because of its weak convergence to x. ∎

Proposition 3.3.3 *If* $\dim X < \infty$, *then weak and strong topologies coincide.*

Proof Surely, a weakly open set is strongly open. But is a strongly open set, weakly open? Let U be strongly open with $x_0 \in U$. So, there is $r > 0$ such that $B(x_0, r) \subseteq U$. Let $\{e_1, \ldots, e_n\}$ be a basis for X with $\|e_i\| = 1$. Let $\{f_1, \ldots, f_n\}$ be the dual basis. In other words, $f_j(e_i) = \delta_{i,j}$. The dual basis has the property that if we can expand any $y \in X$ via: $y = \sum f_i(y)e_i$. Then the set

$$\mathcal{N} = \{x \in X \;:\; |f_i(x - x_0)| < \frac{r}{n} \;\forall i = 1, \ldots, n\}$$

is weakly open. So,

$$x \in \mathcal{N} \;\Rightarrow\; \|x - x_0\|_X = \|\sum_{i=1}^{n} f_i(x - x_0)e_i\| \leq \sum_{i=1}^{n} |f_i(x - x_0)| < r.$$

Hence, $\mathcal{N} \subseteq B(x_0, r) \subseteq U$. Thus, U is weakly open. ∎

Example If $\dim X = \infty$, then $S = \{x \in X \;:\; \|x\|_X = 1\}$ is not weakly closed. In fact, its weak closure is $\hat{B}_X = \{x \in X \;:\; \|x\|_X \leq 1\}$.

Proof of this fact Let $x_0 \in B_X$. We will show every weak neighborhood of x_0 intersects S. Take any U of the form:

$$U = \{x \in X \;:\; |f_i(x - x_0)| < \epsilon, \;\forall i = 1, \ldots, n\}.$$

Then, $\exists y_0 \in X$ such that $f_1(y_0) = \ldots = f_n(y_0) = 0$. If not, then the function $x \mapsto (f_1(x), \ldots, f_n(x))$ would be a one-to-one map, meaning that $\dim X \leq n < \infty$. Therefore,

$$\forall t \in \mathbb{R}, \quad f_i((x_0 + ty_0) - x_0) = tf_i(y_0) = 0.$$

Hence, $x_0 + ty_0 \in U$, $\forall t \in \mathbb{R}$. So, take $g(t) = \|x_0 + ty_0\|$. Then, $g(0) = \|x_0\| < 1$, g is continuous, and $g \to \infty$ as $t \to \infty$. Hence, g must take on all the values between $\|x_0\|_X < 1$ and ∞. Hence, $\exists t_0$ such that $g(t_0) = 1$. So, $x_0+t_0y_0 \in S\cap U$. This proves that the weak closure of S contains B. We will later see that it is B, since B is weakly closed by con

Example $B_X = \{x \in X \ : \ \|x\| < 1\}$ is not weakly open. It has empty interior since every weak neighborhood of $x_0 \in B_X$ contains an element of S.

Theorem 3.3.4 *Let $C \subseteq X$ be a convex set. Then, C is weakly closed if and only if C is strongly closed.*

Proof
($\Longrightarrow$) Since weakly open $\Longrightarrow$ strongly open, taking complements, weakly closed $\Longrightarrow$ strongly closed.

($\Longleftarrow$) Assume C is strongly closed. Then, we show that C is weakly closed. i.e., we show that C^c is weakly open. Let $x_0 \in C^c$. By the Hahn Banach Theorem (Second Geometric Form), $\exists f \in X^*$, $\alpha \in \mathbb{R}$ such that $f(x_0) < \alpha < f(x)$, $\forall x \in C$. So, $N = f^{-1}((-\infty, \alpha))$ is a weakly open set (since it is the inverse image of an open set under a continuous function), containing x_0 and included in C^c. Hence, C^c is weakly open. ∎

Corollary 3.3.5 *Let φ be a convex , lower semi-continuous function (for the strong topology). Then, φ is lower semi-continuous for the weak topology. In particular, if $x_n \rightharpoonup x$, then $\varphi(x) \leq \liminf \varphi(x_n)$.*

Proof φ strongly lower semi-continuous $\Longrightarrow$ $\{\varphi(x) \leq \lambda\}$ is convex and strongly closed. $\Longrightarrow$ The set is weakly closed. Hence, φ is weakly l.s.c. ∎

Remark Therefore, convex, strongly continuous $\Longrightarrow$ weakly l.s.c.

Example $x \mapsto \|x\|_X$ is a convex, continuous function. Hence, it is weakly l.s.c. So, if $x_n \rightharpoonup x$ then, $\|x\| \leq \liminf \|x_n\|$ is reproved.

Theorem 3.3.6 *Let X and Y be two Banach spaces and $T : X \to Y$ linear. Then, T is strongly continuous if and only if it is continuous from $\sigma(X, X^*)$ to $\sigma(Y, Y^*)$.*

Proof
($\Longrightarrow$) Assume that T is strongly continuous. Let $f \in Y^*$. So, take any set in $\sigma(Y, Y^*)$ of the form, $f^{-1}((a, b)) \subset Y$. Then, $T^{-1}(f^{-1}((a, b))) = (f \circ T)^{-1}((a, b))$. But, $f \circ T : X \to Y$ is continuous and linear. Hence, $(f \circ T)^{-1}(a, b)$ is open in $\sigma(X, X^*)$, being the inverse image of an open set under a continuous function. Thus, T is weakly continuous.

($\Longleftarrow$) Conversely, assume that T is weakly continuous. $\Gamma(T)$ is weakly closed (i.e.: closed in $\sigma(X \times Y, (X \times Y)^*)$). So, $\Gamma(T)$ is strongly closed. Hence, T is strongly continuous by the Closed Graph Theorem. ∎

3.4 Weak-* Topologies $\sigma(X^*, X)$

On X^* we can define the weak topology, $\sigma(X^*, X^{**})$. But, $X \subseteq X^{**}$. So, technically, there is something even weaker than the weak topology.

Definition The *weak-* topology* on X^* is defined as the weakest topology which makes all the maps $f \mapsto f(x)$ continuous. $\sigma(X^*, X)$ is weaker than $\sigma(X^*, X^{**})$.

Proposition 3.4.1 *$\sigma(X^*, X)$ is Hausdorff.*

Proof If $f_1, f_2 \in X^*$ and $f_1 \neq f_2$, then $\exists x \in X$ such that $f_1(x) \neq f_2(x)$. So, $\exists \alpha$ such that $f_1(x) < \alpha < f_2(x)$. So, define the following sets:

$$O_1 = \{f \in X^* \ : \ f(x) < \alpha\}, \quad O_2 = \{f \in X^* \ : \ f(x) > \alpha\}.$$

O_1 and O_2 are open in $\sigma(X^*, X)$ and separate f_1 and f_2. ∎

A basis of neighborhoods of f_0 for $\sigma(X^*, X)$ is given by sets of the form:

$$\mathcal{N}_{x_1,\ldots,x_n;\epsilon} = \{f \in X^* \ : \ |(f - f_0)(x_i)| < \epsilon, \ \forall i = 1, \ldots, n\}$$

We say, $f_n \overset{*}{\rightharpoonup} f$ (f_n converges weakly-* to f) if $f_n \to f$ in $\sigma(X^*, X)$. In other words, $\forall x \in X, \ f_n(x) \to f(x)$.

Properties

1. $f_n \overset{*}{\rightharpoonup} f \iff \forall x \in X, f_n(x) \to f(x)$.
2. $f_n \to f$ in X^*
   $\Longrightarrow f_n \rightharpoonup f$ in $\sigma(X^*, X^{**})$
   $\Longrightarrow f_n \overset{*}{\rightharpoonup} f$ in $\sigma(X^*, X)$
3. If $f_n \overset{*}{\rightharpoonup} f$, then $\|f_n\|_{X^*}$ bounded and $\|f_n\|_{X^*} \leq \liminf \|f_n\|_{X^*}$.
4. If $f_n \overset{*}{\rightharpoonup} f$, and $x_n \to x$ in X, then $f_n(x_n) \longrightarrow f(x)$.

Theorem 3.4.2 (Banach-Alaoglu) *Let X^* be the dual of a Banach space. Then,*

$$B_{X^*} = \{f \in X^* \ : \ \|f\|_{X^*} \leq 1\}$$

is compact for the weak- topology.*

Remark Observe right-away that compactness $\not\Rightarrow$ sequential compactness. It is only true if the space is metrizable.

Compare this to Riesz' Theorem which states that the unit ball of a Banach space is strongly compact if and only the dimension is finite.

Proof of Banach-Alaoglu's Theorem Tychonoff's Theorem states that any product of compact spaces is compact for the product topology.

So, apply Tychonoff's Theorem to:

$$A = \prod_{x \in X} B(0, \|x\|_X)$$

A is therefore compact for the product topology.

Elements of A are assignments $x \mapsto g(x)$. So, they are functions of x which satisfy $|g(x)| < \|x\|_X$. Let $\tilde{A}$ be the subset of A containing all linear functions. So, we can write:

$$\tilde{A} = \bigcap_{x,y \in X} A_{x,y} \times \bigcap_{x \in X, \lambda \in \mathbb{K}} B_{\lambda,x},$$

where:

$$\begin{aligned} A_{x,y} &= \{f \in A \ : \ f(x+y) - f(x) - f(y) = 0\} \\ B_{x,\lambda} &= \{f \in A \ : \ f(\lambda x) - \lambda f(x) = 0\}. \end{aligned}$$

These are closed in the product topology. So, $\tilde{A}$ is a closed subset of a compact set and so, $\tilde{A}$ is compact for the product topology. But, the product topology on $\tilde{A}$ is the weak-* topology. Hence, $\tilde{A} = B_{X^*}$ is compact in $\sigma(X^*, X)$. ∎

3.5 Reflexive Spaces

Definition X is said to be reflexive if $X^{**} = X$.

Theorem 3.5.1 (Kakutani) *Let X be a Banach Space. Then, the closed unit ball, $B_X = \{x \in X : \|x\| \leq 1\}$ is compact for the weak topology $\sigma(X, X^*)$ if and only if X is reflexive.*

Before we prove Kakutani's Theorem, we need several lemmas by Helly and Goldstein.

Lemma 3.5.2 (Helly) *Let X be a Banach space, $f_1, \ldots, f_n \in X^*$ and $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$. Then, the following conditions are equivalent:*

1. $\forall \epsilon > 0 \ \exists x_\epsilon \ \|x_\epsilon\| \leq 1$ *such that:*

$$| < f_i, x_\epsilon > - \alpha_i| < \epsilon \ \ \forall i = 1, \ldots, n.$$

2. $\forall \beta_i, \ \ |\sum_{i=1}^n \beta_i \alpha_i| \leq \|\sum_{i=1}^n \beta_i f_i\|_{X^*}$.

Proof (1) $\Longrightarrow$ (2): From (1), we get that $\forall \beta_i$,

$$\left| \sum_{i=1}^n \beta_i < f_i, x_\epsilon > - \sum_{i=1}^n \beta_i \alpha_i \right| < \epsilon \sum_{i=1}^n |\beta_i|$$

$$\begin{aligned}
\left|\sum_{i=1}^{n}\beta_i\alpha_i\right| &\leq \left|\sum_{i=1}^{n} < \beta_i f_i, x_\epsilon >\right| + \epsilon\sum_{i=1}^{n}|\beta_i| \\
&= \left|\left\langle\sum_{i=1}^{n}\beta_i f_i, x_\epsilon\right\rangle\right| + \epsilon\sum_{i=1}^{n}|\beta_i| \\
&\leq \left\|\sum_{i=1}^{n}\beta_i f_i\right\|_{X^*} \|x_\epsilon\|_X + \epsilon\sum_{i=1}^{n}|\beta_i|
\end{aligned}$$

But, $\|x_\epsilon\|_X \leq 1$. So, let $\epsilon \to 0$. Then,

$$\left|\sum_{i=1}^{n}\beta_i\alpha_i\right| \leq \left\|\sum_{i=1}^{n}\beta_i f_i\right\|_{X^*}.$$

(2) $\Longrightarrow$ (1): Assume not. Then, let $\vec{\varphi}(x) = (< f_1, x >, \ldots, < f_n, x >)$. Then, $(\alpha_1, \ldots, \alpha_n) \notin \overline{\vec{\varphi}(B_X)}$. Since $\{(\alpha_1, \ldots, \alpha_n)\} = \{\alpha\}$ is a compact set and $\overline{\vec{\varphi}(B_X)}$ is closed and convex, we can apply the Hahn-Banach Theorem and say that $\exists\gamma$ and $\vec{\beta}$ such that $\vec{\beta}\cdot\vec{\alpha} > \gamma > \vec{\beta}\cdot\vec{\varphi}(x)$ $\forall x \in B_X$. So,

$$\forall x \in B_X, \quad \sum_{i=1}^{n}\beta_i\alpha_i > \gamma > \sum_{i=1}^{n}\beta_i < f_i, x >.$$

Changing x to $-x$ above, we get that:

$$\left\|\sum_{i=1}^{n}\beta_i f_i(x)\right\| < \gamma < \left|\sum_{i=1}^{n}\beta_i\alpha_i\right|.$$

Taking the sup over $x \in B_X$:

$$\left\|\sum_{i=1}^{n}\beta_i f_i\right\|_{X^*} \leq \gamma < \left|\sum_{i=1}^{n}\beta_i\alpha_i\right|,$$

contradicting the assumption made in (2). ∎

Lemma 3.5.3 (Goldstine) *$J(B_X)$ is dense in $B_{X^{**}}$ for $\sigma(X^{**}, X^*)$. Here, $J : X \to X^{**}$, $J(x) =< x, \cdot >$.*

Proof We prove that for every $\eta \in B_{X^{**}}$, every neighborhood of η for $\sigma(X^{**}, X^*)$ intersects $J(B_X)$.

So, take $\eta \in B_{X^{**}}$. We can assume that the neighborhood is:

$$\{\zeta \in X^{**} \ : \ |< \zeta - \eta, f_i >| < \epsilon, \ f_i \in X^*, \ i = 1, \ldots, n\}.$$

So, is there $x \in B_X$ such that $|< x - \eta, f_i >| < \epsilon$ for $i = 1, \ldots, n$? This is equivalent to asking is there $x \in B_X$ such that $|< f_i, x > - < \eta, f_i >| < \epsilon$ for each i? Let $\alpha_i =< \eta, f_i >$. By Helly's Lemma, this can only happen if and only if $|\sum_{i=1}^n \beta_i\alpha_i| \leq \|\sum_{i=1}^n \beta_i f_i\|_{X^*}$. Since, $\eta \in B_{X^{**}}$,

$$\forall \beta_i, \quad \left|\sum_{i=1}^n < \beta_i f_i, \eta >\right| \leq \left\|\sum_{i=1}^n \beta_i f_i\right\|_{X^*}.$$

But, $\alpha_i =< \eta, f_i >$. So, we have that:

$$\left|\sum_{i=1}^n \beta_i\alpha_i\right| = \left|\sum_{i=1}^n \beta_i < f_i, \eta >\right| = \left|\sum_{i=1}^n < \beta_i f_i, \eta >\right| \leq \left\|\sum_{i=1}^n \beta_i f_i\right\|_{X^*}. \quad \blacksquare$$

We are now ready to prove Kakutani's Theorem.

Proof of Kakutani's Theorem
($\Longleftarrow$) If X is reflexive, then apply the Banach-Alaoglu Theorem to X^*. Since $X = (X^*)^*$, the result follows

($\Longrightarrow$) We must show that $X^{**} = X$. But, this is equivalent to showing that $J(B_X) = B_{X^{**}}$ by linearity of J. By Theorem 3.3.6, if T is a linear operator, then it is strong-strong continuous if and only if it is weak-weak continuous.

Hence, J is Continuous from $\sigma(X, X^*)$ to $\sigma(X^{**}, X^{***})$. This is more demanding that J being continuous from $\sigma(X, X^*)$ to $\sigma(X^{**}, X^*)$ since $X^{***} \supseteq X^*$. Therefore, J is continuous $\sigma(X, X^*)$ to $\sigma(X^{**}, X^*)$. Since $J(B_X)$ is compact for $\sigma(X^{**}, X^*)$, it is closed. So, by Goldstein's lemma, $J(B_X)$ is dense in $B_{X^{**}}$ and closed. Hence, $J(B_X) = B_{X^{**}}$. This proves that $J(X) = X^{**}$. Hence, X is reflexive. $\blacksquare$

Corollary 3.5.4 *If M is a closed subspace of a reflexive space X, then M is reflexive.*

Proof B_M is a weakly closed subset of the compact set B_X because it's convex. Hence, B_M is weakly compact. Hence, M is reflexive. $\blacksquare$

Corollary 3.5.5 *Let X be a reflexive Banach space. If C is a closed (strong or weak), convex, bounded set, then C is compact for $\sigma(X, X^*)$.*

Proof C is weakly closed and $C \subseteq mB_X$ for some $m > 0$. Since mB_X is compact for $\sigma(X, X^*)$, C is compact for $\sigma(X, X^*)$ as well. $\blacksquare$

Proposition 3.5.6 *Let X be a reflexive Banach space, and $\varphi \not\equiv +\infty$, a convex, lower semi-continuous function from a closed, convex set A to $(-\infty, +\infty]$ such that either A is bounded or $\lim_{x\in A, \|x\|\to\infty} \varphi(x) = +\infty$. Then, φ achieves its minimum on A.*

Proof

Property A lower semi-continuous function achieves its min on a compact set.

Let $\lambda = \varphi(x_0) < \infty$ for some x_0. Then, if we define:

$$\tilde{A} = \{x \in A : \ \varphi(x) \leq \lambda\},$$

is a convex, strongly closed set (because φ is l.s.c.). Hence, $\tilde{A}$ is weakly closed. Since it's bounded by assumption, it is weakly compact. Hence, since φ is convex, l.s.c. it is weakly l.s.c.. Hence, φ achieves its minimum on $\tilde{A}$. Since $\tilde{A} \subset A$, φ achieves its minimum on A. ∎

3.6 Separable Spaces

Definition X is separable if X has a countable, dense subset.

Theorem 3.6.1 *B_{X^*} is metrizable for the $\sigma(X^*, X)$ topology if and only if X is separable, with the metric given by:*

$$d(f, g) = \sum_{n=1}^{\infty} \frac{1}{2^n} |< f - g, x_n >|,$$

where the $\{x_n\}_n$ is the countable dense set in X.

Remark B_{X^*} is metrizable by not X^*

Corollary 3.6.2 *Let X be separable. Let $\{f_n\}_n$ be a bounded sequence in X^*. Then, there exists a subsequence $\{f_{n_k}\}_k$ converging weakly-*.*

Proof We assume WLOG that $\{f_n\}_n \subset B_{X^*}$. By Banach-Alaoglu, B_{X^*} is weakly-* compact. Since B_{X^*} is metrizable, by Theorem 3.6.1, we have that B_{X^*} is sequentially compact. ∎

Proposition 3.6.3 *Let X be a reflexive space and $\{x_n\}_n$ a bounded sequence in X. Then, there exists a subsequence $\{x_{n_k}\}_k$ which converges in $\sigma(X, X^*)$.*

Proof X reflexive $\implies$ B_X is compact. Let $M = Span\{x_1, x_2, \ldots\}$. Then, $\overline{M}$ is a separable Banach space, which is also reflexive. So, $B_{\overline{M}}$ is compact for $\sigma(X, X^*)$. Hence, we may extract a convergent subsequence. ∎

Remark These two results show that for a reflexive space X, B_X is both compact and sequentially compact.

3.7 Applications

3.7.1 L^p Spaces

For $1 < p < \infty$, the dual of L^p is $L^{p'}$ where $1/p + 1/p' = 1$. So, what is weak convergence in L^p? Answer:

$$f_n \rightharpoonup f \text{ in } L^p \iff \forall g \in L^{p'}, \int f_n g \to \int f g.$$

Recall that the definition of strong L^p convergence is $\int |f_n - f|^p \to 0$.

Example

- Consider $\{f_n(x) = \sin nx\}_{n\in\mathbb{Z}}$ on $[0,1]$. Then, $\forall g \in C^\infty([0,1])$,

  $$\int_0^1 \sin nx g(x)\, dx \to 0.$$

  Since C^∞ is dense in $L^{p'}$, we see that $\sin nx \rightharpoonup 0$ weakly in L^p.

- On the other hand $\exists C > 0$ such that $\forall n$,

  $$\int_0^1 |\sin nx|^2\, dx = C.$$

  Hence, $\{f_n\}_n \not\to 0$ strongly in L^p.

Example For $1 < p < \infty$, L^p is reflexive and separable. Hence, the unit ball B_1 is weakly and weakly sequentially compact.

Example

- $(L^1)^* = L^\infty$, but $(L^\infty)^* \supsetneq L^1$
  (In fact, $(L^\infty)^* = \{$ Bounded Radon Measures $\}$).
- Neither L^∞ nor L^1 is reflexive.
- L^1 is separable, but L^∞ is not.
- L^1 is not the dual of any space.
- B_{L^1} is not even weakly closed. Hence, it's not weakly compact (Take approximate identities and you'll see that $\overline{B_{L^1}} = B_{\{\text{ measures }\}}$.
- B_{L^∞} is weak-* compact, but not weakly compact by Kakutani's Theorem (since L^∞ is not reflexive.

3.7.2 PDE's

Suppose we wish to solve the following PDE:

$$\begin{cases} \triangle u + |u|\cdot u + u = f & \text{on } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases} \tag{3.1}$$

where Ω is a bounded open set in $\mathbb{R}^2$, and f is smooth.

Method of Calculus of Variations: We want to minimize the *energy*:

$$F(u) = \frac{1}{2}\int_\Omega |\nabla u|^2 + \frac{1}{3}\int_\Omega |u|^3 + \frac{1}{2}\int_\Omega |u|^2 - \int_\Omega fu.$$

Assume that u minimizes F. Then, $\forall g \in C_0^\infty(\Omega)(g = 0 \text{ on } \partial\Omega)$, if we set $\varphi(t) = F(u + tg)$,

$$\frac{d}{dt}\Big|_{t=0} \varphi(t) = 0, \quad \text{since } \varphi(t) \geq \varphi(0),\ \forall t.$$

$$\begin{aligned}
F(u+tg) &= \frac{1}{2}\int_\Omega |\nabla(u+tg)|^2 + \frac{1}{3}\int_\Omega |u+tg|^3 + \frac{1}{2}\int_\Omega |u+tg|^2 + \\
&\qquad + \int_\Omega f(u+tg) \\
&= \frac{1}{2}\int_\Omega (|\nabla u|^2 + 2t\nabla u \nabla g) + \frac{1}{3}\int_\Omega (|u|^3 + 3tg|u| \cdot u) \\
&+ \frac{1}{2}\int_\Omega (|u|^2 + 2tgu) - \int_\Omega (fu + tg + O(t^2)) \\
0 &= \frac{d}{dt}\Big|_{t=0} F(u+tg) = \int_\Omega (\nabla u \cdot \nabla g + g|u| \cdot u + ug - fg)
\end{aligned}$$

Integrating the first term by parts and noticing that g vanishes on $\partial\Omega$, we see that:

$$\begin{aligned}
0 &= \int_\Omega [(-\triangle u)g + g|u| \cdot u + ug - fg] \\
&= \int_\Omega [-\triangle u + |u| \cdot u + u - f]\, g
\end{aligned}$$

Since g was an arbitrary member of C_0^∞, we have that u solves 3.1 in the sense of distributions.

So, to summarize, the minimizer of the equation:

$$F(u) = \frac{1}{2}\int_\Omega |\nabla u|^2 + \frac{1}{3}\int_\Omega |u|^3 + \frac{1}{2}\int_\Omega |u|^2 - \int_\Omega fu.$$

gives a weak solution to:

$$\begin{cases} \triangle u + |u| \cdot u + u = f & \text{on } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases}$$

So, now define F to be a function on the Sobolev space $H_0^1(\Omega)$. First, the norm on $H_0^1(\Omega)$ is given by:

$$\|u\|_{H_0^1(\Omega)} = \int_\Omega (|\nabla u|^2 + |u|^2).$$

So, with that norm, $H_0^1(\Omega)$ becomes the closure of $C_0^\infty(\Omega)$. Alternatively, we can define $H_0^1(\Omega)$ to be the set of functions $u \in L^2(\Omega)$ whose weak derivative ∇u is also in $L^2(\Omega)$ and $u = 0$ on $\partial\Omega$.

Fact: If $\dim = 1$, $H^1 \subseteq C^0$.

Proof Let $u \in H_0^1$. Then, $u(x) - u(y) = \int_x^y u'(t)dt$. Hence,

$$|u(x) - u(y)| \leq \left| \int_x^y u'(t)dt \right| \leq \sqrt{y-x} \sqrt{\int_x^y |u'(t)|^2 dt} \leq C\|u\|_{H_0^1(\Omega)}.$$

by Cauchy-Schwartz. Hence, $u \in C^{0,\frac{1}{2}}$, the set of Hölder continuous functions with Hölder exponent $\frac{1}{2}$. ∎

Fact: If dim $= 2$, then H^1 is not a subset of C^0. However, in any any dimension, there is a continuous embedding of H^1 into L^p for $p < \infty$ and $p \leq p^*$ where $1/p^* = 1/2 - 1/d$. We call p^* the critical exponent.. By continuous embedding, we mean that $\|u\|_{L^p} \leq C\|u\|_{H^1}$.

Example So, for example, in dim $= 2$, $p^* = \infty$. Hence, $H^1 \subseteq L^p, \ \forall p < \infty$. In dim $= 3$, we have: $H^1 \subseteq L^p, \ \forall p \leq 6$.

Remark In fact, the embedding $H^1 \subseteq L^p$ is compact (the embedding is a compact operator). This means that $B(0,1)$ in H^1 is mapped into a compact set in L^p. This transforms weak convergence into strong convergence. In other words, if $u_n \rightharpoonup u$ weakly in H^1, then $u_n \to u$ strongly in $L^p, \quad \forall p < p$.

Returning back to the problem, we still have not answered whether $\min F$ is achieved. First, we check that F is *coercive*, i.e. $F \to \infty$ as $\|u\|_{H_0^1(\Omega)} \to \infty$, so that: $\{u : \ F(u) \leq 1\}$ is bounded and nonempty (since $F(0) = 0$).

Proof that F is Coercive

$$F(u) = \frac{1}{2}\int_\Omega |\nabla u|^2 + \frac{1}{3}\int_\Omega |u|^3 + \frac{1}{2}\int_\Omega |u|^2 - \int_\Omega fu.$$

for $\Omega \subseteq \mathbb{R}^2$. If $u \in H_0^1(\Omega)$, then $u \in L^p(\Omega)$ for all $p < \infty$. In particular, $u \in L^3(\Omega)$. So,

$$F(u) \geq \frac{1}{2}\|u\|^2_{H_0^1(\Omega)} - \|f\|_{L^2(\Omega)}\|u\|_{L^2(\Omega)},$$

where $\|u\|_{L^2(\Omega)} \leq \|u\|_{H_0^1(\Omega)}$. Thus,

$$F(u) \geq \frac{1}{4}\|u\|^2_{H_0^1(\Omega)} + \underbrace{\frac{1}{4}\|u\|^2_{H_0^1(\Omega)} - \|f\|_{L^2(\Omega)}\|u\|_{H_0^1(\Omega)}},$$

where the grouped terms are bounded from below by $-\|f\|_{L^2(\Omega)}$. To see this, consider the function $x \mapsto x^2/4 - \|f\|_{L^2(\Omega)}x$, which has a minimum at $x = 2\|f\|_{L^2(\Omega)}$ and takes the value $-\|f\|_{L^2(\Omega)}$ there . Thus, $F(u) \geq \frac{1}{4}\|u\|^2_{H_0^1(\Omega)} - C$ with C independent of u. Therefore, $F(u) \to \infty$ as $\|u\|_{H_0^1(\Omega)} \to \infty$. ∎

Lemma 3.7.1 *F is weakly l.s.c.*

Proof Note that the following functions are all strongly continuous and convex:

$$\begin{aligned} u &\mapsto \frac{1}{4}\int_\Omega (|\nabla u|^2 + |u|^2) \\ u &\mapsto \frac{1}{3}\int_\Omega |u|^3 \\ u &\mapsto -\int_\Omega fu \qquad \blacksquare \end{aligned}$$

Hence, F is weakly lower semi-continuous and convex. Therefore, by Proposition 3.5.6, F achieves its min on $H_0^1(\Omega)$. $\blacksquare$

Chapter 4

Bounded (Linear) Operators and Spectral Theory

4.1 Topologies on Bounded Operators

Let X, Y be Banach spaces and denote by $\mathcal{L}(X,Y)$ to be the space of bounded operators from X to Y, with the norm given by:

$$\|T\|_{\mathcal{L}(X,Y)} = \sup_{\|x\|_X \leq 1} \|Tx\|_Y$$

Definition The topology on $\mathcal{L}(X,Y)$ defined by this norm is called the *uniform topology*. In that topology, $(A,B) \mapsto AB$ is jointly continuous.

Definition We define the *strong topology* as the weakest topology which makes all the:

$$E_x : \mathcal{L}(X,Y) \longrightarrow Y, \quad T \mapsto Tx$$

continuous ($\forall x \in X$). It's the topology of pointwise convergence. However, in this topology, multiplication, $(A,B) \mapsto AB$ is separately continuous, but not jointly continuous.

Definition We define the *weak operator topology* as the weakest topology which makes all of the:

$$E_{x,l} : (X,Y) \longrightarrow \mathbb{C}, \quad T \mapsto < l, Tx >$$

for $x \in X,\ l \in Y^*$, continuous.

Remark It is akin to the convergence of all n matrix entries $< l, Tx >$ of T. So, we write:

$$T_n \xrightarrow{w} T, \quad \text{if } \forall l \in Y^*,\ \forall x \in X, \quad < l, T_n x > \longrightarrow < l, Tx > .$$

uniform $>$ strong $>$ weak.

Example

- Bounded operators on $l_2 = \left\{ \{u_n\}_n : \sum |u_n|^2 < \infty \right\}$ given by:

$$T_n : (u_1, u_2, \ldots) \mapsto \left(\frac{u_1}{n}, \frac{u_2}{n}, \ldots \right).$$

It is not difficult to see that $T_n \longrightarrow 0$ uniformly.

- Consider the deletion operators on l_2 :

$$S_n : (u_1, \ldots, u_n, \ldots) \mapsto (\underbrace{0, 0, \ldots, 0}_{n \text{ times}}, u_{n+1}, u_{n+2}, \ldots)$$

Clearly, $S_n \longrightarrow 0$ strongly. However, $S_n \not\to 0$ uniformly. To see this, fix $n > 0$ consider a sequence u whose l_2 norm is 1, such that $u_i = 0$ for all $i \leq n$. Then, $S_n(u) = u$. Hence, $\|S_n\|_{\mathcal{L}(X)} \geq 1$. On the other hand, for any $u \in l_2$ with l_2 norm 1, $\|S_n(u)\|_{l_2} \leq \|u\|_{l_2}$. Hence, $\|S_n\|_{\mathcal{L}(X)} = 1$. Since n was arbitrary, $\|S_n\|_{\mathcal{L}(X)} = 1$ for all n,

- Now, consider the shift operators W_n given by:

$$W_n : (u_1, u_2, \ldots) \mapsto (\underbrace{0, \ldots, 0}_{n \text{ times}}, u_1, u_2, \ldots)$$

To see that $W_n \longrightarrow 0$ weakly, consider any functional $f : l_2 \to \mathbb{R}$. Then, for any $u \in l_2$,

$$< f, W_n(u) >= f(\underbrace{0, \ldots, 0}_{n \text{ times}}, u_1, u_2, \ldots)) \to 0.$$

On the other hand, it is clear that for any $u \in l_2$, $\|W_n(u)\|_{l_2} = \|u\|_{l_2}$. Hence, $\|W_n\|_{\mathcal{L}(X)} = 1$ for each n. Hence, $W_n \not\to 0$ strongly.

Theorem 4.1.1 *Let H be a Hilbert space and $T_n \in \mathcal{L}(H)$ such that $\forall x, y \in H$, $\langle T_n x, y \rangle_H$ converges as $n \to \infty$, then $\exists T \in \mathcal{L}(H)$ such that $T_n \to T$ in the weak topology.*

Proof Given x, $\forall y \in Y$, $\sup_n |\langle T_n x, y \rangle| < \infty$. Hence, by the Uniform Boundedness Principle,

$$\sup_{\|y\|_H \leq 1} \sup_n |\langle T_n x, y \rangle| < \infty \iff \sup_n \|T_n x\|_H < \infty.$$

This is true for any $x \in H$. So, again applying the Uniform Boundedness Principle, we see that $\sup_n \|T_n\|_{\mathcal{L}(H)} < \infty$. Now, we define $B(x, y) = \lim_{n\to\infty} \langle T_n x, y \rangle$. One can see that B is sesquilinear. Furthermore,

$$|B(x, y)| \leq \limsup_n \|T_n\|_{\mathcal{L}(H)} \|x\|_H \|y\|_H \leq C\|x\|_H \|y\|_H$$

Therefore, by a corollary of the Riesz Representation Theorem (not proven in class), $\exists T \in \mathcal{L}(H)$ such that $B(x, y) = \langle T_n x, y \rangle$. Then, it is easy to see that $T_n \to T$ weakly. ∎

4.2 Adjoint

Definition If $T \in \mathcal{L}(X,Y)$, where X, Y are Banach spaces, the *adjoint*, $T' \in \mathcal{L}(Y^*, X^*)$ defined by:

$$T'(l) = l(Tx) \quad \text{or} \quad < l, Tx > = < T'l, x >$$

Theorem 4.2.1 *Let X, Y be Banach spaces and $T \in \mathcal{L}(X,Y)$. Then, the map given by $T \mapsto T'$ is a linear, isometric isomorphism.*

Proof

$$\begin{aligned}
\|T\|_{\mathcal{L}(X,Y)} &= \sup_{\|x\|_X \leq 1} \|Tx\|_Y = \sup_{\|x\|_X \leq 1} \sup_{\|l\|_{Y^*} \leq 1} |\langle l, Tx \rangle| \\
&= \sup_{\|l\|_{Y^*} \leq 1} \sup_{\|x\|_X \leq 1} |\langle T'l, x \rangle| \\
&= \sup_{\|l\|_{Y^*} \leq 1} \|T'l\|_{X^*} \\
&= \|T'\|_{\mathcal{L}(X,Y)}
\end{aligned}$$

This shows the isometry part. Linearity and isomorphism are both trivial. ∎

If H is a Hilbert space, and C is the canonical isomorphism taking H to H^*, we define the Hilbert space adjoint of $T \in \mathcal{L}(H)$ as $T^* = C^{-1}T'C$ where T' is the Banach space adjoint. With this association, $T^* \in \mathcal{L}(H)$. Equivalently, we can write this relation in the more familiar manner:

$$\forall x, y \in H, \quad \langle x, Ty \rangle = \langle T^*x, y \rangle .$$

It follows that $\|T\| = \|T^*\|$. In fact, we have the following properties:

- $T \mapsto T^*$ is an isomorphism with $(\alpha T)^* = \overline{\alpha} T^*$.
- $(TS)^* = S^*T^*$.
- $\left(T^{-1}\right)^* = (T^*)^{-1}$.

The map $T \mapsto T^*$ is continuous in the uniform and weak topologies, but not in the strong.

Counterexample Shift in l_2 :

$$W_n : (u_1, u_2, \ldots) \mapsto (\underbrace{0, \ldots, 0}_{n \text{ times}}, u_1, u_2, \ldots)$$

So what is the adjoint of W_n?

$$\langle v, W_n u \rangle = \sum_{i=1}^{\infty} \overline{v_{n+i}} u_i = \langle V_n v, u \rangle$$

with $V_n(v_1, v_2, \ldots) = (v_{n+1}, v_{n+2}, \ldots)$. Thus, $W_n^* = V_n$. $V_n \longrightarrow 0$ in the strong topology, but $W_n = V_n^* \not\to 0$ strongly.

Note: $\|T^*T\|_{\mathcal{L}(H)} = \|T\|^2_{\mathcal{L}(H)}$.

Definition

- An operator $T \in \mathcal{L}(H)$ is *self-adjoint* if $T^* = T$.
- An operator P is a *projection* if $P^2 = P$.
- A projection P is *orthogonal* if $P^* = P$.

4.3 Spectrum

Definition Let X be a Banach space, $T \in \mathcal{L}(X)$.

- The *resolvent set* of T, denoted $\rho(T)$ is the set of scalars $\lambda \in \mathbb{R}$ (or $\mathbb{C}$) s.t. $\lambda I - T$ is bijective with a bounded inverse.
- If $\lambda \in \rho(T)$, then $R_\lambda(T) = (\lambda I - T)^{-1}$ is called the *resolvent* of T (at λ).
- If $\lambda \notin \rho(T)$, then λ is in the "*spectrum of* T" $= \sigma(T)$.

Note: From the Open Mapping Theorem, if $\lambda I - T$ is bijective, then its inverse is continuous

Definition

1. $\lambda \in \sigma(T)$ is said to be an *eigenvalue* of T if $\ker(\lambda I - T) \neq \{0\}$

   OR $\lambda I - T$ is not injective

   OR $\exists x \neq 0$ such that $Tx = \lambda x$. If this is the case, we say that x is an *eigenvector*.

   The set of eigenvalues is called the *point spectrum* of T.

2. $\lambda \in \sigma(T)$ which is not an eigenvalue and for which $R(\lambda I - T)$ is not dense is said to be in the *residual spectrum* of T.

In fact, we can draw the following diagram to describe the relationship among the various parts of the spectrum.

	$\nearrow$	Point Spectrum = injectivity violated
Bijectivity violated = Spectrum	$\longrightarrow$	Residual Spectrum = injectivity OK, surjectivity *too* violated
	$\searrow$	other = injectivity OK, surjectivity *slightly* violated

Note: In infinite dimensions, injective $\not\Rightarrow$ bijective since there's no pigeonhole principle.

Theorem 4.3.1 *Let X be a Banach space and $T \in \mathcal{L}(X)$. Then, $\rho(T)$ is open, and $R_\lambda(T) = (\lambda I - T)^{-1}$ is an $\mathcal{L}(X)$-valued analytic function of λ on $\rho(T)$. Moreover, $\forall \lambda, \mu \in \rho(T)$, $R_\lambda(T)$ and $R_\mu(T)$ commute and*

$$R_\lambda(T) - R_\mu(T) = (\mu - \lambda) R_\lambda(T) R_\mu(T).$$

Proof Let $\lambda_0 \in \rho(T)$. Formally, if T were to be taken as a real number, we could write:

$$\begin{aligned} \frac{1}{\lambda - T} &= \frac{1}{\lambda_0 - T + \lambda - \lambda_0} = \frac{1}{(\lambda_0 - T)\left(1 + \frac{(\lambda - \lambda_0)}{(\lambda_0 - T)}\right)} \\ &= \frac{1}{\lambda_0 - T} \sum_{n=0}^{\infty} \frac{(\lambda_0 - \lambda)^n}{(\lambda_0 - T)^n} \end{aligned}$$

Inspired by this calculation, we set

$$\tilde{R}_\lambda(T) = R_{\lambda_0}(T) \sum_{n=0}^{\infty} [R_{\lambda_0}(T)]^n (\lambda - \lambda_0)^n.$$

This series converges absolutely, since:

$$\sum_{n=0}^{\infty} \|R_{\lambda_0}(T)^n\| |(\lambda - \lambda_0)^n| \le \sum_{n=0}^{\infty} \|R_{\lambda_0}(T)\|^n |(\lambda - \lambda_0)|^n$$

if $|\lambda - \lambda_0| \|R_{\lambda_0}(T)\| < 1$. That is, in $B\left(\lambda_0, \frac{1}{\|R_{\lambda_0}\|}\right)$, we can define $\tilde{R}_\lambda(T)$ and

$$\tilde{R}_\lambda(T)(\lambda I - T) = (\lambda I - T)\tilde{R}_\lambda(T) = I.$$

Hece, $\tilde{R}_\lambda(T) = R_\lambda(T)$ and $B\left(\lambda_0, \frac{1}{\|R_{\lambda_0}\|}\right) \subseteq \rho(T)$. This proves that $\rho(T)$ is open and that $R_\lambda(T)$ is analytic in λ with coefficients in $\mathcal{L}(X)$, since we just wrote a representation for $\tilde{R}_\lambda(T)$ in this way. Moreover, to show commutativity and the last identity, note the following:

$$\begin{aligned} R_\lambda(T) - R_\mu(T) &= R_\lambda(T) \underbrace{(\mu I - T) R_\mu(T)}_{=I} - \underbrace{R_\lambda(T)(\lambda I - T)}_{=I} R_\mu(T) \\ &= (\mu - \lambda) R_\lambda(T) R_\mu(T). \end{aligned}$$

Similarly, $R_\lambda(T) - R_\mu(T) = -(R_\mu(T) - R_\lambda(T)) = (\mu - \lambda) R_\mu(T) R_\lambda(T)$. This shows that $R_\lambda(T) R_\mu(T) = R_\mu(T) R_\lambda(T)$. ∎

Theorem 4.3.2 *Let X be a Banach space and $T \in \mathcal{L}(X)$. Then, $\sigma(T)$ is closed, non-empty and included in $\overline{B}(0, \|T\|_{\mathcal{L}(X)})$*

Remark This shows that the spectrum is a non-empty compact subset of a disk.

Proof

- Formally, for any λ, we can write:

$$\frac{1}{\lambda I - T} = \frac{1}{\lambda}\left(\frac{1}{1-\frac{T}{\lambda}}\right) = \frac{1}{\lambda}\sum_{n=0}^{\infty}\frac{T^n}{\lambda^n}.$$

  If $|\lambda| > \|T\|_{\mathcal{L}(X)}$, then $\frac{1}{\lambda}\sum_n T^n/\lambda^n$ converges absolutely and provides and inverse to $\lambda I - T$ (one can just check by multiplying on right and left to get the identity). Hence, if $\lambda > \|T\|_{\mathcal{L}(X)}$, then $\lambda \in \rho(T)$ and $R_\lambda(T) = \frac{1}{\lambda}\sum_n T^n/\lambda^n$. Hence, $\sigma(T) \subset \overline{B}(0, \|T\|_{\mathcal{L}(X)})$.

- The fact that the spectrum is closed is clear from the previous theorem.

- If $\sigma(T)$ were empty, then $R_\lambda(T)$ would be an analytic function on $\mathbb{C}$ and $\lim_{|\lambda|\to\infty} R_\lambda(T) = 0$. Hence, R_λ must be constant in λ (by Liouville's Theorem). Hence, $\forall\lambda, R_\lambda(T) = 0$. This is a contradiction. Hence, $\sigma(T) \neq \emptyset$. ∎

Definition The *spectral radius* of T, $r(T)$, is defined as:

$$r(T) = \sup_{\lambda\in\sigma(T)} |\lambda|.$$

We know that $r(T) \leq \|T\|_{\mathcal{L}(X)}$.

Proposition 4.3.3

$$r(T) = \lim_{n\to\infty} \|T^n\|_{\mathcal{L}(X)}^{1/n}$$

If A is self-adjoint (on a Hilbert Space), then $r(A) = \|A\|_{\mathcal{L}(H)}$.

Proof We admit that $\lim \|T^n\|_{\mathcal{L}(X)}{}^{1/n}$ exists.

$$\begin{aligned} R_\lambda(T) &= \frac{1}{\lambda}\sum_{n=0}^{\infty}\frac{T^n}{\lambda^n} \quad \text{Think of this as a series in } z = \frac{1}{\lambda}. \\ &= z\sum_{n=0}^{\infty} T^n z^n \end{aligned}$$

The radius of convergence is given by:

$$\frac{1}{\limsup \|T^n\|^{1/n}} = \frac{1}{\lim \|T^n\|^{1/n}}.$$

This is called Hadamard's Formula . So, for

$$\left|\frac{1}{\lambda}\right| < \frac{1}{\lim \|T^n\|^{1/n}},$$

$R_\lambda(T)$ converges. Hence, $\forall\lambda > \lim_{n\to\infty}\|T^n\|^{1/n}$, $\lambda \in \rho(T)$. Hence, $r(T) \leq \lim_{n\to\infty}\|T^n\|^{1/n}$.

Conversely, if $|\lambda| > r(T)$, that means $\lambda \in \rho(T)$ and $R_\lambda(T)$ is analytic there. $\Longrightarrow$ $\frac{1}{\lambda}$ has to be in the disc of convergence of:

$$\frac{1}{\lambda}\sum_{n=0}^{\infty}\frac{T^n}{\lambda^n}$$

$$\Rightarrow \left|\frac{1}{\lambda}\right| \leq \frac{1}{\lim\|T^n\|^{1/n}}.$$

Hence, $|\lambda| \geq \lim_{n\to\infty}\|T^n\|^{1/n}$. Therefore, $r(T) \geq \lim\|T^n\|^{1/n}$. We conclude that:

$$r(T) = \lim_{n\to\infty}\|T^n\|^{1/n}.$$

In the case of a self-adjoint operator A on a Hilbert space H, $\|A^2\|_{\mathcal{L}(H)} = \|A^*A\|_{\mathcal{L}(H)} = \|A\|^2_{\mathcal{L}(H)}$ (Check that this is indeed the case!). Then, $\|A^{2n}\|_{\mathcal{L}(H)} = \|A\|^{2n}_{\mathcal{L}(H)}$. Thus, $r(A) = \|A\|$. ∎

Example of the Shift Operator Consider $T : l_1 \to l_1$ given by:

$$T(u_1, u_2, \ldots) = (u_2, u_3, \ldots).$$

Its adjoint from $l_\infty \to l_\infty$ is given by $T'(u_1, u_2, \ldots) = (0, u_1, u_2, \ldots)$.

- **Point Spectrum of** T: $Tu = \lambda u$. For $|\lambda| < 1$, define $u_\lambda = (1, \lambda, \lambda^2, \ldots)$. Then, $u_\lambda \in l_1$. So,

  $$Tu_\lambda = \lambda u_\lambda.$$

  Hence, $\{|\lambda| < 1\} \subseteq \sigma(T)$ and $\|T\| = \|T'\| = 1$. Therefore, $\sigma(T) \subset \overline{B(0,1)}$.

  But, what happens for $|\lambda| = 1$? Then, if we solve $Tu = \lambda u = u$, we get that $|u_1| = |u_2| = |u_3|, = \ldots$. But, this means that either $u_i = 0\ \forall i$, or $u \notin l_1$. Thus, $|\lambda| = 1$ is not in the point spectrum.

- T' **has no point spectrum**: $T'u = \lambda u$ gives:

  $$\begin{aligned} \lambda u_1 &= 0 \\ \lambda u_2 &= u_1 \\ &. \\ &. \end{aligned}$$

  This means that $u_1 = u_2 = \ldots = 0$.

- **If λ is in the point spectrum of** T**, then** Ran $(\lambda I - T')$ **is not dense**: Take $f \in (l_1)^*$:

  $$< \underbrace{(\lambda I - T')}_{\in (l_1)^* = l_\infty}(f), x > = < f, (\lambda I - T)x > .$$

Let $|\lambda| < 1$ and apply this to $x = u_\lambda = (1, \lambda, \lambda^2, \cdots)$. We see that $< (\lambda I - T')(f), u_\lambda >= 0 \ \forall f \in (l_1)^*$. From this, we deduce that Ran $(\lambda I - T')$ is not dense, for if it were, every $L \in (l_1)^*$ could be approximated by functions of the form $(\lambda I - T')f_n$ where $f_n \in (l_1)^*$, leading to the conclusion that $< L, u_\lambda >= 0 \ \ \forall L \in (l_1)^* \implies$ by Hahn Banach that $u_\lambda = 0$, a clear contradiction.

- $\lambda \in$ **residual spectrum of** $T \implies \lambda \in$ **point spectrum of** T': Suppose $\lambda \in$ residual spectrum of T.

  $\implies$ Ran $(\lambda I - T)$ is not dense.

  $\implies \exists f \in (l_1)^*$ s.t. $< f, (\lambda I - T)x >= 0 \ \forall x$.

  $\implies < (\lambda I - T')(f), x >= 0 \ \forall x$.

  $\implies \lambda$ is an eigenvalue of T'.

- **If** $|\lambda| = 1$ **then** $\lambda \in$ **residual spectrum of** T': Take $|\lambda| = 1$. Then the element, $c = (1, \overline{\lambda}, \overline{\lambda}^2, \ldots) \in l_\infty$. We will show that $B(c, \frac{1}{2})$ does not intersect Ran $(\lambda I - T')$. So, the range is not dense and thus, $\lambda \in$ residual spectrum of T'.

  Assume $d \in B(C, \frac{1}{2})$ and $\exists e \in l_\infty$ such that $d = (\lambda I - T')e$. Then,

$$
\begin{aligned}
d_1 &= \lambda e_1 \\
d_2 &= \lambda e_2 - e_1 \\
d_3 &= \lambda e_3 - e_2 \\
&. \\
&.
\end{aligned}
$$

  More generally, we can write: $e_n = \overline{\lambda}^{n+1} \sum_{k=1}^{n} \lambda^k d_k$. We just need to check that this is not in l_∞. First, note that $\lambda^k c_k = 1$ for all k. Also, note that $|d_k - c_k| < 1/2$ since $d \in B(c, \frac{1}{2})$. Therefore, since $|\lambda| = 1$, we get that

$$
1/2 > |\lambda^k d_k - \lambda^k c_k| = |\lambda^k d_k - 1| \implies \Re(\lambda^k d_k) \geq 1/2
$$
$$
\implies \Re\left(\sum_{k=1}^{n} \lambda^k d_k\right) \geq n/2 \implies |e_n| \geq n/2 \implies e \notin l_\infty
$$

  But, this is a contradiction. Hence, $B(c, \frac{1}{2})$ does not intersect Ran $(\lambda I - T')$, and $\lambda \in$ residual spectrum of T'.

Summary of Results for Shift Operator on l_1:

	Spectrum	Point Spectrum	Residual Spectrum
T	$\|\lambda\| \leq 1$	$\|\lambda\| < 1$	$\emptyset$
T'	$\|\lambda\| \leq 1$	$\emptyset$	$\|\lambda\| \leq 1$

In general, for any $T \in \mathcal{L}(X, Y)$ for any Banach spaces X, Y, we have the following:

Proposition 4.3.4

1. *If $\lambda \in$ Residual spectrum of T, then λ is in the point spectrum of T'.*

2. *If $\lambda \in$ point spectrum of T, then $\lambda \in$ point spectrum of T' or $\lambda \in$ Residual spectrum of T'.*

Theorem 4.3.5 *Let H be a Hilbert space and $A \in \mathcal{L}(H)$ be self-adjoint. Then,*

1. *A has no residual spectrum.*

2. *$\sigma(A) \subseteq \mathbb{R}$.*

3. *Eigenvectors corresponding to different eigenvalues are orthogonal.*

Proof

1. If λ were in the residual spectrum of A, then λ would be in the point spectrum of $A^* = A$. But, this is a contradiction since the residual spectrum and the point spectrum are disjoint.

2. $\|Ax - (\lambda + i\mu)x\|^2 = \|Ax - \lambda x\|^2 + \mu^2\|x\|^2 + 2\Re < Ax - \lambda x, i\mu x >$. Now, $< Ax - \lambda x, i\mu x >= i\mu < Ax, x > -i\lambda\mu\|x\|^2 = imaginary$ since $\overline{< Ax, x >} =< x, Ax >=< Ax, x >$, thus showing that $< Ax, x >$ is real. Hence,
$$\|Ax - (\lambda + i\mu)x\|^2 \geq \mu^2\|x\|^2. \tag{4.1}$$

   So, assume $\mu \neq 0$. We will show that $\lambda + i\mu \in \rho(T)$. If $\mu \neq 0$, then, we deduce that $A - (\lambda + i\mu)I$ is one-to-one. Therefore, $\lambda + i\mu$ is not in the point spectrum. So, now we will check that Ran $(A - (\lambda + i\mu)I)$ is closed. Suppose that $y_n = Ax_n - (\lambda + i\mu)x_n \longrightarrow y$. Since $\{y_n\}_n$ is a Cauchy sequence, we apply the inequality, 4.1 to get that:
$$\|y_n - y_m\|^2 \geq \mu^2\|x_n - x_m\|^2,$$
   showing that $\{x_n\}_n$ is also Cauchy, hence $\exists x$ such that $x_n \longrightarrow x$. Moreover, by continuity, $(A - (\lambda + i\mu))x_n \longrightarrow (A - (\lambda + i\mu))x$. Hence, $y = (A - (\lambda + i\mu))x$ and is thus in Ran $(A - (\lambda + i\mu)I)$.

   If Ran $(A - (\lambda + i\mu)I)$ were not dense, then $\lambda + i\mu$ would be in the residual spectrum of A. But, A has no residual spectrum. Hence, Ran $(A - (\lambda + i\mu)I)$ is dense and closed. Therefore, it must be that Ran $(A - (\lambda + i\mu)I) = H \implies A - (\lambda + i\mu)I$ is onto. Therefore, it is invertible since we showed earlier that it is one-to-one. Therefore, $(\lambda + i\mu) \in \rho(T)$. Therefore, $\lambda + i\mu$ is in the spectrum only if $\mu = 0$, as desired. ∎

4.4 Positive Operators and Polar Decomposition (In a Hilbert Space)

Definition $A \in \mathcal{L}(H)$ is said to be *positive* if for every x, $< Ax, x > \geq 0$. We write $A \geq 0$. Also, $A \geq B$ means that $A - B \geq 0$.

Proposition 4.4.1 *Every positive operator on a* complex *Hilbert space is self-adjoint.*

Proof $< Ax, x >$ is real if A is positive. Hence,

$$< Ax, x >= \overline{< Ax, x >} =< x, Ax > .$$

Now, $\forall x, y \in H$, this means that:

$< x+y, A(x+y) >=< A(x+y), x+y > \quad < x-y, A(x-y) >=< A(x-y), x-y >$

Subtracting accordingly, we get that $< x, Ay >=< Ax, y >$. ∎

Note: $\forall A \in \mathcal{L}(H),\ A^*A \geq 0$ since $< x, A^*Ax >=< Ax, Ax >= \|Ax\|^2 \geq 0$.

Proposition 4.4.2 (Existence of Square Roots) *Let A be a positive operator in $\mathcal{L}(H)$. Then, $\exists$ a unique positive operator B such that $A = B^2$*

Proof By scaling, reduce to $\|I - A\| < 1$. Compute $\sqrt{A} = \sqrt{I - (I - A)}$ through the series expansion of $\sqrt{1 - z}$:

$$f(z) = \sqrt{1 - z} = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} z^n,$$

with $f^{(n)}(0) \geq 0 \ \ \forall n$. Hence $\sqrt{A}$ is positive. ∎

Definition $|A| = \sqrt{A^*A} \quad$ (in $\mathcal{L}(H)$).

Definition $U \in \mathcal{L}(H)$ is an *isometry* if $\|Ux\| = \|x\| \ \ \forall x \in H$. It is a *partial isometry* if it is an isometry restricted to $(\ker U)^{\perp}$.

Proposition 4.4.3 *Let U be a partial isometry. Then, $U^*U = P\Big|_{(\ker U)^{\perp}}$ is an orthogonal projection on $(\ker U)^{\perp}$ and $UU^* = P\Big|_{\text{Ran } U}$.*
Conversely, if U satisfies these properties, then U is a partial isometry.

Theorem 4.4.4 (Polar Decomposition) *Let $A \in \mathcal{L}(H)$. Then, there exists a partial isometry U such that $A = U|A|$. This U is uniquely determined by the requirement* $\ker U = \ker A$. *Moreover,* $\text{Ran } U = \overline{\text{Ran } A}$.

Example
$A =$ right shift in l_2
$A^* =$ left shift in l_2.

$$A^*A = I \quad \Longrightarrow \quad |A| = I.$$

In the polar decomposition, $A = U(|A|) = U$. So, we see that $U = A$ is not an isometry since A is not invertible.

Chapter 5

Compact and Fredholm Operators

5.1 Definitions and Basic Properties

Definition Let X and Y be Banach spaces. $T \in \mathcal{L}(X,Y)$ is said to be *compact* if $\overline{T(B_X)}$ is compact.

$\Longleftrightarrow$ T maps bounded sets into precompact sets (i.e. sets with compact closure). $\Longleftrightarrow$ T maps bounded sequences into sequences which have convergent subsequences.

Proposition 5.1.1 *If $x_n \rightharpoonup x$, then $T(x_n) \to T(x)$ strongly in Y.*

Proof If $x_n \rightharpoonup x$ then $\{x_n\}_n$ is bounded. Hence, $\{T(x_n)\}_n$ has a convergent subsequence that converges to some $y \in Y$. Since T is continuous in the strong-strong topology, it is also continuous in the weak-weak topology. Hence, $T(x) = y$ and $T(x_{n_k}) \to T(x)$. A sequence whose every convergent subsequence converges to $T(x)$ and which is bounded, converges to $T(x)$. Hence, $T(x_n) \to T(x)$. ∎

Definition $T \in \mathcal{L}(X,Y)$ is said to be an operator of *finite rank* if $\dim \operatorname{Ran}(T) < \infty$.

Remark Finite rank operators are obviously, compact (since a closed and bounded subset of a finite-dimensional space is compact).

Proposition 5.1.2

1. *If T_n are compact operators in $\mathcal{L}(X,Y)$ and $T_n \to T$ in the $\mathcal{L}(X,Y)-$norm, then T is compact.*
2. *T is compact $\Longrightarrow$ T' is compact.*

3. *If T is compact and S is bounded, then $T \circ S$ and $S \circ T$ are compact.*

Proof

1. For $\epsilon > 0$, $n \geq N$, $\|T_n - T\|_{\mathcal{L}(X,Y)} < \epsilon$. Since T_N is compact, $T_N(B_X)$ is precompact. Hence, it can be covered by a finite number of balls of radius ϵ. So,
$$T_N(B_X) \subset \bigcup_{\text{finite}} B(y; \epsilon).$$
But, $\forall x \in B_X, \|T_N(x) - T(x)\| < \epsilon$. Therefore,
$$T(B_X) \subseteq \bigcup_{\text{finite}} B(y, 2\epsilon).$$
Since this is true $\forall \epsilon > 0$, $T(B_X)$ is precompact.

2. In homework (Due 11/19).

3. Since S bounded, $S(B_X)$ is bounded. Since T compact, therefore, $T(S(B_X))$ is precompact. Hence, $T \circ S$ is compact. On the other hand, if T compact, then $T(B_X)$ is precompact. Since S is bounded, $S(T(B_X))$ is precompact as well by continuity. Hence, $S \circ T$ is compact. ∎

Note: This theorem shows that limits of finite rank operators are compact!

Conversely: Can any compact operator be approximated by finite rank operators? Not always. Yes if we're in a Hilbert space:

Theorem 5.1.3 *Let H be a Hilbert space and $T \in \mathcal{L}(H)$ compact. Then, T is the uniform limit of finite rank operators.*

Proof Let $K = \overline{T(B_H)}$, compact. Given $\epsilon > 0$, there exists a covering of K:
$$K \subset \bigcup_{\text{finite}} B(y, \epsilon).$$
Let Y be the space spanned by the $y_i's$. $\dim Y < \infty$. Let P_Y be the orthogonal projection onto Y. Take $T_\epsilon = P_Y \circ T$. Let $x \in B_H$. Therefore, $\exists i_o$ such that $\|Tx - y_{i_0}\| < \epsilon$. By projection,
$$\begin{aligned} \|P_Y Tx - P_Y(y_{i_0})\| < \epsilon &\implies \|T_\epsilon x - y_{i_0}\| < \epsilon \\ &\implies \|T_\epsilon x - Tx\| < 2\epsilon \\ &\implies \|T_\epsilon - T\|_{\mathcal{L}(H)} < 2\epsilon \quad ∎ \end{aligned}$$

Important Example (Kernel of Integral Operator)
Let $X = \left(C^0([0,1]), \|\cdot\|_\infty\right)$. $K(x,y) \in C^0([0,1] \times [0,1])$. For all $f \in X$, define:
$$T_K f(x) = \int_0^1 K(x,y) f(y)\, dy.$$

Proposition 5.1.4 *For each K defined as above, T_K is a compact operator.*

Proof Say $f \in B_X$, $\|f\|_\infty \leq 1$. Thus, $|T_K f(x)| < \int_0^1 |K(x,y)||f(y)\, dy \leq \|K\|_\infty$, independent of f. Hence, $T_K \in \mathcal{L}(X)$, with $\|T_K\| \leq \|K\|_\infty$.

To show that $T_K(B_X)$ is precompact, we will use Ascoli's Theorem , which says that if a uniformly bounded family is equicontinuous, every subsequence has a limit point. So, all remains to show is that $T_K f$ are an equicontinuous family. First, $\forall f \in B_X,\ \|T_K(f)\| \leq \|K\|_\infty$. Hence, $T_K f$ are uniformly bounded. Remains to show equicontinuity.

$K \in C^0([0,1]\times[0,1])$. Since K is a continuous function on a compact set, it is uniformly continuous on that set. Therefore, $\forall \epsilon > 0, \quad \exists \delta > 0$ such that $\forall x, y \in [0,1]$, with $|x - x'| < \delta, \quad |K(x,y) - K(x',y)| < \epsilon$. Then,

$$|T_K f(x) - T_k f(x')| < \int_0^1 |K(x,y) - K(x',y)||f(y)|\, dy < \epsilon$$

for any $f \in B_X$. Hence, we have shown equicontinuity of the family. Then, the conclusion of Ascoli's Theorem gives us that $T_K(B_X)$ is precompact. ∎

5.2 Riesz-Fredholm Theory

Lemma 5.2.1 (Riesz) *Let X be a Banach space and $M \subseteq X$, a closed linear subspace of X, $M \neq X$. Then, $\forall \epsilon > 0$, there exists $\|x\| = 1$ such that $dist(x, M) \geq 1 - \epsilon$.*

Proof Take $x \in X \setminus M$ and let $d = dist(x, M) \neq 0$ (since M is closed). Therefore, $\exists y \in M$ such that $\|x - y\| < \frac{d}{1-\epsilon}$. Take $v = \frac{x-y}{\|x-y\|}$. Now, we want to calculate $dist(v, M)$.

$\forall m \in M$,

$$\|v - m\| = \frac{\|x - \overbrace{(y + \|x-y\|m)}^{\in M}\|}{\|x-y\|} \geq \frac{dist(x,M)}{\|x-y\|} \geq 1 - \epsilon.$$

So, $dist(v, M) \geq 1 - \epsilon$. Hence, v is the one we want. ∎

Definition Let X be a Banach space and Y be a subspace of X. Then, $Y^\perp$ is the subspace of X^* defined by:

$$Y^\perp = \{f \in X^* \ :\ \forall y \in Y,\ f(y) = 0\}$$

Remark $Y^\perp$ is always closed. If $X = X^*$, then $(Y^\perp)^\perp = \overline{Y}$. If things are closed, $(\ker T')^\perp = \text{Ran } T,\ (\text{Ran } T')^\perp = \ker T$.

Definition $\text{codim } Y = \dim Y^\perp$

Theorem 5.2.2 (Fredholm Alternative) *Let $T \in \mathcal{L}(X)$ be a compact operator on a Banach space. Then,*

1. $\ker(I - T)$ *is finite dimensional.*
2. $\text{Ran } (I - T)$ *is closed and* $= (\ker(I - T'))^{\perp}$
3. $\ker(I - T) = \{0\} \iff \text{Ran } (I - T) = X$.
4. $\dim \ker(I - T) = \dim \ker(I - T')$.

The Alternative Let $A = I - T$. Either, "$\ker A = \{0\}$ and Ran $A = X$" **OR** "$\ker A \neq \{0\}$ and Ran $A \neq X$". In other words, either, $Ax = b$ has a unique solution or $Ax = 0$ has non-trivial solutions.

Motivation T is compact (think of an integral operator). Want to solve, $T\varphi - \varphi = f$. This either has solutions $\forall f$ or $T\varphi = \varphi$ has non-trivial solutions.

Example

- Let $\varphi' - \varphi'' = f, \quad \varphi(0) = \varphi(1) = 0$.
- Or in PDE: $\triangle\varphi - \varphi = f$.

Proof of Fredholm Alternative

1. Let $\mathcal{N} = \ker(I - T)$. Then, $\forall x \in \mathcal{N},\ T(x) = x$. $B_{\mathcal{N}} = T(B_{\mathcal{N}}) \subset T(B_X)$. Hence, $B_{\mathcal{N}}$ is precompact. Recall the theorem of Riesz that $\overline{B_{\mathcal{N}}}$ compact $\iff \dim \mathcal{N} < \infty$.

2. Ran $(I - T)$ is closed: Let $f_n = x_n - T(x_n), \quad f_n \to f$. Is $f \in \text{Ran } (I - T)$? Since by the above, $\ker(I - T)$ is finite-dimensional, hence closed. Hence, $d_n = dist(x_n, \ker(I - T))$ is achieved. So, $\exists v_n \in \ker(I - T)$ such that $d_n = \|x_n - v_n\|,\ v_n = Tv_n$. So, we can write:

$$f_n = x_n - v_n - T(x_n - v_n) \tag{5.1}$$

   If $\|x_n - v_n\|$ is bounded, then by compactness of T, we can assume (up to extraction), that $T(x_n - v_n) \to l$. So, pass to the limit in Eqn. 5.1, to obtain $x_n - v_n \to l + f$. Again, passing to the limit in Eqn. 5.1, we get that $f = l + f - T(l + f) \implies f = (I - T)(l + f)$. So, $f \in \text{Ran } (I - T)$. All that remains to do is to check that $\{\|x_n - v_n\|\}_n$ is bounded.

   Suppose not. Then, divide the quantity in Eqn. 5.1 by $\|x_n - v_n\|$. Then,

$$\underbrace{\frac{x_n - v_n}{\|x_n - v_n\|}}_{\equiv u_n} - T\left(\frac{x_n - v_n}{\|x_n - v_n\|}\right) \to 0 \tag{5.2}$$

   $\{u_n\}_n$ is certainly bounded. By compactness of T we can assume that $T(u_n) \to z$. From Eqn. 5.2, $u_n \to z$. Therefore, by uniqueness of limits, we can conclude that $T(z) = z$. Hence, $z \in \ker(I-T)$. But, $dist(x_n, \ker(I-T)) = \|x_n - v_n\|$. Hence, $dist(u_n, \ker(I-T)) = 1$. But, this contradicts the fact that $u_n \to z \in Ker(I-T)$. This proves that $\|x_n - v_n\|$ is bounded. Hence, we're done.

3. $\underline{\ker(I-T) = \{0\} \iff \text{Ran } (I-T) = X:}$

   ($\Longrightarrow$) Assume not. In other words, $\exists x \in X \setminus \text{Ran } (I-T)$. Let $X_1 = \text{Ran } (I-T)$. It is closed. Therefore, it is a Banach space. Also, $T(X_1) \subseteq X_1$ since if $y = x - T(x)$, then

   $$T(y) = T(x) - T^2(x) = (I-T)(T(x)) \in \text{Ran } (I-T) = X_1.$$

   Now, consider $T\big|_{X_1}$ Then, let $X_2 = \text{Ran } (I - T\big|_{X_1})$. Inductively, let $X_n = (I-T)^n(X)$. Then, $X_n \subsetneq X_{n+1}$ (Why?) If $X_n = X_{n-1}$, then Ran $(I-T)^n =$ Ran $(I-T)^{n-1}$. So, applying this to x, we see that $(I-T)^{n-1}x = (I-T)^n y$ for some y. But, $I-T$ is injective. So, $(I-T)y = x \in Ran(I-T)$. This is a contradiction of the original assumption that $x \in X \setminus \text{Ran } (I-T)$.

   Now, apply Riesz' Lemma (Lemma 5.2.1) and we find a $x_n \in X_n$ such that $\|x_n\| = 1$ and $dist(x_n, x_n + 1) \geq 1/2$. Now, consider x_n, x_m $m < n$. Then,

   $$Tx_n - Tx_m = \underbrace{(T-I)x_n}_{\in X_{n+1}} - \underbrace{(T-I)x_m}_{\in X_{m+1}} + \underbrace{x_n}_{\in X_n} - \underbrace{x_m}_{\in X_m}.$$

   So, $\|Tx_n - Tx_m\| > dist(x_m, X_{m+1}) \geq 1/2$. Hence, $\{Tx_m\}_m$ is not a Cauchy sequence, which contradicts the fact that $\{x_n\}_n$ is bounded and T is compact. Therefore, Ran $(I-T) = X$.

   ($\Longleftarrow$) If Ran $(I-T) = X$, then, $\ker(I-T') = \{0\}$. So, apply the ($\Rightarrow$) direction to T', which is also compact. This gives that Ran $(I-T') = X^*$. Hence, $\ker(I-T) = \{0\}$.

4. $\underline{\dim\ker(I-T) = \dim\ker(I-T')}$ (Check as an exercise!) ∎

5.3 Fredholm Operators

Definition A *Fredholm Operator* is an operator $A \in \mathcal{L}(X, Y)$ such that:

- $\ker(A)$ is finite-dimensional
- Ran (A) is closed and has finite codimension (codim Ran $(A) = \dim(\text{Ran } (A))^\perp$).

The index of A is given by:

$$\text{Ind } (A) = \dim(\ker(A)) - \text{codim } (\text{Ran } (A))$$

Example From Riesz-Fredholm Theorem (by parts (1), (2), and (4)) of Fredholm Alternative), if T is compact, then $I-T$ is Fredholm of index 0.

Theorem 5.3.1

1. *The set of* Fred (X, Y) *is open in* $\mathcal{L}(X, Y)$ *and* $A \mapsto \text{Ind } A$ *is continuous, and therefore constant on each connected component of* Fred (X, Y).

2. *Every Fredholm operator is invertible modulo finite rank operators.* $\exists B \in \mathcal{L}(X, Y)$ *such that* $BA - I_X$ *and* $AB - I_Y$ *have finite rank. Conversely, if* $A \in \mathcal{L}(X, Y)$ *is an operator such that* $\exists B \in \mathcal{L}(X, Y)$ *with* $AB - I$ *and* $BA - I$ *compact, then* A *is Fredholm.*

3. *If* A *is Fredholm and* T *compact, then* $A + T$ *is Fredholm and* $\text{Ind }(A + T) = \text{Ind } A$.

4. *If* A *and* B *are Fredholm, then* AB *is also and* $\text{Ind }(AB) = \text{Ind } A + \text{Ind } B$.

5. *If* A *is Fredholm, then* A' *is Fredholm and* $\text{Ind } A' = -\text{Ind } A$.

Example

- Right-shift in l_p: Consider the operator, $A : (u_1, u_2, \ldots) \mapsto (0, u_1, u_2, \ldots)$. Then, $\ker A = \{0\}$. Also, $\text{Ran } A = \{(u_i)_i : \; u_1 = 0\}$. It is closed and codim Ran $A = 1$. Hence, Ind $A = -1$.

- Lef-shift in l_p: Consider the operator $A : (u_1, u_2, \ldots) \mapsto (u_2, u_3, \ldots)$. Then, $\ker A = \{(u, 0, 0, \ldots) : \; u \in \mathbb{R}\}$. Hence, $\dim\ker A = 1$ and $\text{Ran } A = l_p$. Hence, codim Ran $A = 0$. Therefore, Ind $A = 1$.

- Erasure in l_p: Consider the operator $A : (u_1, u_2, \ldots) \mapsto (0, u_2, u_3, \ldots)$. Then, $\ker A = \{(u, 0, 0, \ldots) : \; u \in \mathbb{R}\}$. $\implies$ $\dim\ker A = 1$. Also, $\text{Ran } A = \{(u_n)_n : \; u_1 = 0\}$. It follows, then, that codim Ran $A = 1$ and Ind $A = 0$.

5.4 Spectrum of Compact Operators

Theorem 5.4.1 (Riesz-Schauder) *Let* $T \in \mathcal{L}(X)$ *be a compact operator and* $\dim X = \infty$. *Then, the following hold:*

- $0 \in \sigma(T)$

- $\sigma(T) \setminus \{0\}$ *consists of eigenvalues of finite multiplicity (i.e. the dimension of the* λ*-eigenspace (*$\ker(T - \lambda I)$*) has finite dimension* $\forall \lambda \in \sigma(T) \setminus \{0\}$.

- $\sigma(T) \setminus \{0\}$ *is either empty, finite or a sequence converging to* 0 *(i.e. it is a discrete set with no limits other than* 0*).*

Proof

- If $0 \notin \sigma(T)$ then T is invertible (i.e.: $\ker T = \{0\}$). Therefore, $T \cdot T^{-1} = I$. $\implies$ since T and T^{-1} are compact, that I is compact. Hence, $\dim X < \infty$ (by Riesz' Theorem), a contradiction! **This shows that in infinite dimension, a compact operator is never invertible.**

- From Riesz-Fredholm Theorem (i.e. Fredholm Alternative), if $\lambda \in \sigma(T) \setminus \{0\}$ then since $\lambda \neq 0$, if $\ker(I - \frac{T}{\lambda}) = \{0\}$, then Ran $\left(I - \frac{T}{\lambda}\right) = X$ (since T compact $\implies \frac{T}{\lambda}$ compact). This would show that $I - \frac{T}{\lambda}$ is invertible, a clear contradiction. So, $\ker\left(I - \frac{T}{\lambda}\right) \neq \{0\}$. Hence, λ is an eigenvalue. Moreover, $\dim(\ker\left(I - \frac{T}{\lambda}\right)) < \infty$ by Riesz-Fredholm.

- Suppose to the contrary, that $\exists$ a sequence of non-zero eigenvalues, $\lambda_n \to \lambda \neq 0$. Each λ_n is an eigenvalue, so take e_n to be an eigenvector. The e_n are linearly independent (to see this, by induction assume that $e_{n+1} = \sum_{i=1}^n \alpha_i e_i$. Then, $\lambda_{n+1}\left(\sum_{i=1}^n \alpha_i e_i\right) = T(e_{n+1}) = \sum_{i=1}^n \lambda_i \alpha_i e_i$. Hence, since $e_1, \ldots, e_n$ are linearly independent, $\lambda_i \alpha_i = \lambda_{n+1}\alpha_i$ for each i. But, we assumed that $\lambda_n \neq \lambda_{n+1}$ we get that $\alpha_i = 0$, a contradiction). So, let $X_n = Span(e_1, \ldots, e_n)$. $X_n \nsubseteq X_{n+1}$. Moreover, $(T - \lambda_n I)X_n \subset X_{n-1}$. By Riesz' Lemma, take a sequence $u_n \in X_n$, such that for each n, $\|u_n\| = 1$ and $dist(u_n, X_{n-1}) \geq 1/2$. Then,

$$\left\|\frac{Tu_n}{\lambda_n} - \frac{Tu_m}{\lambda_m}\right\| = \left\|Tu_n - \frac{\lambda_n u_n}{\lambda_n} - \frac{Tu_m - \lambda_m u_m}{\lambda_m} + u_n - \underbrace{u_m}_{\in X_{n-1}}\right\| = \star.$$

So, take $n > m \implies m \leq n-1 \implies X_m \subseteq X_{n-1}$

$$\implies \frac{Tu_n - \lambda_n u_n}{\lambda_n} \in (T - \lambda_n I)X_n \subset X_{n-1}$$
$$\frac{Tu_m - \lambda_m u_m}{\lambda_m} \in X_m \subset X_{n-1}$$
$$\implies \frac{Tu_n - \lambda_n u_n}{\lambda_n} - \frac{Tu_m - \lambda_m u_m}{\lambda_m} \in X_{n-1}$$
$$\implies \star = \|u_n - \underbrace{\tilde{x}}_{\in X_{n-1}}\| > 1/2$$

If $\lambda_n, \lambda_m \to \lambda \neq 0$, this contradicts the fact that Tu_n is a Cauchy sequence. But, $\|u_n\| = 1$ and T is compact. This is a contradiction. Therefore, $\lambda = 0$. ∎

Remark Conversely, if $\alpha_n \to 0$, one can build a compact operator whose spectrum is exactly that sequence. For example, consider l_2 and take $\{u_n\}_n \mapsto \{\alpha_n u_n\}_n$. This can be approximated by finite rank operators. Hence, it is compact.

5.5 Spectral Decomposition of Compact, Self-Adjoint Operators in Hilbert Space

Proposition 5.5.1 *Let $T \in \mathcal{L}(H)$ be a self-adjoint operator on a Hilbert space (recall that self-adjoint operators in Hilbert space have real spectrum... i.e.: $\sigma(T) \subset \mathbb{R}$). Then if we define:*

$$M = \sup_{\|x\|=1} < x, Tx >, \quad m = \inf_{\|x\|=1} < x, Tx >$$

Then, $\sigma(T) \subset [m, M]$ with $m, M \in \sigma(T)$.

Proof Private Exercise!

Corollary 5.5.2 *If T is self-adjoint and $\sigma(T) = \{0\}$ then $T = 0$.*

Proof If $\sigma(T) = 0$, then $m = M = 0$, in the notation of the preceding proposition. Then, $\forall x$, $< x, Tx >= 0$. So, polarize to get $< x, Ty >= 0, \ \forall x, y \in H$. Hence, $T \equiv 0$. ∎

Theorem 5.5.3 (Hilbert-Schmidt Theorem) *Let T be a compact self-adjoint operator on a Hilbert space. Then, $\exists$ a complete orthonormal basis of H formed of eigenvectors such that*

$$T\varphi_n = \lambda_n \varphi_n \ \ \forall n.$$

(If H is separable, then you can find a countable basis. If H is not, then there is possibly an uncountable basis of $\ker T$). Also,

$$\lim_{n\to\infty} \lambda_n = 0 \quad \lambda_0 = 0, \quad (\sigma(T) \setminus \{0\} = \{\lambda_n\}_n)$$

Proof Take $E_n = \ker(T - \lambda_n I)$. Then, by Riesz-Schauder Theorem, $\dim E_n < \infty$. If $x \in E_n$, $y \in E_m$, for $n \neq m \quad \Rightarrow \ < x, y >= 0$. This can be shown by noting that:

$$\lambda_m < x, y >=< x, \lambda_m y >=< x, Ty >=< Tx, y >=< \lambda_n x, y >= \lambda_n < x, y >$$

But, $\lambda_n \neq \lambda_m$. Hence, $< x, y >= 0$. So, let M be the sum of the E_n and $\ker T$. M is stable under T: $T(M) \subset M$ (since E_n is space of eigenvectors). Hence, $M^\perp$ is also stable under T (take $x \in M, y \in M^\perp \ \Rightarrow < x, y >= 0 \ \Rightarrow \ M$ is stable under $T \ \Rightarrow Tx \in M \ \Rightarrow < x, Ty >=< Tx, y >= 0$ since T is self-adjoint and $Tx \in M,\ y \in M^\perp$. Hence, $Ty \in M^\perp$).

So, consider $T\big|_{M^\perp}$. It's also a compact (self-adjoint) operator.

($\Longrightarrow$) $\sigma(T\big|_{M^\perp}) \setminus \{0\}$ is formed by eigenvalues.

($\Longrightarrow$) in $M^\perp$ there are eigenvectors for T, but they are all in M.

($\Longrightarrow$) The only possibility is $\sigma(T\big|_{M^\perp}) = \{0\}$.

So, by previous corollary, $T\big|_{M^\perp} \equiv 0$. Hence, $M^\perp \subset \ker T \subset M$.

$\Longrightarrow \ M^\perp = \{0\}$ (since $M \cap M^\perp = \{0\}$).

So, choose, an orthonormal basis in each E_n (each is finite dimensional) and an orthonormal basis of $\ker T$. This provides a complete orthonormal family, which is countable if H is separable. ∎

T can be approximated by finite rank operators in the following manner: If $x = \sum_{n=0}^{\infty} x_n, \quad x_n \in E_n, \quad E_0 = \ker T$.

$$Tx = \sum_{n=0}^{\infty} \lambda_n x_n$$
$$T_N x = \sum_{n=0}^{N} \lambda_n x_n$$
$$\|T_N - T\|_{\mathcal{L}(H)} \longrightarrow 0 \quad \text{as } N \to \infty$$

Now, we prove a result that essentially says that: "compact operators on a Hilbert space, can be 'diagonalized' over an orthonormal basis"

Theorem 5.5.4 (Canonical form for Compact Operators) *Let T be a compact operator in $\mathcal{L}(H)$. Then, there exist orthonormal sets (not necessarily complete) $\{\varphi_n\}_n$ and $\{\psi_n\}_n$ and a sequence, $\{\lambda_n\}_n$, with $\lambda_n \to 0$ such that:*

$$T = \sum_{n=0}^{\infty} \lambda_n < \psi_n, \cdot > \varphi_n \quad \textbf{\textit{(SVD)}}$$

The λ_n are eigenvalues of $|T| = \sqrt{T^*T}$ and are called singular values *of T.*

Proof T^*T is compact, self-adjoint. Call its eigenvalues, $\mu_n \to 0$ and let $\{\psi_n\}_n$ be the corresponding orthonormal basis of eigenvectors. Then,

$$T\psi = T\left(\sum_{n=0}^{\infty} < \psi_n, \psi > \psi_n\right) = \sum_{n=0}^{\infty} < \psi_n, \psi > T\psi_n.$$

Let $\varphi_n = \frac{T\psi_n}{\lambda_n}$ where $\lambda_n = \sqrt{\mu_n}$.

$$\Longrightarrow \quad T\psi = \sum_{n=0}^{\infty} < \psi_n, \psi > \lambda_n \varphi_n.$$

Check that that the φ_n defined this way are indeed orthonormal. But, this is clear since the μ_n are the eigenvalues of T^*T. ▮

Appendix A

Definition

- Let X be a metric space. A family of functions, $\{f_\alpha\}_\alpha$ defined on a subset $U \subset X$ is said to be *uniformly bounded* if $\exists C > 0$ such that:

$$\sup_{\substack{\alpha \\ x\in U}} |f_\alpha(x)| \leq C.$$

- Let X be a metric space. A family of functions, $\{f_\alpha\}_\alpha$ defined on a subset $U \subset X$ is said to be *equicontinuous* if $\forall \epsilon > 0$ there exists $\delta > 0$ such that

$$dist(x, y) < \delta \implies \sup_\alpha |f_\alpha(x) - f_\alpha(y)| \leq C,$$

  for all $x, y \in U$.

Theorem A.0.5 (Ascoli's Lemma) *Let $\mathcal{K}$ be a uniformly bounded, equicontinuous, family of functions on a compact metric space X. Then, any sequence contains a subsequence that is uniformly convergent in X to a continuous function.*

Corollary A.0.6 *Let X be a compact metric space. A family $\mathcal{K}$ of functions in X^* is precompact if and only if $\mathcal{K}$ is both uniformly bounded and equicontinuous.*

Index