Metric Spaces

We are going to generalize many of the concepts from basic real analysis to situations involving functions whose inputs (and possibly also outputs) are functions rather than real numbers. In fact we shall study functions $f : X \to Y$, where X and Y are abstract sets endowed with some structure that will allow us to do analysis. In order to define concepts such as continuity, it will be useful to have a precise notion of "distance" between two elements of a set. The framework of a "metric space" (to be defined below) is very natural for studying mathematical concepts based on distance.

The material that follows should be regarded as a review of (or very rapid introduction to) the essentials of metric spaces. Very few proofs or examples will be given here. Students who have not previously studied metric spaces are advised to do some outside reading. Some recommended references are given below.

- W.A. Sutherland, *Introduction to Metric and Topological Spaces*
- Section I.6 of Dunford & Schwartz
- Section 1.7 and Chapter 3 of Friedman
- Chapter 1 of Goffman & Pedrick
- Chapter 1 of Kreyszig
- Chapter 7 of Royden

Definition of a Metric Space

Def: A *metric space* is a pair (X, ρ), where X is a set and $\rho : X \times X \to \mathbb{R}$ satisfies

(i) $\forall x, y \in X \quad \rho(x, y) \geq 0$,

(ii) $\forall x, y \in X, \quad \rho(x, y) = 0 \iff x = y$,

(iii) $\forall x, y \in X, \quad \rho(x, y) = \rho(y, x)$,

(iv) (triangle inequality) $\forall x, y, z \in X, \quad \rho(x, z) \leq \rho(x, y) + \rho(y, z)$.

The function ρ is called a *metric* or *distance* on X and the elements of X are usually called *points*.

Remark: It is useful to observe that the triangle inequality can be reformulated as (iv′) $\forall x, y, z \in X, \quad |\rho(x,z) - \rho(y,z)| \leq \rho(x,y)$.

Indeed, it is immediate that (iv′) implies (iv); it is also straightforward to show that (iv′) follows from (iii) and (iv). We note also that (i) could be omitted from the definition of a metric because it follows from (ii), (iii), and (iv); in particular, for all x, y in X we have

$$2\rho(x,y) = \rho(x,y) + \rho(y,x) \geq \rho(x,x) = 0.$$

It is common practice to talk about "a metric space X" without any explicit reference to the metric. Some caution is required because the same set X can be equipped with more than one metric, and a change of metric can lead to dramatic changes in the metric space structure. However, in situations where there is no danger of ambiguity concerning which metric is intended, we sometimes refer to a metric space X, where X is simply the underlying set of points.

Examples:

(a) $X = \mathbb{R}$, $\rho(x,y) = |x - y|$ for all $x, y \in \mathbb{R}$. This metric is called the *standard metric* on $\mathbb{R}$. There are, of course, numerous other metrics on $\mathbb{R}$.

(b) Let X be any set. The metric ρ defined by

$$\rho(x,y) = \begin{cases} 0 & \text{if } x = y \\ \\ 1 & \text{if } x \neq y \end{cases}$$

is called the *discrete metric* on X. The metric space of part (a) is dramatically different from the metric space obtained by equipping $\mathbb{R}$ with the discrete metric. Many useful examples and counterexamples can be obtained by considering these two spaces.

(c) $X = C[0,1]$, the set of all continuous real-valued functions on $[0,1]$. We shall give two different metrics on X:

$$\rho_1(f,g) = \max\{|f(x) - g(x)| : x \in [0,1]\},$$

$$\rho_2(f,g) = \int_0^1 |f(x) - g(x)| dx.$$

Sets in Metric Spaces

Throughout this section (X, ρ) is a given metric space.

Def: Given $\delta > 0$ and $x \in X$ we define the *ball of radius δ centered at x* by

$$B_\delta(x) = \{y \in X : \rho(y, x) < \delta\}.$$

Def: Let $S \subset X$ and $x_0 \in X$ be given.

(a) We say that x_0 is an *interior point* of S provided there exists $\delta > 0$ such that $B_\delta(x_0) \subset S$. The set of all interior points is called the *interior* of S and is denoted by $\mathrm{int}(S)$.

(b) We say that x_0 is a *point of closure* of S provided that

$$\forall \delta > 0, \quad B_\delta(x_0) \cap S \neq \emptyset.$$

The set of all points of closure of S is called the *closure* of S and is denoted by $\mathrm{cl}(S)$ or $\overline{S}$.

Intuitively, to say that x_0 belongs to the interior of S means that x_0 and all points that are sufficiently close to x_0 belong to S. To say that x_0 belongs to the closure of S means that there are points in S that are arbitrarily close to x_0. The following remark is an easy consequence of the definitions of interior and closure.

Remark: Let $A, B \subset X$.

(a) $\mathrm{int}(A) \subset A \subset \mathrm{cl}(A)$.

(b) If $A \subset B$ then $\mathrm{int}(A) \subset \mathrm{int}(B)$ and $\mathrm{cl}(A) \subset \mathrm{cl}(B)$.

Def: Let $S \subset X$. We say that S is

(a) *open* if $S = \mathrm{int}(S)$.

(b) *closed* if $S = \mathrm{cl}(S)$.

To show that S is open, it suffices to show that $S \subset \mathrm{int}(S)$. To show that S is closed, it suffices to show that $\mathrm{cl}(S) \subset S$. These observations follow immediately from the definitions and part (a) of the preceding remark.

Prop. M.1: For every $x \in X$ and every $\delta > 0$, $B_\delta(x)$ is an open set.

Proof: Let $x \in X, \delta > 0$, and $x_0 \in B_\delta(x)$ be given. We need to show that $x_0 \in \mathrm{int}(B_\delta(x))$, i.e. we must produce $\eta > 0$ such that $B_\eta(x_0) \subset B_\delta(x)$. Since $\rho(x_0, x) < \delta$, we may choose η satisfying

$$0 < \eta < \delta - \rho(x_0, x),$$

and hence also

$$\eta + \rho(x_0, x) < \delta.$$

Then we have $B_\eta(x_0) \subset B_\delta(x)$, by virtue of the triangle inequality. Indeed, if $\rho(y, x_0) < \eta$, then

$$\rho(y, x) \leq \rho(y, x_0) + \rho(x_0, x) < \delta. \quad \square$$

Prop. M.2: Let $S \subset X$. Then $\mathrm{int}(S)$ is an open set and $\mathrm{cl}(S)$ is a closed set.

This result is not quite as trivial as it might appear. What Prop. M.2 is really saying is that $\mathrm{int}(\mathrm{int}(S)) = \mathrm{int}(S)$ and $\mathrm{cl}(\mathrm{cl}(S)) = \mathrm{cl}(S)$.

Proof: Put $U = \mathrm{int}(S)$ and $C = \mathrm{cl}(S)$. To show that U is open, let $x \in U$ be given. We want to show that $x \in \mathrm{int}(U)$. Since $x \in \mathrm{int}(S)$ we may choose $\delta > 0$ such that $B_\delta(x) \subset S$. We want to show that $B_\delta(x) \subset U$. Let $y \in B_\delta(x)$ be given. By Prop. M.1, $B_\delta(x)$ is open, so we may choose $\eta > 0$ such that $B_\eta(y) \subset B_\delta(x) \subset S$. This implies that $y \in \mathrm{int}(S) = U$. We conclude that $B_\delta(x) \subset U$, so that $x \in \mathrm{int}(U)$ and U is open.

To prove that C is closed, let $l \in \mathrm{cl}(C)$ be given. We need to show that $l \in C = \mathrm{cl}(S)$. Let $\delta > 0$ be given. Since $l \in \mathrm{cl}(C)$ we may choose $z \in C$ such that $\rho(z, l) < \frac{\delta}{2}$. Since $z \in \mathrm{cl}(S)$ we may choose $w \in S$ such that $\rho(z, w) < \frac{\delta}{2}$. By the triangle inequality, we have $\rho(l, w) < \delta$ so that $B_\delta(x) \cap S \neq \emptyset$. It follows that $l \in \mathrm{cl}(S) = C$. $\square$

Given $S \subset X$, let $S^c = X \backslash S = \{x \in X : x \notin S\}$ denote the complement of S (relative to X). The following elementary result is very useful; it describes how interiors and closures interact with complements.

Prop. M.3: Let $S \subset X$.

(a) $\mathrm{int}(S^c) = (\mathrm{cl}(S))^c$

(b) $\mathrm{cl}(S^c) = (\mathrm{int}(S))^c$

Proof: We shall make use of the simple fact that $A \subset S^c$ if and only if $A \cap S = \emptyset$. To prove (a), let $x \in X$ be given. Then we have

$$\begin{aligned} x \in \mathrm{int}(S^c) &\Leftrightarrow \exists \delta > 0, \quad B_\delta(x) \subset S^c \\ &\Leftrightarrow \exists \delta > 0, \quad B_\delta(x) \cap S = \emptyset \\ &\Leftrightarrow \text{not } (\forall \delta > 0, \quad B_\delta(x) \cap S \neq \emptyset) \\ &\Leftrightarrow x \in (\mathrm{cl}(S))^c. \end{aligned}$$

Part (b) follows from (a), upon replacing S by S^c and using the fact that $(S^c)^c = S$. $\square$

Cor: Let $S \subset X$.

(a) S is open if and only if S^c is closed.

(b) S is closed if and only if S^c is open.

The empty set and the entire space X are open sets and also closed sets. In certain metric spaces, these are the only two open sets that are also closed. (Such metric spaces are said to be *connected*.) However, it is easy to give examples of metric spaces having additional sets that are both open and closed. In particular, with the discrete metric, every subset of X is both open and closed.

Prop. M.4: Let $A, B \subset X$ and $\mathcal{C}$ be a collection of subsets of X. Then

(a) $\text{cl}(A \cup B) = (\text{cl}(A)) \cup (\text{cl}(B))$,

(b) $\text{cl}(\bigcup_{S\in\mathcal{C}} S) \supset \bigcup_{S\in\mathcal{C}} \text{cl}(S)$,

(c) $\text{cl}(\bigcap_{S\in\mathcal{C}} S) \subset \bigcap_{S\in\mathcal{C}} \text{cl}(S)$,

(d) $\text{int}(A \cap B) = (\text{int}(A)) \cap (\text{int}(B))$,

(e) $\text{int}(\bigcap_{S\in\mathcal{C}} S) \subset \bigcap_{S\in\mathcal{C}} \text{int}(S)$,

(f) $\text{int}(\bigcup_{S\in\mathcal{C}} S) \supset \bigcup_{S\in\mathcal{C}} \text{int}(S)$

Remark: Regarding Prop. M.4, in general:

- Part(a) does not extend to unions of infinite collections of sets, i.e. the inclusion in part (b) can be strict if $\mathcal{C}$ is an infinite collection of sets.
- Strict inclusion can occur in (c) even when $\mathcal{C}$ is a finite collection of sets.
- Part (d) does not extend to intersections of infinite collections of sets, i.e. the inclusion in part (e) can be strict if $\mathcal{C}$ is an infinite collection of sets.
- Strict inclusion can occur in (f) even when $\mathcal{C}$ is a finite collection of sets.

Proof of Prop. M.4: We shall prove (d), (e), and (f). Then (a), (b), and (c) will follow from DeMorgan's Laws and Prop. M.3. We begin with (e). Let $x \in \text{int}(\bigcap_{S\in\mathcal{C}} S)$ be given. Then we may choose $\delta > 0$ such that $B_\delta(x) \subset \bigcap_{S\in\mathcal{C}} S$, i.e. $B_\delta(x) \subset S$ for all $S \in \mathcal{C}$. This implies that $x \in \text{int}(S)$ for all $S \in \mathcal{C}$, i.e. $x \in \bigcap_{S\in\mathcal{C}} \text{int}(S)$.

To prove (d), we observe that $\text{int}(A \cap B) \subset (\text{int}(A)) \cap (\text{int}(A)) \cap (\text{int}(B))$. Let $x \in (\text{int}(A)) \cap (\text{int}(B))$ be given. Then we may choose $\delta_1, \delta_2 > 0$ such that $B_{\delta_1}(x) \subset A$

and $B_{\delta_2}(x) \subset B$. Put $\delta = \min\{\delta_1, \delta_2\}$ and notice that $\delta > 0$. Then $B_\delta(x) \subset A \cap B$ and consequently $x \in \text{int}(A \cap B)$.

To prove (f), let $x \in \underset{S \in \mathcal{C}}{\cup} S$ be given. Then we may choose $\hat{S} \in \mathcal{C}$ such that $x \in \text{int}(\hat{S})$. Consequently we may choose $\delta > 0$ such that $B_\delta(x) \subset \hat{S}$, from which we conclude that $B_\delta(x) \subset \underset{S \in \mathcal{C}}{\cup} S$, i.e. $x \in \underset{S \in \mathcal{C}}{\cup} \text{int}(S)$. $\square$

Prop. M.5:

(a) The union of any collection of open sets is open.

(b) The intersection of any finite collection of open sets is open.

(c) The intersection of any collection of closed sets is closed.

(d) The union of any finite collection of closed sets is closed.

Proof: We shall prove (a) and (b). Then (c) and (d) will follow from DeMorgan's Laws and the corollary to Prop. M.3.

To prove (a), let O be a collection of open subsets of X and put $U = \underset{S \in \mathcal{O}}{\cup} S$. Since $S = \text{int}(S)$ for every $S \in \mathcal{O}$, we have

$$U = \bigcup_{S \in \mathcal{O}} \text{int}(S) \subset \text{int}(\underset{S \in \mathcal{O}}{S}) = U,$$

and U is open.

To prove (b), let $\mathcal{F}$ be a finite collection of open subsets of X. If $\mathcal{F}$ is empty, then the intersection of $\mathcal{F}$ is X and we are done. Suppose $\mathcal{F} = \{S_1, S_2, \cdots, S_n\}$ where each S_i is open. Then, by part (d) of Prop. M.4 (and a straightforward induction argument) we have

$$\text{int}(\bigcup_{i=1}^{N} S_i) = \bigcup_{i=1}^{N} \text{int}(S_i) = \bigcup_{i=1}^{N} S_i. \ \square$$

The following proposition is of interest in its own right and it will help motivate the definitions of interior and closure in a topological space.

Prop. M.6: Let $S \subset X$.

(a) The interior of S is equal to the union of all open subsets of S.

(b) The closure of S is equal to the intersection of all closed sets C such that $S \subset C$.

Proof: To prove (a), let $\mathcal{O}$ denote the collection of all open subsets of S. Let $U \in \mathcal{O}$ be given and observe that $U \subset \text{int}(S)$. (Indeed, given $x \in U$, we may choose $\delta > 0$ such that $B_\delta(x) \subset U \subset S$, which implies that $x \in \text{int}(S)$.) It follows that

$$\bigcup_{U \in \mathcal{O}} U \subset \text{int}(S).$$

On the other hand, $\text{int}(S) \in \mathcal{O}$, so that

$$\text{int}(S) \subset \bigcup_{U \in \mathcal{O}} U.$$

We conclude that

$$\text{int}(S) = \bigcup_{U \in \mathcal{O}} U.$$

To prove (b), we apply the result of part (a) to S^c and use Prop. M.3 and DeMorgan's laws. $\Box$

Def: Let $S \subset X$. We say that S is *dense* (in X) provided $\text{cl}(S) = X$.

Notice that S is dense in X if and only if $B_\delta(x) \cap S \neq \emptyset$ for every $x \in X, \delta > 0$. In other words, S is dense in X if and only if every ball (no matter how small and no matter where the center is) contains points that belong to S.

Def: We say that X is *separable* provided that X has a countable dense subset.

Separable metric spaces are structurally simpler than nonseparable ones. In particular, for a separable metric space there is a countable collection of open sets that can be used to generate every open set.

Prop. M.7: X is separable if and only if there is a countable collection $\mathcal{C}$ of open subsets of X such that every open subset of X can be expressed as a union of members of $\mathcal{C}$.

Proof: Assume that X is separable and choose a countable dense set $S \subset X$. Let

$$\mathcal{C} = \{B_\delta(x) : x \in S, \delta \in \mathbb{Q}, \delta > 0\},$$

and observe that $\mathcal{C}$ is a countable collection of open subsets of X. Let U be an open subset of X. To show that U is a union of members of $\mathcal{C}$ it suffices to show that

$$\forall y \in U,\ \exists V \in \mathcal{C},\ y \in V \subset U. \tag{1}$$

To establish (1), let $y \in U$ be given. Since U is open, we may choose $\eta > 0$ such that $B_\eta(y) \subset U$. Since $\mathbb{Q}$ is dense in $\mathbb{R}$, we may choose $\delta \in \mathbb{Q} \cap (0, \eta]$. Notice that $B_\delta(y) \subset U$. Since $y \in \text{cl}(S)$, we may choose $x \in S$ such that $\rho(x, y) < \frac{\delta}{2}$. Observe that $B_{\frac{\delta}{2}}(x) \in \mathcal{C}$ and that

$$B_{\frac{\delta}{2}}(x) \subset B_\delta(y) \subset U,$$

and consequently (1) is satisfied.

Assume now that $\mathcal{C}$ is a countable collection of open subsets of X having the property that every open subset of X is a union of members of $\mathcal{C}$. We may assume that $\emptyset \notin \mathcal{C}$. (Indeed, if $\emptyset \in \mathcal{C}$, we may remove it from the collection, and we still have a countable collection of open sets having the property that every open set can be expressed as a union of members of $\mathcal{C}$.) For each $U \in \mathcal{C}$, we may choose $x_U \in U$. Put

$$S = \{x_U : U \in \mathcal{C}\},$$

and observe that S is countable. To see that S is dense, let $x \in X$ and $\delta > 0$ be given. Since $B_\delta(x)$ is open, it can be expressed as a union of members of $\mathcal{C}$ and consequently we may choose $U \in \mathcal{C}$ such that $x \in U \subset B_\delta(x)$. It follows that $x_U \in B_\delta(x)$ and consequently $x \in \mathrm{cl}(S)$. $\square$

Def: Let $\mathcal{U}$ be a collection of open subsets of X. We say that $\mathcal{U}$ covers S provided that

$$S \subset \bigcup_{U \in \mathcal{U}} U.$$

Notice that if a collection of open sets covers the entire space X then the union of this collection must equal X. A simple, but very useful, characterization of separable metric spaces is contained in the following proposition.

Prop. M.8: X is separable if and only if for every $\epsilon > 0$, there is a countable collection of balls of radius ϵ that covers X.

Proof: Assume that X is separable and let $S \subset X$ be a countable dense set. Let $\epsilon > 0$ be given. Let $\mathcal{O} = \{B_\epsilon(z) : z \in S\}$ and observe that $\mathcal{O}$ is countable. We need to show that every $x \in X$ belongs to some member of $\mathcal{O}$. Let $x \in X$ be given. Since S is dense in X, we may choose $z_x \in S$ such that $\rho(x, z_x) < \epsilon$. This implies that $x \in B_\epsilon(z_x) \in \mathcal{O}$.

Assume now that for every $\epsilon > 0$ there is a countable collection of balls of radius ϵ that covers X. Then, for every $n \in \mathbb{N}$ we may choose a sequence $\{x_{n,k}\}_{k=1}^\infty$ such that the collection $\{B_{\frac{1}{n}}(x_{n,k}) : k \in \mathbb{N}\}$ of balls covers X. Let

$$S = \{x_{n,k} : n, k \in \mathbb{N}\},$$

and observe that S is countable. To prove that S is dense, let $x \in X, \delta > 0$ be given. We need to show that there exist $n, k \in \mathbb{N}$ such that $\rho(x, x_{n,k}) < \delta$. Choose $N > \frac{1}{\delta}$. Then we may choose $k \in \mathbb{N}$ such $x \in B_{\frac{1}{N}}(x_{N,k})$ and consequently we have $\rho(x, x_{N,k}) < \frac{1}{N} < \delta$. $\square$

Def: Let $S \subset X$. We say that S is *compact* provided that every collection of open sets that covers S has a finite subcollection that also covers S.

The definition of compactness may seem a bit strange if you have not encountered this concept before. However, it is one of the most important notions in basic analysis. Suppose that $K \subset X$ is compact. Then the collection of open sets

$$\{B_{\delta_x}(x) : x \in K\}$$

covers K, no matter how the radii δ_x are chosen. In particular, the radii can be chosen sufficiently small so that quantities of interest are controlled on each ball. By compactness, one can choose a finite set $F \subset K$ such that the collection of balls

$$\{B_{\delta_x}(x) : x \in F\}$$

covers K. In particular, K is covered by a *finite* number of balls of conveniently chosen radii.

Prop. M.9: Let K be a compact subset of X. Then K is closed.

Proof: We may assume that K is nonempty. (If $K = \emptyset$ there is nothing to prove.) By the corollary to Prop. M.3, it suffices to show that K^c is open. Let $z \in K^c$ be given. For each $x \in K$, we put

$$\delta_x = \frac{1}{3}\rho(z,x), \quad W_x = B_{\delta_x}(x), \quad V_x = B_{\delta_x}(z).$$

Observe that $\{W_x : x \in K\}$ is a collection of open sets that covers K. We may choose $x_1, x_2, \cdots, x_N \in X$ such that

$$K \subset \bigcup_{i=1}^{N} W_{x_i}.$$

Now let

$$\delta^* = \min\{\delta_{x_i} : i = 1, 2, \cdots, N\}.$$

By construction $B_{\delta^*}(z) \subset K^c$. It follows that $z \in \mathrm{int}(K^c)$. Since $z \in K^c$ was arbitrary, we conclude that K^c is open. $\square$

We note for future reference that in a general topological space, compact sets need not be closed.

Prop M.10: Let K be a compact subset of X and $S \subset K$. If S is closed, then S is compact.

Proof: Assume that S is closed. Let $\mathcal{O}$ be a collection of open sets that covers S and put $\mathcal{O}' = \mathcal{O} \cup \{S^c\}$. By the corollary to Prop. M.3, S^c is open, so $\mathcal{O}'$ is a collection of open sets that covers K. (In fact, it covers X.) Since K is compact, we may choose a finite subcollection $\mathcal{F}'$ of $\mathcal{O}'$ that also covers K. (This subcollection also covers S, but it may not be a subcollection of $\mathcal{O}$ because S^c may well belong to $\mathcal{F}'$, but not to $\mathcal{O}$.) Now, put $\mathcal{F} = \mathcal{F}' \backslash \{S^c\}$ and observe that $\mathcal{F}$ is a finite subcollection of $\mathcal{O}$ that covers S. $\square$

Let (X, ρ), (Y, σ), and (Z, λ) be metric spaces.

Def: Let $f : X \to Y$ and $x \in X$ be given. We say that f is *continuous at* x provided that for every $\epsilon > 0$ there exists $\delta > 0$ such that $\sigma(f(y), f(x)) < \epsilon$ for all $y \in X$ with $\rho(y, x) < \delta$. We say that f is *continuous* (or *continuous on* X) provided that f is continuous at each point $x \in X$.

Def: We say that $f : X \to Y$ is *uniformly continuous* provided that for every $\epsilon > 0$ there exists $\delta > 0$ such that $\sigma(f(y), f(x)) < \epsilon$ for all $x, y \in X$ with $\rho(x, y) < \delta$.

Clearly, uniform continuity implies continuity; the converse implication is false in general. The difference between these two concepts is that in the definition of continuity, δ can depend on the point x as well as on ϵ, whereas for uniform continuity, δ can depend only on ϵ.

Prop. M.11: $f : X \to Y$ is continuous if and only if for every open set V in (Y, σ) the set

$$\{x \in X : f(x) \in V\}$$

is open in (X, ρ).

Cor: If $f : X \to Y$ and $g : Y \to Z$ are continuous then $g \circ f : X \to Z$ is continuous.

Prop. M.12: Assume that $f : X \to Y$ is continuous and $K \subset X$ is compact in (X, ρ). Then $\{f(x) : x \in K\}$ is compact in (Y, σ).

Prop. M.13: Assume that $f : X \to Y$ is continuous and that X is compact. Then f is uniformly continuous.

We denote the natural numbers by $\mathbb{N} = \{1, 2, 3, \cdots\}$. (In particular, our convention is that 1 is the smallest natural number, and therefore 0 is not a natural number. This will not be an important issue for anything, but is explicitly mentioned here because many authors consider 0 to be a natural number.)

A sequence in (X, ρ) is a mapping $x : \mathbb{N} \to X$. It is customary to write x_n in place of $x(n)$ and to denote the sequence by $\{x_n\}_{n=1}^{\infty}$.

Def: Let $\{x_n\}_{n=1}^{\infty}$ be a sequence in (X, ρ) and $l \in X$ be given. We say that l is a *limit* of $\{x_n\}_{n=1}^{\infty}$ provided that for every $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that $\rho(x_n, l) < \epsilon$ for all $n \geq N$. A sequence in (X, ρ) is said to be *convergent* if it has a limit $l \in X$.

Remark: In a metric space, a sequence can have at most one limit. (We note for future reference that in certain topological spaces, sequences can have more than one limit.)

Proof: To prove the remark, let $\{x_n\}_{n=1}^{\infty}$ be a sequence in X and $l, L \in X$ be given.

Suppose that $x_n \to l$ and $x_n \to L$ as $n \to \infty$. We shall show that $l = L$. Let $\epsilon > 0$ be given. We may choose $N_1, N_2 \in \mathbb{N}$ such that

$$\rho(x_n, l) < \epsilon \quad \text{for all } n \geq N_1 \quad \text{and} \quad \rho(x_n, L) < \epsilon \quad \text{for all } n \geq N_2.$$

Let us put $N = \max(N_1, N_2)$. Then, by the triangle inequality, we have

$$\rho(l, L) \leq \rho(l, x_N) + \rho(x_N, L) < \epsilon + \epsilon = 2\epsilon.$$

It follows that

$$\rho(l, L) < 2\epsilon \quad \text{for all } \epsilon > 0.$$

Since ρ is nonnegative, we conclude that $\rho(l, L) = 0$, and this implies that $l = L$. $\square$

Remark: If l is the limit of a sequence $\{x_n\}_{n=1}^{\infty}$ we write $x_n \to l$ as $n \to \infty$ or $\lim_{n\to\infty} x_n = l$.

Def: A sequence $\{x_n\}_{n=1}^{\infty}$ in (X, ρ) is said to be a *Cauchy sequence* (or *fundamental sequence*) provided that for every $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that $\rho(x_n, x_m) < \epsilon$ for all $m, n \geq N$.

Every convergent sequence in a metric space is a Cauchy sequence. (You should prove this as an exercise for yourself.) In some metric spaces, every Cauchy sequence is convergent. However, there are metric spaces in which some Cauchy sequences fail to be convergent.

Def: A metric space is said to be *complete* if every Cauchy sequence is convergent.

In order to use the definition to prove that a sequence is convergent, one needs to know the limit in advance. In a complete metric space, one can prove that a sequence is convergent, by showing that it is a Cauchy sequence. One does not need to know anything about possible values for the limit to show that a sequence is a Cauchy sequence.

In some sense, Cauchy sequences are "trying to converge". In an incomplete metric space, a Cauchy sequence might be trying to converge to something "outside of the metric space". In fact, it is always possible to make an incomplete metric space into a complete one by adding enough additional elements. This process is called *completion* and will be discussed later. It is some sort of "metatheorem" that if you have a Cauchy sequence $\{x_n\}_{n=1}^{\infty}$ in an incomplete metric space X and there is a "natural candidate" $l \in X$ for the limit of this sequence, then indeed $x_n \to l$ as $n \to \infty$.

Def: By a *subsequence* of $\{x_n\}_{n=1}^{\infty}$ we mean a sequence of the form $\{x_{n_k}\}_{k=1}^{\infty}$ where $\{n_k\}_{k=1}^{\infty}$ is a strictly increasing sequence of natural numbers.

Clearly, every subsequence of a convergent sequence is convergent to the limit of the original sequence. Moreover, every subsequence of a Cauchy sequence is a Cauchy sequence. The following elementary result is sometimes useful.

Prop. M.14: Let $\{x_n\}_{n=1}^{\infty}$ be a sequence in X and $l \in X$ be given. If every subsequence of $\{x_n\}_{n=1}^{\infty}$ has in turn a subsequence that converges to l then $x_n \to l$ as $n \to \infty$.

As an illustration of the "metatheorem" referred to above, we have the following result.

Prop. M.15: Let $\{x_n\}_{n=1}^{\infty}$ be a Cauchy sequence in X, $\{x_{n_k}\}_{k=1}^{\infty}$ be a subsequence of $\{x_n\}_{n=1}^{\infty}$, and $l \in X$ be given. Assume that

$$x_{n_k} \to l \text{ as } n \to \infty.$$

Then $x_n \to l$ as $n \to \infty$.

Proof: Let $\epsilon > 0$ be given. Choose $N_1, N_2 \in \mathbb{N}$ such that

$$\rho(x_n, x_m) < \frac{\epsilon}{2} \text{ for all } m, n \geq N_1,$$

$$\rho(x_{n_k}, l) < \frac{\epsilon}{2} \text{ for all } k \geq N_2.$$

Observe that since $\{n_k\}_{k=1}^{\infty}$ is a strictly increasing sequence of natural numbers, we have

$$n_k \geq k \text{ for all } k \in \mathbb{N}.$$

Put $N = \max(N_1, N_2)$ and observe that $n_N \geq N_1$ and $N \geq N_2$. Then, for all $n \geq N$ we have

$$\rho(x_n, l) \leq \rho(x_n, x_{n_N}) + \rho(x_{n_N}, l) < \frac{\epsilon}{2} + \frac{\epsilon}{2}.$$

We conclude that $x_n \to l$ as $n \to \infty$. $\square$

In metric spaces, sequences can be used to characterize closedness, compactness, and continuity. (However, it is important to note that, in topological spaces, sequences cannot always be used to characterize these properties.)

Prop. M.16: Let $S \subset X$ and $l \in X$ be given. Then $l \in \mathrm{cl}(S)$ if and only if there is a sequence $\{x_n\}_{n=1}^{\infty}$ such $x_n \in S$ for every $n \in \mathbb{N}$ and $x_n \to l$ as $n \to \infty$.

Prop. 3.17: Let $S \subset X$ be given. Then S is closed if and only if for every convergent sequence $\{x_n\}_{n=1}^{\infty}$ such that $x_n \in S$ for every $n \in \mathbb{N}$ we have $\lim\limits_{n\to\infty} x_n \in S$.

Prop. M.18: Let $K \subset X$ be given. Then K is compact if and only if every sequence $\{x_n\}_{n=1}^{\infty}$ such that $x_n \in K$ for all $n \in \mathbb{N}$ has a subsequence that converges to an element of K.

Remark: You may encounter the term "sequentially compact" in the literature. The meaning is not completely standard. Some authors say that a set K is sequentially compact provided every sequence of points from K has a convergent subsequence, while other authors say that K is sequentially compact provided that every sequence of points from K has a subsequence that converges to an element of K. For subsets of metric spaces, the latter concept is equivalent to compactness, while the former is not. (Of course, when applied to an entire metric space, the two notions of sequential compactness are the same.)

Prop. M.19: Let $f : X \to Y$ and $l \in X$ be given. Then f is continuous at l if and only if $f(x_n) \to f(l)$ in (Y, σ) as $n \to \infty$ for every sequence $\{x_n\}_{n=1}^{\infty}$ in (X, ρ) such that $x_n \to l$ as $n \to \infty$.

Although continuous functions map convergent sequences to convergent sequences, they need not map Cauchy sequences to Cauchy sequences. However, uniformly continuous functions do map Cauchy sequences to Cauchy sequences.

Prop. M.20: Assume that $f : X \to Y$ is uniformly continuous and $\{x_n\}_{n=1}^{\infty}$ is a Cauchy sequence in (X, ρ). Then $\{f(x_n)\}_{n=1}^{\infty}$ is a Cauchy sequence in (Y, σ).

Baire Category

Many of the deepest and most important consequences of completeness can be obtained from the following result of Baire.

Baire's Theorem: Let (X, ρ) be a complete metric space and $\{U_n\}_{n=1}^{\infty}$ be a sequence of subsets of X such that for every $n \in \mathbb{N}$, U_n is open and dense in X. Then

$$\bigcap_{n=1}^{\infty} U_n \text{ is dense in } X.$$

For the sake of completeness (no pun intended) we give a proof of Baire's Theorem. The following simple comment may make it easier to follow the proof. Let S be a subset of X. If $z \in \text{int}(S)$ then we may choose $\eta > 0$ such that $\text{cl}(B_\eta(z)) \subset S$. The validity of the comment follows immediately from the observation that

$$\text{cl}(B_\eta(z)) \subset \{y \in X : \rho(y, z) \leq \eta\} \subset B_{2\eta}(z).$$

Proof of Baire's Theorem: Let $x \in X$ and $\delta > 0$ be given. It suffices to show that $B_\delta(x)$ contains a point that belongs to each of the sets U_n. We shall accomplish this by constructing a "shrinking" sequence $\{B_{\delta_n}(x_n)\}_{n=1}^{\infty}$ of balls. We show that the sequence $\{x_n\}_{n=1}^{\infty}$ of centers converges to a point l that belongs to $B_\delta(x)$ and each of the U_n. For convenience, we choose that radii to satisfy $\delta_n \leq \frac{1}{n}$.

Since U_1 is dense in X, we may choose a point $x_1 \in U_1 \cap B_\delta(x)$. Since $U_1 \cap B_\delta(x)$ is open and $x_1 \in U_1 \cap B_\delta(x)$, we may choose $\delta_1 \in (0, 1]$ such that

$$B_{\delta_1}(x_1) \subset U_1 \cap B_\delta(x).$$

Since U_2 is dense, we may choose a point $x_2 \in U_2 \cap B_{\delta_1}(x_1)$. Moreover, since $U_2 \cap B_{\delta_1}(x_1)$ is open, we may choose $\delta_2 \in (0, \frac{1}{2}]$ such that

$$\mathrm{cl}(B_{\delta_2}(x_2)) \subset U_2 \cap B_{\delta_1}(x_1).$$

Proceeding by induction, we obtain, for each $n \in \mathbb{N}$ a point $x_n \in X$ and $\delta_n > 0$ such that

- $B_{\delta_n}(x_n) \subset U_n$,
- $\mathrm{cl}(B_{\delta_{n+1}}(x_{n+1})) \subset B_{\delta_n}(x_n)$,
- $\delta_n \leq \frac{1}{n}$.

Let $N \in \mathbb{N}, m, n \in \mathbb{N}$ with $m, n \geq N$ be given. Then $x_m, x_n \in B_{\delta_N}(x_n)$, so that

$$\rho(x_m, x_n) < 2\delta_N \leq \frac{2}{N},$$

which implies that $\{x_n\}_{n=1}^{\infty}$ is a Cauchy sequence. Since X is complete, we may choose $l \in X$ such that $x_n \to l$ as $n \to \infty$. Since $x_n \in B_{\delta_{N+1}}(x_{N+1})$ for all $n \geq N+1$, it follows that

$$l \in \mathrm{cl}(B_{\delta_{N+1}}(x_{N+1})) \subset B_{\delta_N}(x_N) \subset U_N.$$

Since $N \in \mathbb{N}$ was arbitrary, we conclude that

$$l \in \bigcap_{n=1}^{\infty} U_n.$$

Since $B_{\delta_1}(x_1) \subset B_\delta(x)$ and $l \in B_{\delta_1}(x_1)$, we conclude that $l \in B_\delta(x)$. $\square$

Def: Let (X, ρ) be a metric space and $S \subset X$. We say that S is

(i) *nowhere dense* provided $\mathrm{int}(cl(S)) = \emptyset$.

(ii) *meager* if S can be expressed as a countable union of nowhere dense sets.

(iii) *residual* if S^c is meager.

Remark: It is traditional to say that meager sets are of the *first category* and non-meager sets are of the *second category*.

The following result is an easy consequence of Baire's Theorem.

Baire Category Theorem: Assume that $X \neq \emptyset$ and that (X, ρ) is complete. Then X cannot be expressed as a countable union of nowhere dense sets.

The way that we shall typically use this theorem in practice is as follows. We express

$$X = \bigcup_{n=1}^{\infty} S_n,$$

where each set S_n is closed. If $X \neq \emptyset$ and (X, ρ) is complete, then we can conclude that $\mathrm{int}(S_N) \neq \emptyset$ for some $N \in \mathbb{N}$.

Isometries, Homeomorphisms, and Completions

Let (X, ρ) and (Y, σ) be metric spaces.

Def: A mapping $f : X \to Y$ is said to be an *isometry* provided that $\sigma(f(x), f(y)) = \rho(x, y)$ for all $x, y \in X$.

In other words, isometries are mappings that preserve the distance between each pair of points.

Remark: Every isometry is injective and continuous. If an isometry is surjective (and hence bijective), then the inverse is also an isometry.

Def: We say that (X, ρ) and (Y, σ) are

(i) *homeomorphic* provided that there is a bijection $f : X \to Y$ such that f and f^{-1} are continuous. (Such a mapping f is called a *homeomorphism.*)

(ii) *uniformly homeomorphic* provided that there is a bijection $f : X \to Y$ such that f and f^{-1} are uniformly continuous. (Such a mapping f is called a *uniform homeomorphism.*)

(iii) *isometric* provided that there is a surjective isometry $f : X \to Y$.

Properties (such as compactness and separability) that are preserved under every homeomorphism are called *topological properties.* Some important properties, such as completeness, are not necessarily preserved under homeomorphisms. Properties that are preserved under every uniform homeomorphism are called *uniform properties.* It follows from Prop. M.20 that completeness is a uniform property. Metric spaces that are isometric are essentially indistinguishable from the point of view of metric space theory.

Theorem M.21: Let (X, ρ) be a metric space. Then there is a complete metric space $(\hat{X}, \hat{\rho})$ and a set $\hat{W} \subset \hat{X}$ such that $\hat{W}$ is dense in $(\hat{X}, \hat{\rho})$ and (X, ρ) is isometric to $(\hat{W}, \hat{\rho})$.

Remark: It is standard practice to call the metric space $(\hat{W}, \hat{\rho})$ *the completion* of (X, ρ). Of course, there may be more than one completion of a given metric space, but it is clear that all completions of a given metric space are isomorphic.

Subspaces

Let (X, ρ) be a metric space. If $A \subset X$ then, of course, (A, ρ) is also a metric space. We say that (A, ρ) is a *subspace* of (X, ρ). We need to be a bit careful when talking about topological properties in such situations because the interior and closure of S in (X, ρ) may be different from the interior and closure of S in (A, ρ). A very simple example is given below. If $S \subset A \subset X$, we write $\mathrm{int}^A(S)$ and $\mathrm{cl}^A(S)$ for the interior and closure of S as a subset of the metric space (A, ρ). We refer to $\mathrm{int}^A(S)$ and $\mathrm{cl}^A(S)$ as the interior and closure of S *relative* to A. We say that S is *open relative to* A provided that $S = \mathrm{int}^A(S)$ (i.e., S is an open subset of (A, ρ)) and that S is *closed relative to* A provided that $S = \mathrm{cl}^A(S)$ (i.e., S is a closed subset of (A, ρ)).

Example: Let $X = \mathbb{R}$, equipped with the standard metric and put $A = \mathbb{Q}$. Then we have

$$\mathrm{int}^{\mathbb{R}}(\mathbb{Q}) = \emptyset, \quad \mathrm{cl}^{\mathbb{R}}(\mathbb{Q}) = \mathbb{R}, \quad \mathrm{int}^{\mathbb{Q}}(\mathbb{Q}) = \mathrm{cl}^{\mathbb{Q}}(\mathbb{Q}) = \mathbb{Q}.$$

Prop. M.22: Let $S \subset A \subset X$. Then

(a) S is open relative to A if and only if there is a set $U \subset X$ such that U is open in (X, ρ) and $S = U \cap A$,

(b) S is closed relative to A if and only if there is a set $C \subset X$ such that C is closed in (X, ρ) and $S = C \cap A$,

(c) $\mathrm{cl}^A(S) = A \cap \mathrm{cl}^X(S)$.

Prop. M.23: Let $A \subset X$. If (A, ρ) is complete then A is closed in (X, ρ).

Prop. M.24: Assume that (X, ρ) is complete and that $A \subset X$ is closed in (X, ρ). Then (A, ρ) is complete.

We have seen in the simple example above that if $S \subset A \subset X$ then S may be open in (A, ρ) but fail to be open in (X, ρ). A similar comment applies to sets that are closed in (A, ρ). Compactness, however, is a more robust property as the proposition below indicates.

Prop. M.25: Let $S \subset A \subset X$. Then S is compact in (X, ρ) if and only if S is compact in (A, ρ).

Remark: Many authors define compactness first for entire metric spaces (X, ρ) and then define a set $S \subset X$ to be compact provided that the metric space (S, ρ) is compact. It follows from Prop. 4.5 that this definition of compact set is equivalent to the one given earlier.

Prop. M.26: Let $A \subset X$. If (X, ρ) is separable, then so is (A, ρ).

Prop. M.27: Let $A \subset X$ and put $\overline{A} = \mathrm{cl}^X(A)$. Then (A, ρ) is sepparable if and only if $(\overline{A}, \rho)$ is separable.

It is convenient to define separability for subsets of (X, ρ).

Def: Let $S \subset X$. We say that S is *separable* provided that the metric space (S, ρ) is separable. (Notice that this is equivalent to the requirement that there is a countable set $D \subset S$ such that $\mathrm{cl}^S(D) = S$.)

Since $\mathrm{cl}^A(S) = A \cap \mathrm{cl}^X(S)$ when $S \subset A \subset X$ we have the following simple proposition.

Prop. M.28: Let $S \subset A \subset X$. Then S is separable as a subset of (X, ρ) if and only if it is separable as a subset of (A, ρ).

Totally Bounded Sets and Compactness

Let (X, ρ) be a metric space.

Def: A set $S \subset X$ is said to be *bounded* provided that there exists $M \in \mathbb{R}$

$$\rho(x, y) \leq M \quad \text{for all } x, y \in S.$$

It is easy to see that a set S is bounded if and only if there exist $x \in X, \delta > 0$ such that $S \subset B_\delta(x)$. We have already seen that compact sets are closed. It is straightforward to show that every compact set is bounded. If $\mathbb{R}$ is equipped with the standard metric, then every closed bounded set is compact. In general, however, even if (X, ρ) is assumed to be complete, there may be closed bounded sets that fail to be compact. There is a stronger property, known as total boundedness, such that in a complete metric space, a set is compact if and only if it is closed and totally bounded.

Def: A set $S \subset X$ is said to be *totally bounded* provided that for every $\epsilon > 0$, S can be covered by a finite number of balls of radius ϵ.

It is clear that every subset of a totally bounded set is totally bounded. If a set S is totally bounded, then $\mathrm{cl}(S)$ is also totally bounded because the closure of a finite union is the union of the closures and $\mathrm{cl}(B_\eta)(x) \subset B_{2\eta}(x)$. Since this observation will be used later, we state it as a proposition.

Prop. M.29: Let (X, ρ) be a metric space and $S \subset X$. Then S is totally bounded if and only if $\mathrm{cl}(S)$ is totally bounded.

Like compactness and separability, the property of being totally bounded does not depend on which larger space the set lives in, provided that the larger metric spaces under consideration are all subspaces of a single metric space. More precisely we have the following analogue of Propositions M.25 and M.28.

Prop. M.30: Let (X, ρ) be a metric space and $S \subset A \subset X$. Then S is totally bounded in (X, ρ) if and only if it is totally bounded in (A, ρ).

Remark: Many authors define total boundedness first for entire metric spaces (X, ρ) and then define a set $S \subset X$ to be totally bounded provided that the metric space (S, ρ) is totally bounded. In view of Prop. M.30, this definition of totally bounded set is equivalent to the one given here.

Every compact metric space is complete and separable.

Prop. M.31: If (X, ρ) is compact, then it is complete.

Proof: Assume that (X, ρ) is compact and let $\{x_n\}_{n=1}^{\infty}$ be a Cauchy sequence in X. Then, by Prop. 3.6 there is a convergent subsequence $\{x_{n_k}\}_{k=1}^{\infty}$. By Prop. M.15 the sequence $\{x_n\}_{n=1}^{\infty}$ is convergent. $\square$

Prop. M.32: If (X, ρ) is compact then it is separable.

Proof: Assume that (X, ρ) is compact and let $\epsilon > 0$ be given. Then, since the collection $\{B_\delta(x) : x \in X\}$ of open sets covers X, we may choose a finite subcollection that covers X. Separability of X follows from Prop. M.8. $\square$

It is easy to see that every compact set is totally bounded. Indeed, suppose that $K \subset X$ is compact and let $\epsilon > 0$ be given. Then the collection $\{B_\epsilon(x) : x \in K\}$ of open sets covers K and consequently has a finite subcollection that also covers K. It is also straightforward to show that every totally bounded set is bounded. The following characterization of compact subsets of complete metric spaces is of fundamental importance.

Theorem M.33: Assume that (X, ρ) is complete and let $S \subset X$. Then S is compact if and only if S is closed and totally bounded.

Lecture Notes for Week 1

Roughly speaking *functional analysis* is the study of algebra, topology, and geometry of linear spaces that need not be finite dimensional, and of mappings between such spaces. In classical functional analysis, emphasis is usually placed on linear mappings.

The prerequisites for this course are a thorough understanding of linear algebra and topology (especially metric spaces), and familiarity and comfort with the Lebesgue integral. Students are (of course) expected to be proficient at writing careful and complete mathematical proofs. I have posted a set of notes that contains all of the results concerning metric spaces that I anticipate needing during the course.

I want to present a certain significant body of knowledge and also to familiarize you with a number of tools and standard arguments that are useful for proving results in functional analysis. In order to balance these objectives, I will skip the proofs of certain theorems (and simply give references instead) if I feel that the techniques needed for the proof have been (or will be) well exposed through other results whose proofs are discussed in detail elsewhere in the course.

The study of functional analysis was motivated by the fact that certain equations (such as Volterra integral equations) involving unknown functions have a structural similarity with systems of ordinary systems of (algebraic) linear equations. In order to exploit this structural similarity, one is led to consider linear spaces (or vector spaces) whose elements are functions rather than "ordinary vectors". Such *function spaces* are inherently infinite dimensional (meaning that there is no finite set of elements whose linear combinations span the entire space).

We shall study linear spaces (over $\mathbb{R}$ or $\mathbb{C}$), equipped with various topologies that are naturally adapted to the linear structure. We shall begin by looking at topologies that are induced by norms.

In finite-dimensional linear spaces (over $\mathbb{R}$ or $\mathbb{C}$), all norms are equivalent, all linear mappings are continuous, and the closed unit ball is compact. The situation is dramatically different in infinite dimensions. Moreover, every finite-dimensional linear space (over $\mathbb{R}$ or $\mathbb{C}$) is complete (in the sense that every Cauchy sequence is convergent). Infinite-dimensional normed linear spaces need not be complete. A number of important results concerning infinite-dimensional normed linear spaces require one or more of the spaces in question to be complete.

Baire Category

Many of the deepest and most important consequences of completeness can be obtained from the following result of Baire.

Theorem 1.1 (Baire): Let (X, ρ) be a complete metric space and $\{U_n\}_{n=1}^{\infty}$ be a sequence of subsets of X such that for every $n \in \mathbb{N}$, U_n is open and dense in X. Then

$$\bigcap_{n=1}^{\infty} U_n \text{ is dense in } X.$$

For the sake of completeness (no pun intended) we give a proof of Baire's Theorem. The following simple comment may make it easier to follow the proof. Let S be a subset of X. If $z \in \operatorname{int}(S)$ then we may choose $\eta > 0$ such that $\operatorname{cl}(B_\eta(z)) \subset S$. The validity of the comment follows immediately from the observation that

$$\operatorname{cl}(B_\eta(z)) \subset \{y \in X : \rho(y, z) \leq \eta\} \subset B_{2\eta}(z).$$

Proof of Baire's Theorem: Let $x \in X$ and $\delta > 0$ be given. It suffices to show that $B_\delta(x)$ contains a point that belongs to each of the sets U_n. We shall accomplish this by constructing a "shrinking" sequence $\{B_{\delta_n}(x_n)\}_{n=1}^{\infty}$ of balls. We show that the sequence $\{x_n\}_{n=1}^{\infty}$ of centers converges to a point l that belongs to $B_\delta(x)$ and to each of the U_n. For convenience, we choose the radii to satisfy $\delta_n \leq \frac{1}{n}$.

Since U_1 is dense in X, we may choose a point $x_1 \in U_1 \cap B_\delta(x)$. Since $U_1 \cap B_\delta(x)$ is open and $x_1 \in U_1 \cap B_\delta(x)$, we may choose $\delta_1 \in (0, 1]$ such that

$$B_{\delta_1}(x_1) \subset U_1 \cap B_\delta(x).$$

Since U_2 is dense, we may choose a point $x_2 \in U_2 \cap B_{\delta_1}(x_1)$. Moreover, since $U_2 \cap B_{\delta_1}(x_1)$ is open, we may choose $\delta_2 \in (0, \frac{1}{2}]$ such that

$$\operatorname{cl}(B_{\delta_2}(x_2)) \subset U_2 \cap B_{\delta_1}(x_1).$$

Proceeding by induction, we obtain, for each $n \in \mathbb{N}$ a point $x_n \in X$ and $\delta_n > 0$ such that

- $B_{\delta_n}(x_n) \subset U_n$,
- $\operatorname{cl}(B_{\delta_{n+1}}(x_{n+1})) \subset B_{\delta_n}(x_n)$,
- $\delta_n \leq \frac{1}{n}$.

Let $m, n, N \in \mathbb{N}$ with $m, n \geq N$ be given. Then $x_m, x_n \in B_{\delta_N}(x_n)$, so that

$$\rho(x_m, x_n) < 2\delta_N \leq \frac{2}{N},$$

which implies that $\{x_n\}_{n=1}^{\infty}$ is a Cauchy sequence. Since X is complete, we may choose $l \in X$ such that $x_n \to l$ as $n \to \infty$. Since $x_n \in B_{\delta_{N+1}}(x_{N+1})$ for all $n \geq N+1$, it follows that

$$l \in \mathrm{cl}(B_{\delta_{N+1}}(x_{N+1})) \subset B_{\delta_N}(x_N) \subset U_N.$$

Since $N \in \mathbb{N}$ was arbitrary, we conclude that

$$l \in \bigcap_{n=1}^{\infty} U_n.$$

Since $B_{\delta_1}(x_1) \subset B_\delta(x)$ and $l \in B_{\delta_1}(x_1)$, we conclude that $l \in B_\delta(x)$. $\square$

Definition 1.2: Let (X, ρ) be a metric space and $S \subset X$. We say that S is

(i) *nowhere dense* provided $\mathrm{int}(cl(S)) = \emptyset$.

(ii) *meager* if S can be expressed as a countable union of nowhere dense sets.

(iii) *residual* if S^c is meager.

Remark 1.3: It is traditional to say that meager sets are of the *first category* and non-meager sets are of the *second category*.

The following result is an easy consequence of Baire's Theorem.

Theorem 1.4 (Baire Category Theorem): Assume that $X \neq \emptyset$ and that (X, ρ) is complete. Then X cannot be expressed as a countable union of nowhere dense sets.

The way that we shall typically use this theorem in practice is as follows. We express

$$X = \bigcup_{n=1}^{\infty} S_n,$$

where each set S_n is closed. If $X \neq \emptyset$ and (X, ρ) is complete, then we can conclude that $\mathrm{int}(S_N) \neq \emptyset$ for some $N \in \mathbb{N}$.

Linear Spaces and Norms

We shall use the symbol $\mathbb{K}$ to denote a field that is either $\mathbb{R}$ or $\mathbb{C}$.

Definition 1.5: Let X be a linear space over $\mathbb{K}$. By a *norm* on X we mean a function $\|\cdot\| : X \to \mathbb{R}$ satisfying

(i) $\forall x \in X, \quad \|x\| \geq 0,$

(ii) $\forall x \in X, \quad \|x\| = 0 \ \Leftrightarrow \ x = 0,$

(iii) $\forall x \in X, \alpha \in \mathbb{K}, \quad \|\alpha x\| = |\alpha| \cdot \|x\|$,

(iv) (triangle inequality) $\forall x, y \in X, \quad \|x + y\| \leq \|x\| + \|y\|$.

Definition 1.6: By a *normed linear space* (abbreviated *NLS*) we mean a pair $(X, \|\cdot\|)$ where X is a linear space over $\mathbb{K}$ and $\|\cdot\|$ is a norm on X.

Remark 1.7: It is useful to note that the triangle inequality (iv) can be reformulated as

(iv′) (reverse triangle inequality) $\forall x, y \in X, \quad |\|x\| - \|y\|| \leq \|x - y\|$.

In other words, if (i), (ii), and (iii) hold, then (iv) holds if and only if (iv′) holds. (You should verify this yourself as a simple exercise.)

In a NLS $(X, \|\cdot\|)$, the function $\rho : X \times X \to \mathbb{R}$ defined by

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

is a metric on X. It is called the *metric induced by the norm.* Metric and topological properties of $(X, \|\cdot\|)$ (such as completeness, continuity, compactness, etc.) are understood to be defined in terms of the metric induced by the norm unless stated otherwise. It follows immediately from (iv′) that the mapping $x \to \|x\|$ is continuous from (X, ρ) to $\mathbb{R}$. We note there are metrics on linear spaces that do not come from any norm.

Definition 1.8: A complete normed linear space is called a *Banach space* (or *B-space* for short).

There is sometimes confusion over terminology regarding subspaces of Banach spaces. It can happen that $(X, \|\cdot\|)$ is a Banach space, Y is a linear subspace of X, but $(Y, \|\cdot\|)$ is incomplete and hence is not a Banach space.

It seems useful to have a term to indicate that a subset of a linear space is itself a linear space, but need not have any special topological properties. Therefore we make the following definition.

Definition 1.9: Let X be a linear space. By a *linear manifold* in X we mean a nonempty subset of X that is closed under addition and scalar multiplication.

Proposition 1.10: Let $(X, \|\cdot\|)$ be a Banach space and Y be a linear manifold in X. Then $(Y, \|\cdot\|)$ is a Banach space if and only if Y is a (topologically) closed subset of X.

Proof: Assume first that Y is a closed subset of X and let $\{y_n\}_{n=1}^{\infty}$ be a Cauchy sequence in $(Y, \|\cdot\|)$. Then $\{y_n\}_{n=1}^{\infty}$ is also a Cauchy sequence in $(X, \|\cdot\|)$. Since

$(X, \|\cdot\|)$ is complete, we may choose $z \in X$ such that $\|y_n - z\| \to 0$ as $n \to \infty$. Since Y is closed, we conclude that $z \in Y$ and consequently $\{y_n\}_{n=1}^{\infty}$ is convergent (to z) in $(Y, \|\cdot\|)$. It follows that $(Y, \|\cdot\|)$ is complete.

Conversely, assume that $(Y, \|\cdot\|)$ is complete and let $\{y_n\}_{n=1}^{\infty}$ be a sequence such that $y_n \in Y$ for all $n \in \mathbb{N}$ and that $\{y_n\}_{n=1}^{\infty}$ is convergent in $(X, \|\cdot\|)$. Let $z = \lim\limits_{n\to\infty} y_n$. We need to show that $z \in Y$. Since $(Y, \|\cdot\|)$ is complete, we may choose $w \in Y$ such that $\|y_n - w\| \to 0$ as $n \to \infty$. Since

$$\|z - w\| \leq \|y_n - z\| + \|y_n - w\| \quad \text{for all } n \in \mathbb{N},$$

we conclude that $z = w$ and consequently $z \in Y$. $\square$

Remark 1.11: The proof of Proposition 1.10 shows that if $(X, \|\cdot\|)$ is a normed linear space and Y is a linear manifold in X such that $(Y, \|\cdot\|)$ is complete, then Y is a closed subset of $(X, \|\cdot\|)$.

In finite-dimensional NLS, every linear manifold is a closed set. If $(X, \|\cdot\|)$ is an infinite-dimensional NLS and Y is a linear manifold in X, then Y may or may not be a closed subset of $(X, \|\cdot\|)$. In the infinite-dimensional case, many important results concerning linear manifolds require topological closedness. Since a topologically closed linear manifold in a Banach space $(X, \|\cdot\|)$ is itself a Banach space when equipped with $\|\cdot\|$, the following definition seems reasonable. (In, general a substructure of a mathematical structure should itself satisfy the defining properties of the structure in question. In particular, a subspace of a Banach space should be a Banach space. The definition below is consistent with this idea and does not conflict with any conventions in the literature that I am aware of.)

Definition 1.12: Let $(X, \|\cdot\|)$ be a normed linear space. By a *closed subspace* of $(X, \|\cdot\|)$ we mean a linear manifold Y in X such that Y is a closed subset of X.

Example 1.13: Let $\mathbb{K} = \mathbb{R}$ and X be the set of all convergent sequences $x = (x_k | k \in \mathbb{N})$ of real numbers, equipped with the norm defined by

$$\|x\| = \sup\{|x_k| : k \in \mathbb{N}\}.$$

Sequences in X will be indicated by the notation $\{x^{(n)}\}_{n=1}^{\infty}$. You should verify as an exercise that $(X, \|\cdot\|)$ is complete. Let V denote the linear manifold consisting of all real sequences having only finitely many nonzero terms, i.e. the set of all sequences $(x_k | k \in \mathbb{N})$ such that $\{k \in \mathbb{N} : x_k \neq 0\}$ is finite. Consider the sequence $\{y^{(n)}\}_{n=1}^{\infty}$ in V defined by

$$y_k^{(n)} = \begin{cases} \frac{1}{k} & \text{if } k \leq n \\ \\ 0 & \text{if } k > n. \end{cases}$$

It is straightforward to show that $\{y^{(n)}\}_{n=1}^{\infty}$ is a Cauchy sequence in $(X, \|\cdot\|)$ and converges to the element $y \in X$ given by $y_k = \frac{1}{k}$ for all $k \in \mathbb{N}$. Since $y \notin V$, we conclude that V is not closed in $(X, \|\cdot\|)$ and consequently $(V, \|\cdot\|)$ is not complete.

Let W denote the set of all real sequences that converge to 0. It is straightforward to show that W is the (topological) closure of V in $(X, \|\cdot\|)$. It follows that $(W, \|\cdot\|)$ is complete. $\square$

Linear Independence and Bases

The situation with NLS is especially interesting in cases where the underlying linear space X is "infinite dimensional". We need to talk about concepts such as "linear combination", "linear independence", "basis", and "dimension" in a way that is suited to infinite-dimensional as well as finite-dimensional spaces.

Let X be a linear space over $\mathbb{K}$ and let $(x_i | i \in I)$ be a *family* of elements of X. Here I can be any *index set* and we allow for the possibility that $x_i = x_j$ for distinct $i, j \in I$.

Definition 1.14: By a *linear combination* of $(x_i | i \in I)$ we mean an expression of the form

$$\sum_{i \in J} \alpha_i x_i,$$

where J is a finite subset of I and $\alpha_i \in \mathbb{K}$ for all $i \in J$.

Observe that in the definition of linear combination, only *finite* sums are allowed.

Definition 1.15: The set of all linear combinations of $(x_i | i \in I)$ is called the *span* of $(x_i | i \in I)$ and is denoted $\operatorname{span}(x_i | i \in I)$.

Definition 1.16: The family $(x_i | i \in I)$ is said to be *linearly independent* provided that for every finite set $J \subset I$ and for every family $(\alpha_i | i \in J)$ of elements of $\mathbb{K}$ the condition

$$\sum_{i \in J} \alpha_i x_i = 0$$

implies that $\alpha_i = 0$ for all $i \in J$.

In other words, to say that a family is linearly independent means that no nontrivial linear combination can be equal to zero.

Remark 1.17: If $(x_i | i \in I)$ is linearly independent then it is injective, i.e., if $i, j \in I$ with $i \neq j$ then $x_i \neq x_j$.

Definition 1.18: The family $(x_i | i \in I)$ is said to be *linearly dependent* if it is not linearly independent.

Clearly, then, a family is linearly dependent if and only if there is a nontrivial linear combination that is equal to zero.

Definition 1.19: The family $(x_i|i \in I)$ is said to be a *Hamel basis* for X provided that $(x_i|i \in I)$ is linearly independent and $span(x_i|i \in I) = X$.

We note that a family $(x_i|i \in I)$ is a Hamel basis if and only if every element of X can be expressed as a linear combination of members of the basis in precisely one way.

Remark 1.20: The notions of linear combination, span, linear independence, linear dependence, and Hamel basis apply to subsets $S \subset X$ as well as families by regarding a set S as a family $(x|x \in S)$ through self-indexing.

Using Zorn's lemma, it is possible to show that every linear space has a Hamel basis. We give a slightly more useful result.

Proposition 1.21: Let S be a linearly independent subset of X. Then there is a Hamel basis B for X such that $S \subset B$.

Proof: Let $\mathcal{S}$ denote the collection of all linearly independent sets $A \subset X$ such that $S \subset A$, partially ordered by set inclusion. Let $\mathcal{C}$ be a chain (i.e., a totally ordered subset of $\mathcal{S}$). Then $\cup\mathcal{C}$ is an upper bound for $\mathcal{C}$. (See Problem 1 of Assignment 2.) It follows from Zorn's lemma that $\mathcal{S}$ has a maximal element B. To see that $\text{span}(B) = X$, suppose that $\text{span}(B)$ is a proper subset of X. Then we may choose $x \in X\backslash\text{span}(B)$. Since $x \notin \text{span}(B)$, it follows that $B \cup \{x\}$ is linear independent and this contradicts the maximality of B. $\square$

We shall have an extended discussion of Zorn's lemma when we prove the Hahn-Banach theorem.

Since $\emptyset$ is linearly independent, we have the following corollary.

Corollary 1.22: Every linear space over $\mathbb{K}$ has a Hamel basis.

Remark 1.23: It can be shown that every Hamel basis for a given linear space has the same cardinality. (See, for example, Goffman & Pederick.)

Definition 1.24: We say that X is *finitely generated* if there is a finite set $S \subset X$ such that $\text{span}(S) = X$.

Proposition 1.25: Assume that X is finitely generated. Then all Hamel bases for X are finite and have the same number of elements. (See, for example, *Linear Algebra* by Charles Curtis.)

Definition 1.26: If X is finitely generated, we say that X is *finite dimensional* and the number of elements in a Hamel basis for X is called the *dimension* of X. If X is not finitely generated, we say that X is *infinite dimensional.*

The following simple result is an immediate consequence of the definition of Hamel basis.

Proposition 1.27 : Let X be a linear space over $\mathbb{K}$ and $(x_i|i \in I)$ be a Hamel basis for X. Then, for every $i \in I$, there is a linear mapping $\alpha_i : X \to \mathbb{K}$ such that for every $x \in X$, (i) and (ii) below hold:

(i) $\{i \in I : \alpha_i(x) \neq 0\}$ is finite.

(ii) $x = \sum_{i \in I} \alpha_i(x) x_i$.

Definition 1.28: We refer to $(\alpha_i|i \in I)$ as the *family of coefficient mappings* for the basis $(x_i|i \in I)$.

Remark 1.29: It is interesting to note that no infinite-dimensional Banach space can have a countable Hamel basis.

To see why the remark is true, assume that $(X, \|\cdot\|)$ is an infinite-dimensional Banach space and suppose that $(x_i|i \in \mathbb{N})$ is a Hamel basis for X. For each $n \in \mathbb{N}$ let $V_n = \text{span}(x_1, x_2, \cdots, x_n)$. It is straightforward to show that $\text{int}(V_n) = \emptyset$. We shall see shortly that every finite-dimensional linear manifold is closed. It follows that each V_n is nowhere dense. On the other hand, we must have

$$X = \bigcup_{n=1}^{\infty} V_n,$$

which is impossible since X is complete. (Such a representation would violate the Baire Category Theorem.)

Remark 1.30: The preceding remark shows that there is no norm under which the space of all real sequences having only finitely many nonzero terms can be complete. (Indeed, the list $(e^{(i)}|i \in \mathbb{N})$, where $e_i^{(i)} = 1$ and $e_k^{(i)} = 0$ when $i \neq k$ is a Hamel basis for this space.) Similarly, there is no norm under which the space of all real polynomials can be complete.

Let $(X, \|\cdot\|)$ be a NLS and $\{x_n\}_{n=1}^{\infty}$ be a sequence in X. We associate with $\{x_n\}_{n=1}^{\infty}$ the sequence $\{s_n\}_{n=1}^{\infty}$ of partial sums given by

$$s_n = \sum_{k=1}^{n} x_k \text{ for all } n \in \mathbb{N}.$$

Definition 1.31: We say that the sequence $\{x_n\}_{n=1}^{\infty}$ is *summable*, or equivalently, that the series $\displaystyle\sum_{n=1}^{\infty} x_n$ is *convergent* provided that the sequence $\{s_n\}_{n=1}^{\infty}$ of partial sums is convergent. In this case we write

$$\sum_{n=1}^{\infty} x_n = \lim_{n\to\infty} s_n.$$

Definition 1.32: We say that the sequence $\{x_n\}_{n=1}^{\infty}$ is *absolutely summable*, or equivalently that the series $\sum_{n=1}^{\infty} x_n$ is *absolutely convergent* provided that the series $\sum_{n=1}^{\infty} \|x_n\|$ is convergent to a real number.

A very useful characterization of completeness is contained in the following simple proposition, whose proof is part of Assignment 2.

Prop. 1.33: A normed linear space $(X, \|\cdot\|)$ is complete if and only if every absolutely summable sequence is summable.

Hamel bases are of limited use in the study of infinite-dimensional NLS. No infinite-dimensional Banach space has a countable Hamel basis. Moreover, since the notion of a Hamel basis is a purely algebraic concept, the coefficients in the representation for a given vector x cannot be controlled in terms of $\|x\|$. There is another type of basis, known as a Schauder basis, that is very useful for certain types of computations in infinite-dimensional NLS.

Definition 1.34: A sequence $\{x_n\}_{n=1}^{\infty}$ is called a *Schauder basis* for $(X, \|\cdot\|)$ provided that for every $y \in X$ there is exactly one sequence $\{\alpha_n\}_{n=1}^{\infty}$ in $\mathbb{K}$ such that

$$\sum_{n=1}^{\infty} \alpha_n x_n = y.$$

Proposition 1.35: If $(X, \|\cdot\|)$ has a Schauder basis then $(X, \|\cdot\|)$ is separable.

The proof of Proposition 1.35 is also part of Assignment 2. Notice that a NLS that has a Schauder basis can (in some sense) be identified with a sequence space.

Completeness and Compactness in Finite-Dimensional NLS

The next two propositions express important and well-known properties of finite-dimensional NLS. I assume that students are already familiar with these results.

Prop. 1.36: Every finite-dimensional NLS is complete.

Remark 1.37: It follows from Proposition 1.36 that every finite-dimensional linear manifold in a NLS is closed. It is customary to refer to finite-dimensional linear manifolds as finite-dimensional subspaces. (Since they are automatically closed, there is no danger of ambiguity with the use of the term subspace.)

Another very important property of finite-dimensional NLS is that closed bounded sets are compact.

Proposition 1.38 (Heine-Borel Theorem): Let $(X, \|\cdot\|)$ be a finite-dimensional normed linear space, and $S \subset X$. Then S is compact if and only if S is both closed and bounded.

Proposition 1.38 has numerous important consequences. We shall soon show that in infinite-dimensional NLS, there are always closed bounded sets that fail to be compact. This lack of compactness in infinite dimensions is perhaps the most significant difference between finite-dimensional and infinite-dimensional NLS.

We complete this section with a lemma that can be used to prove proposition 1.36 and the difficult direction of proposition 1.38.

Lemma 1.39: Let $(X, \|\cdot\|)$ be a normed linear space and let $N \in \mathbb{N}$ be given. Assume that $(x_i | i = 1, 2, \cdots, N)$ is a linearly independent list of elements of X. Then there exists a constant $c > 0$ (depending on the list of vectors) such that

$$\|\sum_{i=1}^{N} \alpha_i x_i\| \geq c \sum_{i=1}^{N} |\alpha_i| \text{ for all } \alpha_1, \alpha_2, \cdots, \alpha_N \in \mathbb{K}.$$

If you have not seen this result before, I suggest that you work out a proof as an exercise for yourself. (Suggestion: If no such c exists then you can construct a sequence $\{\alpha^{(n)}\}_{n=1}^{\infty}$ in $\mathbb{K}^N$ such that

$$\sum_{i=1}^{N} |\alpha_i^{(n)}| = 1 \text{ for all } n \in \mathbb{N}$$

and

$$\|\sum_{i=1}^{N} \alpha_i^{(n)} x_i\| \to 0 \text{ as } n \to \infty.$$

This should lead to a contradiction.)

Lecture Notes for Week 2

Linear Spaces and Norms (Continued)

Lemma 2.1 (Riesz): Let $(X, \|\cdot\|)$ be a NLS and let Y, Z be linear manifolds in X such that Y is closed, $Y \subset Z$, and $Y \neq Z$. Let $\theta \in (0,1)$ be given. Then there exists $z \in Z$ such that $\|z\| = 1$ and $\|z - y\| \geq \theta$ for all $y \in Y$.

Proof: Choose $v \in Z \backslash Y$. Put

$$m = \inf\{\|y - v\| : y \in Y\},$$

and notice that $m > 0$ since Y is closed. By the definition of m, and since $\theta < 1$, we may choose $\hat{y} \in Y$ such that

$$m \leq \|\hat{y} - v\| \leq \frac{m}{\theta}.$$

Put $c = \|\hat{y} - v\|^{-1}$, $z = c(\hat{y} - v)$, and notice that $\|z\| = 1$. Let $y \in Y$ be given. Since $\hat{y} - c^{-1}y \in Y$ we have

$$\begin{aligned} \|z - y\| &= \|c(\hat{y} - v) - y\| \\ &= c\|\hat{y} - c^{-1}y - v\| \\ &= c\|v - (\hat{y} + c^{-1}y)\| \\ &\geq cm = \frac{m}{\|\hat{y} - v\|} \geq \frac{m}{m/\theta} = \theta. \quad \square \end{aligned}$$

Proposition 2.2: Let $(X, \|\cdot\|)$ be a NLS and put $B = \{x \in X : \|x\| \leq 1\}$. If B is compact then X is finite dimensional.

Proof: We prove the contrapositive implication. Assume that X is infinite dimensional. We shall use induction to construct a sequence $\{x_n\}_{n=1}^{\infty}$ such that $\|x_n\| = 1$ for all $n \in \mathbb{N}$ and $\|x_m - x_n\| \geq \frac{1}{2}$ for all $m, n \in \mathbb{N}$ with $m > n$. The existence of such a sequence implies that B is not compact, because each $x_n \in B$, but there can be no subsequence that is a Cauchy sequence. Choose any $x_1 \in X$ with $\|x_1\| = 1$. Let $n \in \mathbb{N}$ with $n \geq 2$ be given and assume that we have constructed $x_1, x_2, \cdots, x_n \in X$ with $\|x_i\| = 1$ for $i = 1, 2, \cdots, n$ such that

$$\|x_n - y\| \geq \frac{1}{2} \text{ for all } y \in \operatorname{span}(x_1, x_2, \cdots, x_{n-1}).$$

Since $\mathrm{span}(x_1, x_2, \cdots, x_n)$ is finite dimensional, it is a closed subspace. By Riesz's lemma, we may choose $x_{n+1} \in X$ with $\|x_{n+1}\| = 1$ such that

$$\|x_{n+1} - y\| \geq \frac{1}{2} \text{ for all } y \in \mathrm{span}(x_1, x_2, \cdots, x_n).$$

By induction, this procedure produces a sequence with the desired properties. $\square$

Remark 2.3: It is an immediate consequence of Proposition 2.2 that if S is a compact set in an infinite-dimensional NLS then $\mathrm{int}(S) = \emptyset$ and consequently (since S is closed) we can conclude that S is nowhere dense.

Remark 2.4: The proof of Proposition 2.2 also shows that if $(X, \|\cdot\|)$ is infinite dimensional, then the set $S = \{x \in X : \|x\| = 1\}$ is not compact.

Linear Mappings

Let X, Y, Z be linear spaces over the same field which we assume is either $\mathbb{R}$ or $\mathbb{C}$. Recall that a mapping $T : X \to Y$ is said to be *linear* provided that

$$T(\alpha x + \beta y) = \alpha T(x) + \beta T(y) \text{ for all } \alpha, \beta \in \mathbb{K},\ x, y \in X.$$

For a linear mapping T it is traditional to write Tx in place of $T(x)$. Linear mappings are frequently referred to as *linear operators*. When we talk about linear mappings between two different linear spaces it should be understood that the fields for the linear spaces are the same.

Definition 2.5: The *null space* of a linear mapping $T : X \to Y$ is defined by

$$\mathcal{N}(T) = \{x \in X : Tx = 0\}.$$

The term *kernel* is frequently used as a synonym for null space.

Definition 2.6: The *range* of a linear mapping $T : X \to Y$ is defined by

$$\mathcal{R}(T) = \{Tx : x \in X\}.$$

Some Basic Facts about Linear Mappings

We record some simple facts about linear mappings that are assumed to be familiar. The proofs of these facts for general linear spaces are the same as typically given in Linear Algebra for finite-dimensional spaces.

Let X, Y, and Z be linear spaces over the same field and assume that $T : X \to Y$ and $L : Y \to Z$ are linear mappings. Then we have

(a) $\mathcal{N}(T)$ is a linear manifold in X.

(b) $\mathcal{R}(T)$ is a linear manifold Y.

(c) T is injective if and only if $\mathcal{N}(T) = \{0\}$.

(d) If T is bijective then $T^{-1} : Y \to X$ is linear.

(e) If X is finite dimensional then $\mathcal{R}(T)$ is finite dimensional and $\dim(\mathcal{R}) \leq \dim(X)$. (In fact, if X is finite dimensional then $\dim(\mathcal{R}(T)) + \dim(\mathcal{N}(T)) = \dim(X)$.

(f) The mapping $LT : X \to Z$ defined by $LT(x) = L(T(x))$ for all $x \in X$ is linear.

(g) If T and L are bijective then LT is bijective and $(LT)^{-1} = T^{-1}L^{-1}$.

Definition 2.7: Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be NLS and $T : X \to Y$ be linear. We say that T is *bounded* provided that there is a constant M such that

$$\|Tx\|_Y \leq M\|x\|_X \text{ for all } x \in X.$$

Convention: In dealing with linear operators between different NLS it is customary to drop the subscripts on the norms and simply write $\|Tx\|$ and $\|x\|$ when there is no danger of confusion.

Prop 2.8: Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be NLS and $T : X \to Y$ be linear. The following four statements are equivalent.

(i) There exists a point $x_0 \in X$ such T is continuous at x_0.

(ii) T is continuous (on X).

(iii) T is uniformly continuous (on X).

(iv) T is bounded.

Proof: The implications (iv) $\Rightarrow$ (iii) $\Rightarrow$ (ii) $\Rightarrow$ (i) are immediate, so it suffices to prove (i) $\Rightarrow$ (iv). Let $x_0 \in X$ be given and assume that T is continuous at x_0. We may choose $\delta > 0$ such that $\|Tx - Tx_0\| < 1$ for all $x \in X$ with $\|x - x_0\| < \delta$. Let $y \in X\backslash\{0\}$ be given and put

$$x = x_0 + \frac{\delta}{2\|y\|}y.$$

Since $\|x - x_0\| < \delta$, we have $\|Tx - Tx_0\| < 1$. On the other hand, we have

$$\|Tx - Tx_0\| = \frac{\delta}{2\|y\|}\|Ty\|.$$

We conclude that

$$\|Ty\| \leq \frac{2}{\delta}\|y\|. \quad \square$$

Remark 2.9: The scalar field $\mathbb{K}$ is a (one-dimensional) linear space over itself. It therefore makes sense to talk about linear mappings from a linear space X to $\mathbb{K}$.

Proposition 2.10: Let $(X, \|\cdot\|)$ be a NLS and assume that $L : X \to \mathbb{K}$ is linear. Then L is continuous if and only if $\mathcal{N}(L)$ is closed.

The proof of Proposition 2.10 is Problem 12 on Assignment 2.

Corollary 2.11: Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be NLS and $T : X \to Y$ be a linear operator. Assume that X is finite dimensional. Then T is continuous.

Proof: Choose a (Hamel) basis $(x_i | i = 1, \cdots, N)$ for X and let $(\alpha_i | i = 1, \cdots, N)$ be the family of coefficient mappings for this basis. (See Definition 1.28.) For every $i = 1, \cdots, N$ we have that $\mathcal{N}(\alpha_i)$ is a finite-dimensional linear manifold and is therefore closed by Remark 1.37. It follows from Corollary 2.11 that α_i is continuous for each $i = 1, \cdots, N$. By the linearity of T, we have

$$Tx = \sum_{i=1}^{N} \alpha_i(x) T x_i.$$

Since each α_i is continuous, we conclude that T is continuous. $\square$

Equivalence of Norms

In many situations the given norm on a linear space is inconvenient for certain purposes. This leads to a natural question: *Can we replace the norm with another one that is more convenient and is such that the space retains its important properties when we change to the new norm?*

Definition 2.12: Let X be a linear space over $\mathbb{K}$ and $\|\cdot\|_a$ and $\|\cdot\|_b$ be norms on X. We say that these norms are equivalent provided that there are constants $M, m > 0$ such that

$$m\|x\|_a \le \|x\|_b \le M\|x\|_a \ \text{ for all } x \in X.$$

Remark 2.13: Notice that if the equation above holds then

$$M^{-1}\|x\|_b \le \|x\|_a \le m^{-1}\|x\|_b \ \text{ for all } x \in X,$$

which shows that Definition 2.12 is symmetric in $\|\cdot\|_a$ and $\|\cdot\|_b$.

Using Proposition 2.8 (and the fact that the identity mapping $I : X \to X$ is linear) we obtain the following two remarks.

Remark 2.14: Let X be a linear space with two norms $\|\cdot\|_a$ and $\|\cdot\|_b$. Then these norms are equivalent if and only if the identity operator is continuous from $(X, \|\cdot\|_a)$ to $(X, \|\cdot\|_b)$ and from $(X, \|\cdot\|_b)$ to $(X, \|\cdot\|_a)$.

Remark 2.15: If a linear space is equipped with two equivalent norms, then the associated metric spaces are uniformly homeomorphic.

Proposition 2.16: Let X be a finite-dimensional linear space over $\mathbb{K}$. Then all norms on X are equivalent.

Proposition 2.16 follows immediately from Corollary 2.11 and Remark 2.14. It can also be established using Lemma 1.38.

Spaces of Bounded Linear Mappings

Definition 2.17: Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be NLS over the same field $\mathbb{K}$. We denote the set of all bounded linear mappings from X to Y by $\mathcal{L}(X;Y)$. We define the function $\|\cdot\|_{\mathcal{L}(X;Y)} : \mathcal{L}(X;Y) \to \mathbb{R}$ by

$$\|T\|_{\mathcal{L}(X;Y)} = \sup\{\|Tx\|_Y : x \in X,\ \|x\|_X \leq 1\}.$$

Remark 2.18: It is clear that $\mathcal{L}(X;Y)$ is a linear space.

The following result is an easy consequence of the definition of $\|\cdot\|_{\mathcal{L}(X;Y)}$.

Proposition 2.19: Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be NLS. Then $\|\cdot\|_{\mathcal{L}(X;Y)}$ is a norm on $\mathcal{L}(X, Y)$.

Remark 2.20: When there is no danger of ambiguity, it is customary to omit subscripts on the norms for X, Y and $\mathcal{L}(X;Y)$.

Remark 2.21: If $X \neq \{0\}$ and $T \in \mathcal{L}(X;Y)$ then

$$\|T\| = \sup\left\{\frac{\|Tx\|}{\|x\|} : x \in X,\ x \neq 0\right\} = \sup\{\|Tx\| : x \in X,\ \|x\| = 1\}.$$

Remark 2.22: If $(Z, \|\cdot\|_Z)$ is a third NLS over the same field as X and Y and $T \in \mathcal{L}(X;Y),\ L \in \mathcal{L}(Y;Z)$ then $LT \in \mathcal{L}(X;Z)$ and $\|LT\| \leq \|L\| \cdot \|T\|$.

Proposition 2.23: Assume that $(Y, \|\cdot\|)$ is complete. Then $\mathcal{L}(X;Y)$ is complete under $\|\cdot\|_{\mathcal{L}(X;Y)}$.

Proof: Let $\{T_n\}_{n=1}^\infty$ be a Cauchy sequence in $\mathcal{L}(X;Y)$. Let $x \in X$ be given. Then $\{T_n x\}_{n=1}^\infty$ is a Cauchy sequence in Y; since Y is complete this sequence is convergent. Define $L : X \to Y$ by

$$Lx = \lim_{n\to\infty} T_n x.$$

It is clear that L is linear. We need to show that L is bounded and that $\|T_n - L\| \to 0$ as $n \to \infty$. Let $\epsilon > 0$ be given and choose $N \in \mathbb{N}$ such that $\|T_n - T_m\| < \frac{\epsilon}{2}$ for all

$m, n \geq N$. Let $x \in X$ be given. Then for all $n \geq N$ we have

$$\|T_n x\| \leq \|T_N x\| + \|(T_n - T_N)x\| \leq \left(\|T_N\| + \frac{\epsilon}{2}\right) \|x\|.$$

Letting $n \to \infty$ and using continuity of the norm we obtain

$$\|Lx\| \leq \left(\|T_N\| + \frac{\epsilon}{2}\right) \|x\| \quad \text{for all } x \in X,$$

so that L is bounded.

To show that $\|T_n - L\| \to 0$ as $n \to \infty$, let $x \in X$ and $n \geq N$ be given. Then we have

$$\begin{aligned} \|T_n x - Lx\| &= \lim_{m \to \infty} \|T_n x - T_m x\| \\ &\leq \limsup_{m \to \infty} \|T_n - T_m\| \cdot \|x\| \\ &\leq \frac{\epsilon}{2} \|x\|. \end{aligned}$$

Taking the supremum over all $x \in X$ with $\|x\| \leq 1$, we obtain

$$\|T_n - L\| \leq \frac{\epsilon}{2} < \epsilon \quad \text{for all } n \geq N. \quad \square$$

Lecture Notes for Week 3

The Principle of Uniform Boundedness

In practice, pointwise bounds are generally much easier to obtain than uniform bounds, but uniform bounds are much more useful. The *Banach-Steinhaus Theorem* (also know as the *Principle of Uniform Boundedness*) tells us that in certain important situations, pointwise bounds for a family bounded linear operators actually imply uniform bounds on the family of operator norms. Dunford & Schwartz list the Principle of Uniform Boundedness as one of *Three Basic Principles of Linear Analysis* (the other two being the *Hahn-Banach Theorem* and the *Open Mapping Theorem*).

Theorem 3.1 (Banach-Steinhaus): Let X be a Banach space and Y be a NLS. Let $(T_\alpha | \alpha \in I)$ be a family in $\mathcal{L}(X;Y)$. (Here, I can be any index set.) Assume that for every $x \in X$, we have

$$\sup\{\|T_\alpha x\| : \alpha \in I\} < \infty.$$

Then we also have

$$\sup\{\|T_\alpha\| : \alpha \in I\} < \infty.$$

Proof: For each $k \in \mathbb{N}$ let

$$A_k = \{x \in X : \|T_\alpha x\| \leq k \text{ for all } \alpha \in I\}.$$

Each A_k is a closed set because it is the intersection over α of the closed sets $\{x \in X : \|T_\alpha x\| \leq k\}$. The hypotheses of the theorem imply that

$$X = \bigcup_{k=1}^{\infty} A_k.$$

Since X is complete, the Baire Category theorem allows us to choose $N \in \mathbb{N}$ such that

$$\text{int}(A_N) \neq \emptyset.$$

Therefore, we may choose $x_0 \in A_N$, $\delta > 0$ such that $B_\delta(x_0) \subset A_N$. Then we have

$$\|T_\alpha x\| \leq N \text{ for all } \alpha \in I,\ x \in B_\delta(x_0).$$

Let $y \in X \backslash \{0\}$ be given and put

$$\gamma = \frac{\delta}{2\|y\|}, \quad z = x_0 + \gamma y.$$

Observe that $z \in B_\delta(x_0)$ so that

$$\|T_\alpha z\| \leq N \text{ for all } \alpha \in I.$$

Since $y = \gamma^{-1}(z - x_0)$ we have

$$T_\alpha y = \frac{1}{\gamma}(T_n z - T_n x_0) \quad \text{for all } \alpha \in I.$$

We conclude that for all $\alpha \in I$,

$$\|T_\alpha y\| \leq \frac{1}{\gamma}(\|T_\alpha z\| + \|T_\alpha x_0\|) \leq \frac{2}{\gamma}N = \frac{4N\|y\|}{\delta}.$$

Taking the supremum over all $y \in X$ with $\|y\| \leq 1$, we arrive at

$$\|T_\alpha\| \leq \frac{4N}{\delta} \quad \text{for all } \alpha \in \mathbb{N}. \;\; \square$$

Remark 3.2: The Banach-Steinhaus Theorem is also called the *Principle of Uniform Boundedness*.

Completeness of X is essential in the Banach-Steinhaus Theorem, as the following example illustrates.

Example 3.3: Let X be the set of all $\mathbb{K}$-valued sequence having finite support (i.e., the set of all sequences $x : \mathbb{N} \to \mathbb{K}$ such that $\{k \in \mathbb{N} : x_k \neq 0\}$ is finite), equipped with the norm $\|\cdot\|$ defined by

$$\|x\| = \sup\{|x_k| : k \in \mathbb{N}\},$$

and let $Y = \mathbb{K}$ equipped with $|\cdot|$. Notice that Y is complete, but X is not.

For each $n \in \mathbb{N}$ define the linear mapping $T_n : X \to Y$ by

$$T_n x = n x_n \quad \text{for all } x \in X.$$

Observe that

$$|T_n x| \leq n\|x\| \quad \text{for all } x \in X, \; n \in \mathbb{N},$$

so each $T_n \in \mathcal{L}(X, Y)$. For every $x \in X$, the sequence $(n x_n : n \in \mathbb{N})$ converges to 0 (since it has finite support), and consequently

$$\sup\{|T_n x| : n \in \mathbb{N}\} < \infty.$$

Consider the sequence $\{x^{(n)}\}_{n=1}^\infty$ in X defined by

$$x_k^{(n)} = \begin{cases} 1 & \text{if } k = n \\ 0 & \text{if } k \neq n. \end{cases}$$

Notice that $\|x^{(n)}\| = 1$ and $|T_n x^{(n)}| = n$ for all $n \in \mathbb{N}$. It follows that $\|T_n\| \geq n$ for all $n \in \mathbb{N}$ and consequently

$$\sup\{\|T_n\| : n \in \mathbb{N}\} = \infty. \;\; \square$$

Let X be a linear space over $\mathbb{K}$. Since $\mathbb{K}$ is a linear space over itself, it is possible to study linear mappings from X to $\mathbb{K}$. Such mappings are called *linear functionals*.

Definition 3.4: The set of all linear functionals on X is called the *algebraic dual* of X and is denoted by $X^{\#}$.

Remark 3.5: It is clear that $X^{\#}$ is a linear space over $\mathbb{K}$. Other notations are frequently used for algebraic duals.

The algebraic dual of an infinite-dimensional normed linear space is generally "too large" to be useful in analysis. We will not use the algebraic dual very much in this course. We introduce a smaller dual space, known as the topological dual, or simply the dual space, which will be used extensively.

Definition 3.6: Let $(X, \|\cdot\|)$ be a normed linear space. The set of all bounded linear functionals on X is called the *dual* of X and is denoted by X^*. We equip X^* with the norm $\|\cdot\|_*$ defined by

$$\|x^*\|_* = \sup\{|x^*(x)| : x \in X,\ \|x\| \leq 1\}.$$

Remark 3.7: Since $(\mathbb{K}, |\cdot|)$ is complete, it follows from Proposition 2.23 that $(X^*, \|\cdot\|_*)$ is complete, even if $(X, \|\cdot\|)$ is not.

Remark 3.8: If there is any danger of confusion with the algebraic dual, we shall refer to X^* as the *topological dual* of X. The notation X' is frequently used to denote the topological dual of X.

Definition 3.9: The mapping $\langle\cdot,\cdot\rangle : X^* \times X \to \mathbb{K}$ defined by

$$\langle x^*, x\rangle = x^*(x) \text{ for all } x \in X, x^* \in X^*$$

is called the *duality pairing* for X.

For each $x \in X$ the mapping from X^* to $\mathbb{K}$ that carries x^* into $\langle x^*, x\rangle$ is linear in x^*. This observation leads to a canonical linear injection J of X into $X^{**} = (X^*)^*$. More precisely, we define $J : X \to X^{**}$ by

$$(J(x))(x^*) = \langle x^*, x\rangle \text{ for all } x \in X, x^* \in X^*.$$

The mapping J is called the *canonical embedding of X into X^{**}*.

Remark 3.10: In order to show that J is injective we need to know that given $x \in X\backslash\{0\}$, there exists $x^* \in X^*$ such that $\langle x^*, x\rangle \neq 0$. Although the existence of such functionals may seem obvious, the proof is not completely straightforward. The Hahn-Banach theorem (to be discussed later) ensures the existence of linear functionals having many useful properties. In particular, it is a consequence of the Hahn-Banach Theorem that $\|J(x)\|_{**} = \|x\|$ for all $x \in X$.

Definition 3.11: A normed linear space $(X, \|\cdot\|)$ is said to be *reflexive* if the canonical embedding of X into X^{**} is surjective.

Reflexivity is a very important and subtle concept. We shall explore it extensively later on in the course.

Sequence Spaces

We now give a systematic introduction to some standard spaces of sequences of elements of $\mathbb{K}$. The spaces are very useful for constructing interesting examples.

We denote by $\mathbb{K}^{\mathbb{N}}$ the set of all sequences $x : \mathbb{N} \to \mathbb{K}$. It is clear that $\mathbb{K}^{\mathbb{N}}$ is a linear space over $\mathbb{K}$. There is no standard norm for this space.

We denote by $\mathbb{K}^{(\mathbb{N})}$ the set of all members of $\mathbb{K}^{\mathbb{N}}$ having finite support, i.e. the set of all $x \in \mathbb{K}^{\mathbb{N}}$ such that $\{n \in \mathbb{N} : x_n \neq 0\}$ is finite. It is clear that $\mathbb{K}^{(\mathbb{N})}$ is a linear space over $\mathbb{K}$. For each $n \in \mathbb{N}$ we define $e^{(n)} \in \mathbb{K}^{(\mathbb{N})}$ by

$$e_k^{(n)} = \begin{cases} 1 & \text{if } k = n \\ \\ 0 & \text{if } k \neq n. \end{cases}$$

Remark 3.12: It is easy to see that $\{e^{(n)}\}_{n=1}^{\infty}$ is a Hamel basis for $\mathbb{K}^{(\mathbb{N})}$. It follows that there is no norm that will render $\mathbb{K}^{(\mathbb{N})}$ complete.

The space l^∞: We denote by l^∞ the set of all bounded members of $\mathbb{K}^{\mathbb{N}}$. It is clear that l^∞ is a linear space over $\mathbb{K}$. We equip l^∞ with the norm defined by

$$\|x\|_\infty = \sup\{|x_n| : n \in \mathbb{N}\}.$$

It is easy to check that $\|\cdot\|_\infty$ is a norm and it is straightforward to show that $(l^\infty, \|\cdot\|_\infty)$ is complete.

The space c: We denote by c the set of all convergent members of $\mathbb{K}^{\mathbb{N}}$ and we equip c with $\|\cdot\|_\infty$. It is clear that c is a linear space over $\mathbb{K}$. Notice that $c \subset l^\infty$. It is straightforward to show that $(c, \|\cdot\|_\infty)$ is complete. Consequently c is a closed subspace of $(l^\infty, \|\cdot\|_\infty)$.

The space c_0: We denote by c_0 the set of all $x \in \mathbb{K}^{\mathbb{N}}$ such that $x_n \to 0$ as $n \to \infty$. It is clear that c_0 is a linear space. Notice that $c_0 \subset c$. It is straightforward to show that $(c_0, \|\cdot\|_\infty)$ is complete. Consequently c_0 is a closed subspace of $(l^\infty, \|\cdot\|_\infty)$ and of $(c, \|\cdot\|_\infty)$.

The spaces l^p, $1 \le p < \infty$: For each $p \in [1,\infty)$ we denote by l^p the set of all $x \in \mathbb{K}^{\mathbb{N}}$ such that

$$\sum_{n=1}^{\infty} |x_n|^p < \infty,$$

and we define $\|\cdot\|_p : l^p \to \mathbb{R}$ by

$$\|x\|_p = \left(\sum_{n=1}^{\infty} |x_n|^p \right)^{1/p}.$$

It is clear that l^1 is a linear space over $\mathbb{K}$ and that $\|\cdot\|_1$ is a norm on l^1. It is not difficult to show that $(l^1, \|\cdot\|_1)$ is complete. It is true that l^p is a linear space over $\mathbb{K}$ and that $\|\cdot\|_p$ is a norm on l^p for each $p \in [1,\infty)$, but this is not completely obvious and a proof will be given shortly. These spaces are also complete. Notice that for $1 \le p_1 \le p_2 < \infty$ we have

$$l^{p_1} \subset l^{p_2} \subset c_0 \subset c \subset l^{\infty}.$$

Remark 3.13: Let $p \in [1,\infty]$ and $x \in l^p$ be given. It is straightforward to show that

$$\|x\|_p \ge \|x\|_\infty.$$

If we allow for the norms to take infinite values, then this inequality is valid for all $x \in \mathbb{K}^{\mathbb{N}}$. Moreover, if there exists $q \ge 1$ such $x \in l^p$ for all $p > q$ then

$$\|x\|_p \to \|x\|_\infty \quad \text{as } p \to \infty.$$

Before proving that l^p is in fact a linear space and that $\|\cdot\|$ is a norm, we record some important properties of the sequence spaces introduced above.

Remark 3.14: From now on, when we talk about the spaces c, c_0, and l^p, $1 \le p \le \infty$ it is to be understood that they are equipped with the norms described above.

Properties of the Sequence Spaces

Completeness: c, c_0, and l^p, $1 \le p \le \infty$ are complete.

Separability: c, c_0, and l^p, $1 \le p < \infty$ are separable. l^∞ is not separable.

Schauder Bases: The sequence $\{e^{(n)}\}_{n=1}^{\infty}$ is a Schauder basis for c_0 and for l^p, $1 \le p < \infty$. The space c has Schauder bases; you are asked to find one in Assignment 3.

Reflexivity: The spaces l^p, $1 < p < \infty$ are reflexive. The spaces c, c_0, l^1, and l^∞ are not reflexive.

Remark 3.15:

(a) Let X denote one of the spaces $c, c_0, l^p, 1 \le p \le \infty$. To prove that X is complete, one takes a Cauchy sequence $\{x^{(n)}\}_{n=1}^{\infty}$ in X and observes that for each $k \in \mathbb{N}$, $\{x_k^{(n)}\}_{n=1}^{\infty}$ is a Cauchy sequence in $\mathbb{K}$; since $\mathbb{K}$ is complete, one may choose $x_k \in \mathbb{K}$ such that $x_k^{(n)} \to x_k$ as $n \to \infty$. One then shows that the sequence $x \in \mathbb{K}^{\mathbb{N}}$ satisfies $x \in X$ and $\|x^{(n)} - x\|_X \to 0$ as $n \to \infty$.

(b) Separability of c, c_0, and l^p, $1 \le p < \infty$ follows from the existence of Schauder bases.

(c) To see that l^∞ is not separable, let $S = \{x \in l^\infty : x_n \in \{0,1\}$ for all $n \in \mathbb{N}\}$ and observe that S is uncountable. Observe further that $\|x - y\|_\infty = 1$ for all $x, y \in S$ with $x \neq y$. Consequently, each (open) ball of radius 1 contains at most one element of S and we conclude that no countable collection of open balls of radius 1 can cover l^∞. Proposition M8 implies that l^∞ is not separable.

In order to prove that the spaces l^p, $1 < p < \infty$ are linear spaces and that $\|\cdot\|_p$ is a norm on l^p, it suffices to prove that

$$\|x + y\|_p \le \|x\|_p + \|y\|_p \quad \text{for all } x, y \in l^p.$$

The other properties are immediate. The inequality above is called Minkowski's inequality. In order to establish this inequality, we start with an important algebraic inequality, known as Young's inequality, and then prove a fundamental inequality, called Holder's inequality, for l^p spaces. Holder's inequality will be used to establish Minkowski's inequality.

Proposition 3.16 (Young's Inequality): Let $p, q \in (1, \infty)$ with $\frac{1}{p} + \frac{1}{q} = 1$ and $a, b \ge 0$ be given. Then

$$ab \le \frac{a^p}{p} + \frac{b^q}{q}$$

(with equality if and only if $b = a^{p-1}$).

Remark 3.17: For each $p \in (1, \infty)$ there is exactly one $q \in (1, \infty)$ satisfying $\frac{1}{p} + \frac{1}{q} = 1$, namely

$$q = \frac{p}{p-1}.$$

Proof of Young's Inequality: We assume that $a, b > 0$. (If $a = 0$ or $b = 0$, there is nothing to prove.) Put $\gamma = p - 1$. In the $x - y$ plane, consider the rectangle with vertices $(0,0), (a,0), (a,b), (0,b)$ and sketch the graph of $y = x^\gamma$ for $0 \le x \le \max\{a, b^{1/\gamma}\}$. By looking at areas we see that

$$ab \le \int_0^a x^\gamma dx + \int_0^b y^{1/\gamma} dy,$$

(with equality if and only if $a = b^{1/\gamma}$.) Evaluating the integrals we obtain

$$ab \le \frac{a^{\gamma+1}}{\gamma+1} + \frac{b^{\frac{1}{\gamma}+1}}{\frac{1}{\gamma}+1}.$$

Substituting $p = \gamma + 1$ and $q = \frac{p}{p-1}$ gives the desired result. $\square$

Notice that for $p, q \in (1, \infty)$ with $p^{-1} + q^{-1} = 1$, Young's Inequality implies the existence of constants $C_p, K_q > 0$ such that

$$ab \le C_p a^p + K_q b^q \quad \text{for all } a, b \ge 0.$$

In particular the standard form of Young's Inequality uses $C_p = p^{-1}$ and $K_q = q^{-1}$. If either one of these constants is reduced, without increasing the other one, then the inequality above will fail. In applications it is sometimes essential to reduce one of the constants, say C_p, at the expense of increasing the other. There is a simple scaling argument to see how this works.

Remark 3.18: Given $\epsilon > 0$, $p, q \in (1, \infty)$ with $p^{-1} + q^{-1} = 1$, and $a, b \ge 0$, observe that

$$\begin{aligned} ab &= [(\epsilon p)^{\frac{1}{p}} a][(\epsilon p)^{-\frac{1}{p}} b] \\ &\le \epsilon a^p + \frac{b^q}{q(\epsilon p)^{\frac{q}{p}}}. \end{aligned}$$

Proposition 3.19 (Holder's Inequality): Let $p, q \in (1, \infty)$ with $\frac{1}{p} + \frac{1}{q} = 1$ and $x \in l^p, y \in l^q$ be given. Then

$$\sum_{i=1}^{\infty} |x_i y_i| \le \|x\|_p \|y\|_q.$$

Proof: We may assume that $x \neq 0$ and $y \neq 0$. Let us put

$$u = \frac{x}{\|x\|_p}, \quad v = \frac{y}{\|y\|_q}.$$

Notice that $u \in l^p$, $v \in l^q$, and $\|u\|_p = \|v\|_q = 1$, from which we conclude that

$$\frac{\|u\|_p^p}{p} + \frac{\|v\|_q^q}{q} = 1.$$

It suffices to show that

$$\sum_{i=1}^{\infty} |u_i v_i| \le 1.$$

By Young's Inequality, we have

$$|u_i v_i| \le \frac{|u_i|^p}{p} + \frac{|v_i|^q}{q} \quad \text{for all } i \in \mathbb{N}.$$

Let $N \in \mathbb{N}$ be given. Then we have

$$\sum_{i=1}^{N} |u_i v_i| \leq \frac{1}{p} \sum_{i=1}^{N} |u_i|^p + \frac{1}{q} \sum_{i=1}^{N} |v_i|^q \leq \frac{\|u\|_p^p}{p} + \frac{\|v\|_q^q}{q} = 1.$$

Letting $N \to \infty$, we conclude that

$$\sum_{i=1}^{\infty} |u_i v_i| \leq 1 \quad \square.$$

Remark 3.20: Holder's inequality remains valid for $p = 1$ and $q = \infty$. The proof is immediate.

Proposition 3.21 (Minkowski's Inequality): Let $p \in [1, \infty]$ be given. Then

$$\|x + y\|_p \leq \|x\|_p + \|y\|_p \text{ for all } x, y \in l^p.$$

Proof: The cases $p = 1$ and $p = \infty$ are immediate, so we assume that $p \in (1, \infty)$. Also, if $x + y = 0$ there is nothing to prove, so we assume that $x + y \neq 0$. Let $N \in \mathbb{N}$ be given, sufficiently large so that

$$\sum_{i=1}^{N} |x_i + y_i|^p > 0.$$

Choose $q \in (1, \infty)$ such that $\frac{1}{p} + \frac{1}{q} = 1$. Then we have

$$\begin{aligned}
\sum_{i=1}^{N} |x_i + y_i|^p &= \sum_{i=1}^{N} |x_i + y_i| \cdot |x_i + y_i|^{p-1} \\
&\leq \sum_{i=1}^{N} |x_i||x_i + y_i|^{p-1} + \sum_{i=1}^{N} |y_i||x_i + y_i|^{p-1} \\
&\leq \left(\sum_{i=1}^{N} |x_i|^p \right)^{1/p} \left(\sum_{i=1}^{N} |x_i + y_i|^p \right)^{1/q} + \left(\sum_{i=1}^{N} |y_i|^p \right)^{1/p} \left(\sum_{i=1}^{N} |x_i + y_i|^p \right)^{1/q}.
\end{aligned}$$

Here we have used Holder's inequality and the fact that $(p-1)q = p$. Dividing both sides by

$$\left(\sum_{i=1}^{N} |x_i + y_i|^p \right)^{1/q},$$

we obtain

$$\left(\sum_{i=1}^{N} |x_i + y_i|^p\right)^{1-\frac{1}{q}} \leq \left(\sum_{i=1}^{N} |x_i|^p\right)^{1/p} + \left(\sum_{i=1}^{N} |y_i|^p\right)^{1/p}.$$

Since $1 - \frac{1}{q} = \frac{1}{p}$ we conclude that

$$\left(\sum_{i=1}^{N} |x_i + y_i|^p\right)^{1/p} \leq \left(\sum_{i=1}^{N} |x_i|^p\right)^{1/p} + \left(\sum_{i=1}^{N} |y_i|^p\right)^{1/p}.$$

Letting $N \to \infty$ we obtain the desired result. $\square$

The Open Mapping Theorem

The main tool used in the proof of the Principle of Uniform Boundedness was the Baire Category Theorem. Another very important consequence of the Baire Category Theorem is the so-called Open Mapping Theorem. This theorem says that if a bounded linear mapping between two Banach spaces is surjective, then it maps open sets to open sets.

Theorem 3.22 (Open Mapping Theorem): Let X, Y be Banach spaces, $T \in \mathcal{L}(X; Y)$ be given and assume that is surjective. Let $\mathcal{O}$ be an open subset of X. Then $\{Tx : x \in \mathcal{O}\}$ is an open subset of Y.

Before proving this theorem, we record a very useful (and immediate) corollary and make a few observations.

Corollary 3.23: (Bounded Inverse Theorem) Let X, Y be Banach spaces and assume that $T \in \mathcal{L}(X; Y)$ is bijective. Then T^{-1} is bounded.

By combining Remark 2.14 with Corollary 3.23, we obtain the following result.

Corollary 3.24: Let X be a linear space and let $\|\cdot\|_1$ and $\|\cdot\|_2$ be norms on X such that $(X, \|\cdot\|_1)$ and $(X, \|\cdot\|_2)$ both are complete. Assume that there exists $M \in \mathbb{R}$ such that

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

Then $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent.

In Corollary 3.24, it is essential that X be complete under both norms, as the following example shows.

Example 3.25: Let $\mathbb{K} = \mathbb{R}$ and $X = C([0,1]; \mathbb{R})$ (i.e., the set of continuous functions from $[0,1]$ to $\mathbb{R}$). Consider the norms on X given by

$$\|f\|_\infty = \max\{|f(x)| : x \in [0,1]\}, \quad \|f\|_1 = \int_0^1 |f(x)|\,dx, \quad \text{for all } f \in X.$$

It is easy to see that $(X, \|\cdot\|_\infty)$ is complete because this norm corresponds to uniform convergence. The space $(X, \|\cdot\|_1)$ is incomplete. Observe that

$$\|f\|_1 \leq \|f\|_\infty \ \text{ for all } f \in X.$$

However, it is easy to see that these two norms are not equivalent. Consider the sequence $\{f_n\}_{n=1}^\infty$ of continuous functions defined by

$$f_n(x) = \begin{cases} n - n^2 x & \text{for } 0 \leq x \leq \frac{1}{n} \\ \\ 0 & \text{for } \frac{1}{n} \leq x \leq 1, \end{cases}$$

and observe that $\|f_n\|_\infty = n$ and $\|f_n\|_1 = 1$ for all $n \in \mathbb{N}$.

Remark 3.26: Let $X = C([0,1];\mathbb{R}$ equipped with $\|\cdot\|_\infty$, $Y = C([0,1];\mathbb{R})$ equipped with $\|\cdot\|_1$, and define $T \in \mathcal{L}(X;Y)$ by $Tf = f$ for all $f \in X$. Clearly T is bijective. Example 3.25 shows that T does not map every open set in X to an open set in Y because T^{-1} is not continuous. This shows that completeness of the space Y is essential in the Open Mapping Theorem. You will be asked what happens to the Open Mapping Theorem in Y is complete but X is incomplete in Assignment 3.

It is easy to see that the conclusion of the Banach-Steinhaus can fail if T is not surjective – even if X and Y are both complete.

Example 3.27: Let X and Y be Banach spaces and assume that $Y \neq 0$. Define $T : X \to Y$ by $Tx = 0$ for all $x \in X$. Clearly T is linear and continuous, but $\{Tx : x \in X\} = \{0\}$ which is not open, even though X is open.

Lecture Notes for Week 4

The Open Mapping Theorem (Continued)

The proof of the Open Mapping Theorem will be broken down into several pieces. Some new notation will be helpful.

General Notation (to be used later as well): Let Z, W be linear spaces, $A \subset Z$, $a \in Z$, $\alpha \in \mathbb{K}$, and $T : Z \to W$ be given. Then

$$\alpha A = \{\alpha z : z \in Z\}, \quad a + A = \{a + z : z \in A\}, \quad T[A] = \{T(z) : z \in A\}.$$

Local Notation (to be used only in our discussion of the open mapping theorem): Given $\epsilon, \delta > 0$ we put

$$U_\epsilon = B_\epsilon^X(0), \quad V_\delta = B_\delta^Y(0).$$

Throughout this section $\overline{A}$ denotes the closure of a subset A of a NLS.

Lemma 4.1: Let X be a NLS, Y be a Banach space, and assume that $T \in \mathcal{L}(X;Y)$ is surjective. Let $\epsilon > 0$ be given. Then there exists $\delta > 0$ such that

$$\overline{T[U_{2\epsilon}]} \supset V_\delta.$$

Proof: Observe that

$$X = \bigcup_{n=1}^\infty nU_\epsilon.$$

Since T is surjective (and linear) we may write

$$Y = \bigcup_{n=1}^\infty nT[U_\epsilon].$$

Since Y is complete, we may invoke the Baire Category Theorem to choose $N \in \mathbb{N}$ such that

$$\mathrm{int}(N\overline{T[U_\epsilon]}) \neq \emptyset.$$

(Here we have made use of the simple fact that $\mathrm{cl}(NS) = N\mathrm{cl}(S)$ for any set $S \subset Y$.) Choose $z \in Y$ and $r > 0$ such that

$$B_r^Y(z) \subset N\overline{T[U_\epsilon]}.$$

Let us put

$$\delta = \frac{r}{N}, \quad y_0 = \frac{z}{N},$$

so that

$$B_\delta^Y(y_0) \subset \overline{T[U_\epsilon]}. \tag{1}$$

Define the sets $P \subset Y$ and $Q \subset X$ by

$$P = \{y_1 - y_2 : y_1, y_2 \in B_\delta^Y(y_0)\},$$

$$Q = \{u_1 - u_2 : u_1, u_2 \in U_\epsilon\}.$$

It follows from (1) and the definitions of P and Q that

$$P \subset \overline{T[Q]}. \tag{2}$$

To see why (2) is true, let $y \in P$ be given. Then we may choose $y_1, y_2 \in B_\delta^Y(y_0)$ and sequences $\{u_1^{(n)}\}_{n=1}^\infty$ and $\{u_2^{(n)}\}_{n=1}^\infty$ in U_ϵ such that

$$y = y_1 - y_2, \quad Tu_1^{(n)} \to y_1 \text{ and } Tu_2^{(n)} \to y_2 \text{ as } n \to \infty.$$

Then we have $u_1^{(n)} - u_2^{(n)} \in Q$ for all $n \in \mathbb{N}$ and

$$T(u_1^{(n)} - u_2^{(n)}) = Tu_1^{(n)} - Tu_2^{(n)} \to y_1 - y_2 2 = y \text{ as } n \to \infty.$$

Notice that $Q \subset U_{2\epsilon}$ by the triangle inequality and consequently

$$\overline{T[Q]} \subset \overline{T[U_{2\epsilon}]}. \tag{3}$$

Finally, we observe that $V_\delta \subset P$ since every $y \in V_\delta$ can be expressed in the form $y = (y + y_0) - y_0$. Using (2) and (3), we conclude that

$$V_\delta \subset P \subset \overline{T[U_{2\epsilon}]}. \ \square$$

Lemma 4.2: Let X and Y be Banach spaces and assume that $T \in \mathcal{L}(X;Y)$ is surjective. Let $\epsilon_0 > 0$ be given. Then there exists $\delta_0 > 0$ such that

$$T[U_{2\epsilon_0}] \supset V_{\delta_0}.$$

Proof: Choose a real sequence $\{\epsilon_n\}_{n=1}^{\infty}$ such that $\epsilon_n > 0$ for all $n \in \mathbb{N}$ and

$$\sum_{n=1}^{\infty} \epsilon_n < \epsilon_0. \tag{4}$$

Using Lemma 4.1, we may choose a real sequence $\{\delta_n\}_{n=0}^{\infty}$ such that $\delta_n > 0$ for all $n \in \mathbb{N} \cup \{0\}$,

$$\overline{T[U_{\epsilon_n}]} \supset V_{\delta_n} \quad \text{for all } n \in \mathbb{N} \cup \{0\}, \tag{5}$$

and $\delta_n \to 0$ as $n \to \infty$.

Let $y \in V_{\delta_0}$ be given. We need to produce $x \in U_{2\epsilon_0}$ such that $y = Tx$. Using (5) with $n = 0$ we may choose $x_0 \in U_{\epsilon_0}$ such $\|y - Tx_0\|$ is as small as we please. In particular, we may choose $x_0 \in U_{\epsilon_0}$ such that $\|y - Tx_0\| < \delta_1$. Now $y - Tx_0 \in V_{\delta_1}$. Using (5) with $n = 1$, we may choose $x_1 \in U_{\epsilon_1}$ such that $\|y - Tx_0 - Tx_1\| < \delta_2$. Now we have $y - Tx_0 - Tx_1 \in V_{\delta_2}$, so we may use (5) with $n = 2$ to choose $x_2 \in U_{\epsilon_2}$ with $\|y - Tx_0 - Tx_1 - Tx_2\| < \delta_3$, and so forth. By induction, we can construct a sequence $\{x_n\}_{n=0}^{\infty}$ such that

$$x_n \in U_{\epsilon_n}, \quad \|y - T\left(\sum_{i=0}^{n} x_i\right)\| < \delta_{n+1} \quad \text{for all } n \in \mathbb{N} \cup \{0\}. \tag{6}$$

Observe that

$$\sum_{i=0}^{n} \|x_i\| < \epsilon_0 + \epsilon_1 + \cdots + \epsilon_n < 2\epsilon_0,$$

and consequently the sequence $\{x_n\}_{n=0}^{\infty}$ is absolutely summable. Since X is complete, it follows from Proposition 1.33 that the sequence $\{x_n\}_{n=0}^{\infty}$ is summable. Put

$$x = \sum_{n=0}^{\infty} x_n,$$

and notice that $x \in U_{2\epsilon_0}$ by virtue of (4) and the fact that each $x_n \in U_{\epsilon_n}$. Since T is continuous and $\delta_n \to 0$ as $n \to \infty$, we deduce from (6) that $y = Tx$ and the proof is complete. $\square$

Proof of the Open Mapping Theorem: Let $y \in T[\mathcal{O}]$ be given. We need to find $\delta > 0$ such that $B_\delta^Y(y) \subset T[\mathcal{O}]$. To this end, we choose $x \in \mathcal{O}$ such that $y = Tx$. Since $\mathcal{O}$ is open, we may may choose $\epsilon > 0$ such that

$$B_\epsilon^X(x) = x + U_\epsilon \subset \mathcal{O}.$$

By Lemma 4.2, we may choose $\delta > 0$ such that

$$T[U_\epsilon] \supset V_\eta.$$

Then we have

$$T[\mathcal{O}] \supset T[x + U_\epsilon] = y + T[U_\epsilon] \supset y + V_\delta = B^Y_\delta(y). \quad \square$$

Closed Graph Theorem

A very important consequence of the open mapping theorem is the so-called *closed Graph Theorem*. Before stating this theorem, we need to talk a bit about *product spaces*.

Let X and Y be linear spaces over $\mathbb{K}$ and let

$$X \times Y = \{(x, y) : x \in X, y \in Y\}.$$

If we define addition and scalar multiplication componentwise, then $X \times Y$ is a linear space over $\mathbb{K}$. If X and Y are normed, then we can define a natural norm $\|(\cdot, \cdot)\|_{X\times Y}$ on $X \times Y$ by

$$\|(x, y)\|_{X\times Y} = \|x\|_X + \|y\|_Y \ \text{ for all } (x, y) \in X \times Y. \tag{7}$$

Remark 4.3: It is easy to see that $X \times Y$ is complete in the norm defined by (7) if and only if X and Y both are complete.

It is easy (and sometimes very useful) to construct equivalent norms on $X \times Y$.

Remark 4.4: Let ν be a norm on $\mathbb{R}^2$. Then the norm on $X \times Y$ defined by

$$|||(x, y)||| = \nu(\|x\|_X, \|y\|_Y)$$

is equivalent to $\|(\cdot, \cdot)\|_{X\times Y}$.

Let $T : X \to Y$ be a linear mapping. The *graph* of T is the linear manifold in $X \times Y$ defined by

$$\mathrm{Gr}(T) = \{(x, Tx) : x \in X\}.$$

If X and Y are normed and $T \in \mathcal{L}(X; Y)$ then $\mathrm{Gr}(T)$ is a closed subset of $X \times Y$. In general, closedness of the graph of a linear operator *does not* imply continuity of the operator; however, if X and Y are both complete, a closed graph implies continuity of a linear operator. This is the content of the next theorem.

Theorem 4.5 (Closed Graph Theorem): Let X and Y be Banach spaces and assume that $T : X \to Y$ is linear and that $\mathrm{Gr}(T)$ is a closed subset of $X \times Y$. Then T is continuous.

Proof: Define the linear mapping $L : \mathrm{Gr}(T) \to X$ by

$$L(x, Tx) = x \ \text{ for all } x \in X.$$

Observe that $(\mathrm{Gr}(T), \|(\cdot,\cdot)\|_{X\times Y})$ is a Banach space since $\mathrm{Gr}(T)$ is a closed subspace of $X \times Y$. Clearly L is bijective. Notice that L is continuous because

$$\|L(x, Tx)\| = \|x\| \leq \|x\| + \|Tx\| = \|(x, Tx)\| \text{ for all } x \in X.$$

By the bounded inverse theorem $L^{-1} : X \to \mathrm{Gr}(T)$ is continuous. Consequently we may choose $K \in \mathbb{R}$ such that

$$\|L^{-1}x\| \leq K\|x\| \text{ for all } x \in X.$$

Since $L^{-1}x = (x, Tx)$ for all $x \in X$ we deduce that

$$\|(x, Tx)\| = \|x\| + \|Tx\| \leq K\|x\| \text{ for all } x \in X.$$

It follows that

$$\|Tx\| \leq K\|x\| \text{ for all } x \in X. \quad \square$$

The Closed Graph Theorem is very convenient for a large number of situations. You will encounter several such situations in Assignment 3.

Hahn-Banach Theorems

There are several closely related results, each of which is sometimes referred to as “the” Hahn-Banach theorem. “Algebraic forms”, or “extension forms” of the Hahn-Banach theorem are concerned with extending linear functionals defined on a linear manifold in a linear space to be defined on the entire space in such a way that certain inequalities, or bounds, are preserved. These extension theorems imply very powerful results concerning the separation of convex sets by linear functionals (or by hyperplanes). The separation theorems are sometimes referred to as “geometric forms” or “separation forms” of the Hahn-Banach Theorem.

We shall begin by proving an extension theorem for real linear spaces. Then we prove an extension theorem for real or complex linear spaces and investigate important consequences of this theorem in normed linear spaces (real or complex). After discussing the extension theorems, we shall investigate the separation of convex sets.

Theorem 4.6 (Hahn-Banach Theorem (Real Linear Space)): Let X be a real linear space and assume that $p : X \to \mathbb{R}$ satisfies

(i) $\forall x, y \in X, \quad p(x + y) \leq p(x) + p(y)$,

(ii) $\forall x \in X, \alpha \geq 0, \quad p(\alpha x) = \alpha p(x)$.

Let Y be a linear manifold in X and $f : Y \to \mathbb{R}$ be a linear functional satisfying $f(x) \leq p(x)$ for all $x \in Y$. Then there exists a linear functional $F : X \to \mathbb{R}$ such that $F(x) = f(x)$ for all $x \in Y$ and $F(x) \leq p(x)$ for all $x \in X$.

Remark 4.7: It is very easy to produce a linear functional $F : X \to \mathbb{R}$ satisfying $F(x) = f(x)$ for all $x \in Y$. The deep part of the theorem is that the extension can be made so as to preserve the inequality $F(x) \leq p(x)$.

Proof of Theorem 4.6: Let $\mathcal{E}$ be the set of all linear functionals $g : \mathcal{D}(g) \to \mathbb{R}$ such that $\mathrm{Gr}(f) \subset \mathrm{Gr}(g)$ and $g(x) \leq p(x)$ for all $x \in \mathcal{D}(g)$. (Notice that the condition $\mathrm{Gr}(f) \subset \mathrm{Gr}(g)$ means precisely that g is an extension of f, i.e. $Y \subset \mathcal{D}(g)$ and $g(x) = f(x)$ for all $x \in Y$.) Let us equip $\mathcal{E}$ with the partial order $\leq$ defined by

$$g_1 \leq g_2 \;\Leftrightarrow\; \mathrm{Gr}(g_1) \subset \mathrm{Gr}(\mathrm{g}_2).$$

Let $\mathcal{C}$ be a chain in $(\mathcal{E}, \leq)$. Then

$$\bigcup_{g \in \mathcal{C}} \mathrm{Gr}(g)$$

is the graph of a linear functional $\tilde{g}$ and $\tilde{g}$ is an upper bound for $\mathcal{C}$. By Zorn's lemma, $(\mathcal{E}, \leq)$ has a maximal element F. We need to show that $\mathcal{D}(F) = X$.

Suppose that $\mathcal{D} \neq X$. Then we may choose $y_1 \in X \backslash \mathcal{D}(F)$. Notice that $y_1 \neq 0$ and put

$$Y_1 = \mathrm{span}(\mathcal{D}(F) \cup \{y_1\}).$$

Given $x \in Y_1$ there is a unique decomposition

$$x = y + \alpha y_1$$

with $y \in \mathcal{D}(F)$ and $\alpha \in \mathbb{R}$. (More precisely, for each $x \in Y_1$ there is exactly one pair $(y, \alpha) \in \mathcal{D}(F) \times \mathbb{R}$ such that $x = y + \alpha y_1$.) For each $c \in \mathbb{R}$ define the linear functional $g_c : Y_1 \to \mathbb{R}$ by

$$g_c(y + \alpha y_1) = F(y) + \alpha c.$$

We want to choose c so that

$$g_c(x) \leq p(x) \;\text{ for all } x \in \mathcal{D}(g_c) = Y_1.$$

(This will contradict the maximality of F.)

Let $w, z \in \mathcal{D}(F)$ be given. Then we have

$$\begin{aligned} F(w) - F(z) &= F(w - z) \\ &\leq p(w - z) = p(w + y_1 - y_1 - z) \\ &\leq p(w + y_1) + p(-y_1 - z) \end{aligned}$$

It follows that

$$-p(-y_1 - z) - F(z) \le p(w + y_1) - F(w) \quad \text{for all } w, z \in \mathcal{D}(F) \tag{8}$$

Put

$$m = \sup\{-p(-y_1 - z) - F(z) : z \in \mathcal{D}(F)\},$$

$$M = \inf\{p(w + y_1) - F(w) : w \in \mathcal{D}(F)\}.$$

It follows from (8) that $m \le M$. Choose c with $m \le c \le M$. Then we have

$$-p(-y_1 - z) - F(z) \le c \quad \text{for all } z \in \mathcal{D}(F), \tag{9}$$

$$c \le p(w + y_1) - F(w) \quad \text{for all } w \in \mathcal{D}(F). \tag{10}$$

Let $y \in \mathcal{D}(F)$ and $\alpha < 0$ be given. Putting $z = \alpha^{-1}y$ in (9) we obtain

$$-p(-y_1 - \alpha^{-1}y) - F(\alpha^{-1}y) \le c.$$

Multiplying by $-\alpha > 0$ and using (ii) we obtain

$$-p(\alpha y_1 + y) + F(y) \le -\alpha c.$$

We conclude that

$$g_c(y + \alpha y_1) = F(y) + \alpha c \le p(y + \alpha y_1).$$

Now let $\alpha > 0$ be given and put $w = \alpha^{-1}y$ in (10). This gives

$$c \le p(\alpha^{-1}y + y_1) - F(\alpha^{-1}y).$$

Multiplying through by α and using (ii) we obtain

$$\alpha c \le p(y + \alpha y_1) - F(y).$$

Once again, we have

$$g_c(y + \alpha y_1) = F(y) + \alpha c \leq p(y + \alpha y_1).$$

(For $\alpha = 0$ there is nothing to worry about – we already know that $g_c(y) = F(y) \leq p(y)$.)

This shows that g_c is a proper linear extension of F satisfying $g_c(x) \leq p(x)$ for all $x \in \mathcal{D}(g_c)$. This contradicts the maximality of F. We conclude that $\mathcal{D}(F) = X$. $\square$

Lecture Notes for Week 5

Hahn-Banach Theorems (Continued)

We now give a generalization of Theorem 4.6 that applies to both real and complex linear spaces.

Theorem 5.1 (Hahn-Banach Theorem (Real or Complex Linear Space)): Let X be a linear space over $\mathbb{K}$ and assume that $p : X \to \mathbb{R}$ satisfies

(i) $\forall x, y \in X, \quad p(x+y) \leq p(x) + p(y)$,

(ii) $\forall x \in X, \alpha \in \mathbb{K}, \quad p(\alpha x) = |\alpha| p(x)$.

Let Y be a linear manifold in X and $f : Y \to \mathbb{K}$ be a linear functional satisfying $|f(x)| \leq p(x)$ for all $x \in Y$. Then there is a linear functional $F : X \to \mathbb{K}$ satisfying $F(x) = f(x)$ for all $x \in Y$ and $|F(x)| \leq p(x)$ for all $x \in X$.

Remark 5.2: Functionals p satisfying (i) and (ii) above are called *seminorms*. We will discuss them in detail later in the course. It is worthwhile to observe now that (i) and (ii) imply that p is nonnegative and that $p(0) = 0$ because

$$\forall x \in X, \quad 2p(x) = p(x) + p(-x) \geq p(x - x) = p(0) = 0 \cdot p(x) = 0.$$

Proof of Theorem 5.1: **Case 1**: $\mathbb{K} = \mathbb{R}$. Using Theorem 4.6, we may choose a linear functional $F : X \to \mathbb{R}$ such that $F(x) = f(x)$ for all $x \in Y$ and $F(x) \leq p(x)$ for all $x \in X$. Notice that

$$-F(x) = F(-x) \leq p(-x) = p(x) \text{ for all } x \in X.$$

It follows that $|F(x)| \leq p(x)$ for all $x \in X$.

Case 2: $\mathbb{K} = \mathbb{C}$. Let us put

$$f_1(x) = \mathrm{Re}(f(x)), \quad f_2(x) = \mathrm{Im}(f(x)) \text{ for all } x \in Y,$$

so that

$$f_1, f_2 : Y \to \mathbb{R} \quad \text{and}$$

$$f(x) = f_1(x) + i f_2(x) \text{ for all } x \in Y.$$

Since $f(ix) = i f(x)$ for all $x \in X$ we have

$$f_1(ix) + i f_2(ix) = i f_1(x) - f_2(x) \text{ for all } x \in Y,$$

which implies that

$$f_2(x) = -f_1(ix) \text{ for all } x \in Y.$$

Observe further that

$$f_1(x+y) = f_1(x) + f_1(y), \quad f_2(x+y) = f_2(x) + f_2(y) \text{ for all } x, y \in X,$$

and

$$f_1(cx) = cf_1(x), \quad f_2(cx) = cf_2(x) \text{ for all } x \in X, \ c \in \mathbb{R}.$$

Let X_r denote the real linear space obtained by restricting X to real scalars. (Of course, X_r and X have the same elements, but they have a different linear structure; the addition function is the same, but scalar multiplication is restricted to $\mathbb{R} \times X$.) Let Y_r denote the linear manifold in X_r obtained by restricting Y to real scalars. Then $f_1 : Y_r \to \mathbb{R}$ is a linear functional satisfying $f_1(x) \leq p(x)$ for all $x \in Y_r$. By Theorem 4.6, we may choose a linear functional $F_1 : X_r \to \mathbb{R}$ satisfying $F_1(x) = f_1(x)$ for all $x \in Y_r$ and $F_1(x) \leq p(x)$ for all $x \in X_r$. Now let us define a functional $F : X \to \mathbb{C}$ by

$$F(x) = F_1(x) - iF_1(ix) \text{ for all } x \in X.$$

We want to show that F is linear (with respect to the complex linear structure) and that $|F(x)| \leq p(x)$ for all $x \in X$. It is clear that $F(x+y) = F(x) + F(y)$ for all $x, y \in X$. Let $x \in X$ and $\alpha \in \mathbb{C}$ be given and put

$$a = \mathrm{Re}(\alpha), \quad b = \mathrm{Im}(\alpha).$$

Then we have

$$\begin{aligned} F(\alpha x) &= F_1(ax + ibx) - iF_1(iax - bx) \\ &= aF_1(x) + bF_1(ix) - iaF_1(ix) + ibF_1(x). \end{aligned}$$

On the other hand we have

$$\begin{aligned} \alpha F(x) &= (a + ib)(F_1(x) - iF_1(ix)) \\ &= aF_1(x) + ibF_1(x) - iaF_1(x) + bF_1(ix). \end{aligned}$$

It follows that $F(\alpha x) = \alpha F(x)$ and $F : X \to \mathbb{C}$ is linear.

Clearly, we have $F(x) = f(x)$ for all $x \in Y$. It remains to show that $|F(x)| \leq p(x)$ for all $x \in X$. Let $x \in X$ be given and choose $\theta \in \mathbb{R}$ such that

$$F(x) = |F(x)|e^{i\theta}.$$

Then we have

$$|F(x)| = F(x)e^{-i\theta} = F(e^{-i\theta}x).$$

Since $|F(x)|$ is real , we must have $F(e^{-i\theta}x) = F_1(e^{-i\theta}x)$, so that

$$|F(x)| = F_1(e^{-i\theta}x) \leq p(e^{-i\theta}x) = p(x). \quad \square$$

If X is normed, then a very natural (and useful) choice for the functional p in the preceding theorem is $p(x) = m\|x\|$ for some $m > 0$. This observation leads us to the next result which says that a continuous linear functional defined on a linear manifold always has a norm-preserving extension to a linear functional defined on the entire space.

Theorem 5.3 (Hahn-Banach Theorem (Normed Linear Space)): Let $(X, \|\cdot\|)$ be a normed linear space, Y be a linear manifold in X and $f : Y \to \mathbb{K}$ be a continuous linear functional. Put

$$m = \sup\{|f(x)| : x \in Y, \|x\| \leq 1\}.$$

Then there exists $x^* \in X^*$ such that $\|x^*\| = m$ and $\langle x^*, x\rangle = f(x)$ for all $x \in Y$.

Proof: Define $p : X \to \mathbb{K}$ by $p(x) = m\|x\|$ for all $x \in X$. Then p satisfies conditions (i) and (ii) of Theorem 5.1 and we have $|f(x)| \leq m\|x\| = p(x)$ for all $x \in Y$. By Theorem 5.1, we may choose a linear functional $x^* : X \to \mathbb{K}$ satisfying $x^*(x) = f(x)$ for all $x \in Y$ and

$$|x^*(x)| \leq m\|x\| \;\text{ for all} x \in X.$$

It follows that $x^* \in X^*$ and $\|x^*\| \leq m$. We also have

$$\|x^*\| = \sup\{|x^*(x)| : x \in X, \|x\| \leq 1\} \geq \sup\{|x^*(x)| : x \in Y, \|x\| \leq 1\} = m,$$

and consequently $\|x^*\| = m$. $\square$

The three theorems given above fall into the category of algebraic (or extension) forms of the Hahn-Banach Theorem. These results have some important consequences concerning the existence of "interesting" and "useful" continuous linear functionals.

Proposition 5.4: Let $(X, \|\cdot\|)$ be a normed linear space and let $x_0 \in X\backslash\{0\}$ be given. Then there exists $x_0^* \in X^*$ such that $\|x_0^*\| = 1$ and $\langle x_0^*, x_0\rangle = \|x_0\|$.

Remark 5.4: The conclusion of Proposition 5.4 remains valid when $x_0 = 0$ provided that $X \neq 0$.

Proof of Proposition 5.4: Put $Y = \operatorname{span}(\{x_0\})$ and define $f : Y \to \mathbb{K}$ by

$$f(x) = \alpha(x)\|x_0\| \;\text{ for all } x \in Y,$$

where, for each $x \in Y$, $\alpha(x)$ is the unique element of $\mathbb{K}$ such that

$$x = \alpha(x)x_0.$$

Observe that $f(x_0) = \|x_0\|$. Let us put

$$m = \sup\{|f(x)| : x \in Y, \|x\| \leq 1\}.$$

Since

$$|f(x)| = |\alpha(x)| \cdot \|x_0\| = \|\alpha(x)x_0\| = \|x\| \text{ for all } x \in Y,$$

we conclude that $m = 1$. By Theorem 5.3, we may choose $x_0^* \in X^*$ such that $\|x_0^*\| = 1$ and $\langle x_0^*, x\rangle = f(x)$ for all $x \in Y$. It follows that $\langle x_0^*, x_0\rangle = \|x_0\|$. $\square$

Corollary 5.5: Let $((X, \|\cdot\|)$ be a normed linear space and let $x_0 \in X$ be given. If $\langle x^*, x_0\rangle = 0$ for all $x^* \in X^*$ then $x_0 = 0$.

Proposition 5.6: Let $(X, \|\cdot\|)$ be a normed linear space and let $x_0 \in X$ be given. Then

$$\|x_0\| = \sup\{|\langle x^*, x_0\rangle| : x^* \in X^*, \|x^*\| \leq 1\}.$$

Proof: If $x_0 = 0$, the result is immediate, so assume that $x_0 \neq 0$. Choose $x_0^* \in X^*$ as given by Proposition 5.4. Then we have

$$\sup\{|\langle x^*, x_0\rangle| : x^* \in X^*, \|x^*\| \leq 1\} \geq |\langle x_0^*, x_0\rangle| = \|x_0\|.$$

On the other hand, since $|\langle x^*, x_0\rangle| \leq \|x^*\| \cdot \|x_0\|$ for all $x^* \in X^*$, we have

$$\sup\{|\langle x^*, x_0\rangle| : x^* \in X^*, \|x^*\| \leq 1\} \leq \|x_0\|. \ \square$$

Proposition 5.7: Let X be a normed linear space and Y be a linear manifold in X. Let $d > 0$ and $x_0 \in X$ be given. Assume that

$$\inf\{\|y - x_0\| : y \in Y\} = d.$$

Then there exists $x^* \in X^*$ such that $\langle x^*, x_0\rangle = 1$, $\|x^*\| = \frac{1}{d}$ and $\langle x^*, y\rangle = 0$ for all $y \in Y$.

The proof of Proposition 5.7 is part of Assignment 4.

The following immediate consequence of Proposition 5.7 will be used frequently.

Corollary 5.8: Let $(X, \|\cdot\|)$ be a normed linear space, Y be a closed subspace of X, and $x_0 \in X \backslash Y$ be given. Then there exists $x^* \in X^*$ such that $\langle x^*, x_0\rangle = 1$ and $\langle x^*, y\rangle = 0$ for all $y \in Y$.

Convex Sets

The notions of *convex set* and *convex function* (or convex functional) play important roles in many branches of mathematics, including functional analysis. A subset S of a linear space is called convex provided that for every pair of points from S, the line segment joining these points lies entirely in S. We shall develop some basic properties of convex sets now and discuss convex functions later.

Definition 5.9: A set $K \subset X$ is said to be *convex* provided that

$$\forall x, y \in K, t \in [0,1], \quad tx + (1-t)y \in K.$$

Notice that the definition of convex set involves only real scalars. Consequently, a subset of a complex linear space X is convex if and only if the set is a convex subset of X_r, the restriction of X to real scalars. For this reason, many authors consider only real spaces when talking about convex sets (or are "a bit loose") about specifying whether the scalar field is $\mathbb{R}$ or $\mathbb{C}$. Here, unless stated otherwise, we allow the scalar field to be either $\mathbb{R}$ or $\mathbb{C}$.

We record below some elementary results concerning convex sets. The proofs of these results are very simple and are left as exercises.

Proposition 5.10: The intersection of any collection of convex sets is convex.

Proposition 5.11: Let K be a convex subset of X and let $x_1, x_2, \cdots, x_N \in K$ and $t_1, t_2, \cdots, t_N \geq 0$ be given and assume that $t_1 + t_2 + \cdots + t_N = 1$. Then $t_1 x_1 + t_2 x_2 + \cdots + t_N x_N \in K$.

The idea of the proof of Proposition 5.11 is to use induction on N.

Proposition 5.12: Let K_1, K_2 be convex subsets of X and let $\lambda \in \mathbb{K}$ be given. Then $K_1 + K_2$ is convex and λK_1 is convex.

Proposition 5.13: Let K be a convex subset of X and assume that $T : X \to Y$ is linear. Then $T[K]$ is convex.

Proposition 5.14: Let C be a convex subset of Y and assume that $T : X \to Y$ is linear. Then $\{x \in X : Tx \in C\}$ is convex.

In analogy with linear combination and span, it is useful to define *convex combination* and *convex hull.*

Definition 5.15 Let $(x_i | i \in I)$ be a family of elements of X. (Here, I can be any index set.) By a *convex combination* of $(x_i | i \in I)$ we mean a sum of the form

$$\sum_{i \in J} t_i x_i$$

where J is a finite subset of I, $(t_i | i \in J)$ is a family of nonnegative real numbers, and

$$\sum_{i \in J} t_i = 1.$$

Just as in the case of linear combination, the sum in the definition of convex combination is assumed to be finite.

Definition 5.16: Let $(x_i | i \in I)$ be a family of elements of X. The set of all convex combinations of $(x_i | i \in I)$ is called the *convex hull* of $(x_i | i \in I)$ and is denoted $\mathrm{co}(x_i | i \in I)$.

Remark 5.17: The notions of convex combination and convex hull apply to sets $S \subset X$ by making S into a family through self-indexing.

Proposition 5.18: Let $S \subset X$ be given. Then

$$\mathrm{co}(S) = \cap \mathcal{C},$$

where $\mathcal{C}$ is the collection of all convex sets $K \subset X$ such that $S \subset K$.

Proof: Notice that $\mathcal{C} \neq \emptyset$ since $X \in \mathcal{C}$ and by Proposition 5.10, $\cap \mathcal{C}$ is convex. Since every member of $\mathcal{C}$ includes S, we have $S \subset \cap \mathcal{C}$, so it follows from Proposition 5.11 and convexity of $\cap \mathcal{C}$ that $\mathrm{co}(S) \subset \mathcal{C}$. To establish the reverse inclusion, it suffices to show that $\mathrm{co}(S)$ is convex. To this end, let $v, w \in \mathrm{co}(S)$ be given. Then, we may choose (finite) families $(x_i | i = 1, 2, \cdots, m)$, $(y_j | j = 1, 2. \cdots, n)$ of points in S and corresponding families $(\nu_i | i = 1, 2, \cdots, m)$, $(\tau_j | j = 1, 2, \cdots, n)$ of nonnegative real numbers such that

$$\sum_{i=1}^{m} \nu_i = \sum_{j=1}^{n} \tau_j = 1,$$

and

$$\sum_{i=1}^{m} \nu_i x_i = v, \quad \sum_{j=1}^{n} \tau_j y_j = w.$$

Let $t \in [0, 1]$ be given and observe that $t\nu_i \geq 0$, $(1-t)\tau_j \geq 0$ for all $i = 1, 2, \cdots, m$, $j = 1, 2, \cdots, n$, and

$$\sum_{i=1}^{m} t\nu_i + \sum_{j=1}^{m} (1-t)\tau_j = 1.$$

It follows that $tv + (1-t)w$ is a convex combination of $\{x_1, \cdots, x_m, y_1, \cdots, y_n\}$ because

$$tv + (1-t)w = \sum_{i=1}^{m} t\nu_i x_i + \sum_{j=1}^{n} (1-t)\tau_i y_j. \quad \square$$

Remark 5.19: It is more common to define the convex hull of a set S to be the intersection of all convex sets including S and then prove that the convex hull of S is equal to the set of all convex combinations of S.

Convex Sets, Internal Points, and Separation by Linear Functionals

Let X be a linear space over $\mathbb{K}$

Definition 5.20: Let $S \subset X$. A point $x_0 \in X$ is said to be an *internal point* of S provided that for every $x \in X$, there exists $\epsilon > 0$ such that

$$x_0 + \lambda x \in S \tag{1}$$

for all $\lambda \in \mathbb{K}$ with $|\lambda| < \epsilon$.

Every internal point of S must belong to S. If X is a NLS and x_0 is an interior point of S, then x_0 is an internal point of S. However, even in finite dimensions, an internal point need not be an interior point. (A simple example is given with $X = \mathbb{R}^2$, $x_0 = 0$ and S is the set described in polar coordinates by $0 \leq r \leq \theta, \quad 0 < \theta \leq 2\pi$.) However, for convex sets in finite-dimensional NLS, every internal point is an interior point.)

Remark 5.21: In the definition of internal point, some authors require that (1) hold only for real λ (with $|\lambda| < 1$) even when $\mathbb{K} = \mathbb{C}$. These two definitions are not equivalent in general. However, as we shall show below, for convex sets, these two definitions of internal point are equivalent. This will not be an important issue for us, because we shall use the notion of internal point only for convex sets.

Proposition 5.22: Let X be a linear space over $\mathbb{C}$ and assume that $K \subset X$ is convex. Let $x_0 \in X$ be given and assume that for every $x \in X$ there exists $\epsilon > 0$ such that

$$x_0 + \lambda x \in K \quad \text{for all } \lambda \in \mathbb{R} \text{ with } |\lambda| < \epsilon. \tag{2}$$

Then x_0 is an internal point of K.

Proof: Let $v \in X$ be given. We apply (2) using $x = v$ and $x = iv$ to choose $\epsilon_1, \epsilon_2 > 0$ such that

$$x_0 + cv \in K \quad \text{for all } c \in \mathbb{R} \text{ with } |c| < \epsilon_1, \tag{3}$$

$$x_0 + idv \in K \quad \text{for all } d \in \mathbb{R} \text{ with } |d| < \epsilon_2. \tag{4}$$

Put $\epsilon = \frac{1}{2}\min\{\epsilon_1, \epsilon_2\}$ and let $\lambda \in \mathbb{C}$ be given with $|\lambda| < \epsilon$. Then we may choose $a, b \in \mathbb{R}$ with $|2a| < \epsilon_1$ and $|2b| < \epsilon_2$ such that $\lambda = a + ib$. It follows that $x_0 + 2av \in K$, $x_0 + 2ibv \in K$. Since K is convex, we have

$$x_0 + \lambda v = \frac{1}{2}(x_0 + 2av) + \frac{1}{2}(x_0 + 2ibv) \in K. \ \square$$

Definition 5.23: Let K be a convex subset of X having 0 as an internal point. The *Minkowski functional* for K (about 0) is the function $p^K : X \to \mathbb{R}$ defined by

$$p^K(x) = \inf\{t \in (0, \infty) : t^{-1}x \in K\} = \inf\{s \in (0, \infty) : x \in sK\} \quad \text{for all } x \in X.$$

Remark 5.24: The function p^K defined above is sometimes called the *support function* or *gauge* of K. It can be defined for more general sets, but then it will not have all of the properties that we develop below. Also, one should be aware that the term "support function" of a convex set has other meanings as well.

Example 5.25: Let X be a normed linear space and let $K = B_1(0) = \{x \in X : \|x\| < 1\}$. Then

$$p^K(x) = \inf\{s \in (0,\infty) : x \in sK\} = \inf\{s \in (0,\infty) : \|x\| < s\} = \|x\| \quad \text{for all } x \in X.$$

Before proving a basic result about the Minkowski functional, it is useful to introduce some more definitions.

Definition 5.26: Let S be a subset of X. We say that S is *absorbing* provided that for every $x \in X$, there exists $\epsilon > 0$ such that

$$\lambda x \in S$$

for all $\lambda \in \mathbb{K}$ with $|\lambda| < \epsilon$.

Remark 5.27

(a) It is immediate that a set S is absorbing if and only if 0 is an internal point.

(b) Some authors consider only real scalars in the definition of absorbing set, even when the scalar field is $\mathbb{C}$. In view of Proposition 5.21, this gives the same notion of "absorbing" for convex sets.

(c) Absorbing sets are sometimes referred to as *radial sets*.

Definition 5.28: Let S be a subset of X. We say that S is *balanced* provided that $\lambda S \subset S$ for all $\lambda \in \mathbb{K}$ with $|\lambda| \le 1$.

Remark 5.29: Let S be a subset of X and let $\alpha \in \mathbb{K}$ be given with $|\alpha| = 1$ be given. Then $\alpha S = S$ because $\alpha S \subset S$ and $\alpha^{-1} S \subset S$.

Remark 5.30: Balanced sets are sometimes called *circled sets*.

Remark 5.31: The terms "absorbing set", "circled set", and "radial set" are sometimes given different definitions than those mentioned above because the conditions are required to hold only for a more restricted set of scalars. The different definitions that I have encountered all seem to agree with the ones given here for convex sets. If you are using applying these concepts to sets that need not be convex, you should be very careful about the precise forms of the definitions.

Lemma 5.32: Let K be a subset of X. Assume that K is convex and absorbing. Then

(a) $p^K(ax) = ap^K(x)$ for all $x \in X, a \ge 0$,

(b) $\{x \in X : p^K(x) < 1\} \subset K \subset \{x \in X : p^K(x) \leq 1\}$,

(c) $p^K(x+y) \leq p^K(x) + p^K(y)$ for all $x, y \in X$.

If, in addition, K is balanced then

(d) $p^K(\alpha x) = \|\alpha| p^K(x)$ for all $x \in X$, $\alpha \in \mathbb{K}$.

Proof: For each $x \in X$, define the set

$$H^K(x) = \{t \in (0,\infty) : t^{-1}x \in K\}.$$

Since K is convex and $0 \in K$, we see that each $H^K(x)$ is an interval of the form $(p^K(x), \infty)$ or $[p^K(x), \infty)$. Moreover, for each $a > 0$ we have $H^K(ax) = aH^K(x)$ for all $x \in X$. (Notice that $p^K(0) = 0$ and $H^K(0) = (0,\infty)$.) By definition, we have

$$p^K(x) = \inf H^K(x) \text{ for all } x \in X.$$

Parts (a) and (b) of the lemma follow easily from these observations. To prove (c), let $x, y \in X$ and $c > p^K(x) + p^K(y)$ be given. Then we may write $c = a + b$ with $a > p^K(x)$ and $b > p^K(y)$. Observe that

$$\frac{x+y}{c} = \frac{x+y}{a+b} = \frac{a(a^{-1}x) + b(b^{-1}y)}{a+b}. \tag{5}$$

Since $a^{-1}x, b^{-1}y \in K$ and K is convex, it follows from (5) that $c^{-1}(x+y) \in K$. We conclude that $p^K(x+y) \leq c$. Since this holds for all $c > p^K(x) + p^K(y)$, we conclude that $p^K(x+y) \leq p^K(x) + p^K(y)$.

To prove (d), assume that K is balanced. Let $x \in X, \alpha \in \mathbb{K}$ be given. If $\alpha = 0$, we are done. Assume that $\alpha \neq 0$. Since K is balanced, we have $|\alpha|^{-1}\alpha K = K$. Consequently, we have

$$\begin{aligned}
p^K(\alpha x) &= \inf\{t \in (0,\infty) : \alpha x \in tK\} \\
&= \inf\{t \in (0,\infty) : \alpha x \in t\tfrac{\alpha}{|\alpha|}K\} \\
&= \inf\{t \in (0,\infty) : x \in \tfrac{t}{|\alpha|}K\} \\
&= \inf\{|\alpha|s \in (0,\infty) : x \in sK\} \\
&= |\alpha| \inf\{s \in (0,\infty) : x \in sK\} = |\alpha| p^K(x). \quad \square
\end{aligned}$$

Definition 5.33: Let M, N be subsets of X. A linear functional $F : X \to \mathbb{K}$ is said to *separate* M and N provided there exists $c \in \mathbb{R}$ such that either

$$\mathrm{Re}(F(x)) \leq c \text{ for all } x \in M \text{ and } \mathrm{Re}(F(y)) \geq c \text{ for all } y \in N,$$

or
$$\mathrm{Re}(F(x)) \geq c \ \text{ for all } x \in M \ \text{ and } \ \mathrm{Re}(F(y)) \leq c \ \text{ for all } y \in N.$$

Lemma 5.34: Let M, N be nonempty subsets of X and assume that $F : X \to \mathbb{K}$ is linear. Then the following three statements are equivalent:

(i) F separates M and N.

(ii) F separates $M - N$ and $\{0\}$.

(iii) Either
$$\mathrm{Re}(F(x)) - \mathrm{Re}(F(y)) \leq 0 \ \text{ for all } x \in M, y \in N,$$
or
$$\mathrm{Re}(F(x)) - \mathrm{Re}(F(y)) \geq 0 \ \text{ for all } x \in M, y \in N.$$

The proof of Lemma 5.34 is quite elementary and is left as an exercise.

Lecture Notes for Week 6

Convex Sets, Internal Points, and Separation by Linear Functionals (Continued)

We now state and prove a basic result concerning separation of convex sets in a linear space.

Theorem 6.1: Let X be a linear space over $\mathbb{K}$ and K_1, K_2 be nonempty convex subsets of X. Assume that $K_1 \cap K_2 = \emptyset$ and that K_1 has an internal point. Then there is a nonzero linear functional that separates K_1 and K_2.

Proof: Assume first that $\mathbb{K} = \mathbb{R}$. Choose an internal point $x_1 \in K_1$ and any point $x_2 \in K_2$. Put $x_0 = x_1 - x_2$. Then $K_1 - K_2$ is convex and has x_0 as an internal point. (The convexity of $K_1 - K_2$ follows from Proposition 5.12. Notice that $x_1 - x_2$ is an internal point of $-x_2 + K_1$ because for every $x \in X$ and $\lambda \in \mathbb{R}$ we have $x_1 - x_2 + \lambda x \in -x_2 + K_1$ if and only if $x_1 + \lambda x \in K_1$. Since $-x_2 + K_1 \subset K_1 - K_2$, it follows that $x_1 - x_2$ is an internal point of $K_1 - K_2$.) Put

$$K = -x_0 + (K_1 - K_2)$$

and observe that K is convex and has 0 as an internal point. Since $K_1 \cap K_2 = \emptyset$, we know that $0 \notin K_1 - K_2$. This implies that $-x_0 \notin K$ and consequently $p^K(-x_0) \geq 1$ by part (b) of Lemma 5.32. Let $Y = \mathrm{span}(\{x_0\})$ and define the linear functional $f : Y \to \mathbb{R}$ by

$$f(\alpha x_0) = -\alpha \text{ for all } \alpha \in \mathbb{R}.$$

By parts (a) and (c) of Lemma 5.32 we know that p^K can be used as the comparison functional in the Hahn-Banach theorem for real linear spaces. If $\alpha \geq 0$, then

$$f(\alpha x_0) \leq 0 \leq p^K(\alpha x_0).$$

If $\alpha < 0$, then we have

$$f(\alpha x_0) = -\alpha \leq -\alpha p^K(-x_0) = p^K(\alpha x_0),$$

since $p^K(-x_0) \geq 1$ and nonnegative scalars can be factored out of p^K. It follows that

$$f(x) \leq p(x) \text{ for all } x \in Y.$$

By the Hahn-Banach theorem for real spaces, we may choose a linear functional $F : X \to \mathbb{R}$ such that $F(x) = f(x)$ for all $x \in Y$ and

$$F(x) \leq p^K(x) \text{ for all } x \in X.$$

We need to show that F separates K_1 and K_2. Let $x \in K_1$, $y \in K_2$ be given and notice that $x - y - x_0 \in K$ which tells us that $p^K(x - y - x_0) \le 1$. Using the linearity of F, we find that

$$F(x) - F(y) - F(x_0) = F(x - y - x_0) \le p^K(x - y - x_0) \le 1.$$

Since $F(x_0) = -1$, we obtain

$$F(x) - F(y) \le 0 \ \text{ for all } x \in K_1, y \in K_2.$$

By Lemma 5.34, F separates K_1 and K_2.

Assume now that $\mathbb{K} = \mathbb{C}$. Let X_r denote the linear space obtained by restricting X to the subfield of real scalars. Then K_1, K_2 are nonempty disjoint convex subsets of X_r and K_1 has an internal point. We may apply the separation theorem for real spaces to obtain a linear functional $F_1 : X_r \to \mathbb{R}$ satisfying

$$F_1(x) - F_1(y) \le 0 \ \text{ for all } x \in K_1, y \in K_2.$$

Now F_1 satisfies

$$F_1(x + y) = F_1(x) + F_1(y), \quad F_1(ax) = aF_1(x) \ \text{ for all } x, y \in X, a \in \mathbb{R}.$$

Let us define $F : X \to \mathbb{C}$ by

$$F(z) = F_1(z) - iF_1(iz) \ \text{ for all } z \in \mathbb{C}.$$

It is straightforward to check that F is a linear functional (with respect to the linear structure over $\mathbb{C}$). Moreover, since $\mathrm{Re}(F(z)) = F_1(z)$ for all $z \in X$, we have

$$\mathrm{Re}(F(x)) - \mathrm{Re}(F(y)) \le 0 \ \text{ for all } x \in K_1, y \in K_2,$$

so that F separates K_1 and K_2 by Lemma 5.34. $\square$

Theorem 6.1 has numerous important applications and numerous important generalizations and variants. Several remarks pertaining to generalizations and variants are given below.

Remark 6.2: Convexity of both sets is essential in Theorem 6.1, even when X is finite dimensional. To see why, let $X = \mathbb{R}^2$,

$$K_1 = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 \le e^{-x_1}\}, \quad K_2 = \{(0, 2)\}.$$

You should convince yourself that any line which is above $(0, 1)$ and below $(0, 2)$ will have a nonempty intersection with K_2.

Remark 6.3: If X is finite dimensional the assumption that K_1 has an internal point can be dropped and the assumption that $K_1 \cap K_2 = \emptyset$ can be weakened.

Remark 6.4: The separation ensured by Theorem 6.1 need not be strict, even when X is finite dimensional. (Here by strict separation we mean that either

$$\mathrm{Re}(F(x)) - \mathrm{Re}(F(y)) < 0 \text{ for all } x \in M,\ y \in N$$

or

$$\mathrm{Re}(F(x)) - \mathrm{Re}(F(y)) > 0 \text{ for all } x \in M,\ y \in N.)$$

Strict separation or even separation with a "gap" can be ensured by additional assumptions.

Remark 6.5: There are numerous important variants of the basic separation theorem. See *Convex Analysis and Variational Problems* by I. Ekeland and R. Temam, or *Convex Analysis* by R. T. Rockafellar.

If X is normed, it is often desirable to separate two convex sets by a continuous linear functional. The assumptions made on K_1 and K_2 in Theorem 6.1 are not sufficient for this purpose. However, if the assumption that K_1 has an internal point is strengthened to $\mathrm{int}(K_1) \neq \emptyset$ then the separation of K_1 and K_2 can be achieved by an element of X^*.

Theorem 6.6: Let X be a normed linear space over $\mathbb{K}$ and let K_1, K_2 be nonempty convex subset of X. Assume that $K_1 \cap K_2 \neq \emptyset$ and that $\mathrm{int}(K_1) = \emptyset$. Then there is a nonzero continuous linear functional that separates K_1 and K_2.

The proof of Theorem 6.6 is part of Assignment 4.

Seminorms and Convexity

Let X be a linear space over $\mathbb{K}$.

Definition 6.7: A function $p : X \to \mathbb{R}$ is said to be a *seminorm* provided that

(i) $\forall x, y \in X$, we have $p(x + y) \leq p(x) + p(y)$,

(ii) $\forall x \in X, \alpha \in \mathbb{K}$, we have $p(\alpha x) = |\alpha| p(x)$.

Notice that seminorms are precisely the kind of comparison functions appearing in the Hahn-Banach theorem for real or complex spaces.

Proposition 6.8: Assume that $p : X \to \mathbb{R}$ is a seminorm. Then

(a) $p(0) = 0$,

(b) $p(x) \geq 0$ for all $x \in X$,

(c) $|p(x) - p(y)| \leq p(x - y)$,

(d) Put $Y = \{x \in X : p(x) = 0\}$. Then Y is a linear manifold,

(e) Put $K = \{x \in X : p(x) < 1\}$. Then, K is convex, balanced, and absorbing. Moreover, $p = p^K$, where p^K is the Minskowski functional for K.

Proof: To prove (a), we let $x \in X$ be given and notice that $p(0) = p(0x) = 0p(x) = 0$.

To prove (b) we let $x \in X$ be given and observe that

$$2p(x) = p(x) + p(-x) \geq p(x - x) = 0.$$

To prove (c), let $x, y \in X$ be given. Then we have

$$p(x) - p(y) = p(x - y + y) - p(y) \leq p(x - y) + p(y) - p(y) = p(x - y),$$

$$p(y) - p(x) = p(y - x + x) - p(x) \leq p(y - x) + p(x) - p(x) = p(y - x) = p(x - y).$$

It follows that $|p(x) - p(y)| \leq p(x - y)$.

To prove (d), we first observe that $0 \in Y$ by (a). Now, let $x, y \in Y, \alpha, \beta \in \mathbb{K}$ be given. Then, since $p(x) = p(y) = 0$, we have

$$0 \leq p(\alpha x + \beta y) \leq p(\alpha x) + p(\beta y) = |\alpha| p(x) + |\beta| p(y) = 0,$$

which implies $\alpha x + \beta y \in Y$.

To show that K is convex, let $x_1, x_2 \in K, t \in [0, 1]$ be given. Then, since $0 \leq p(x_1), p(x_2) < 1$ and $0 \leq t, (1 - t) \leq 1$, we have

$$p(tx_1 + (1 - t)x_2) \leq p(tx_1) + p((1 - t)x_2) = tp(x_1) + (1 - t)p(x_2) < 1.$$

To show that K is balanced, let $\alpha \in \mathbb{K}$ with $|\alpha| \leq 1$ and $x \in \alpha K$ be given. We need to show that $x \in K$. Choose $y \in K$ such that $x = \alpha y$, and note that $p(y) < 1$. Then we have

$$p(x) = p(\alpha y) = |\alpha| p(y) < 1,$$

which tells us that $x \in K$.

To show that K is absorbing, let $x \in X$ be given. If $p(x) = 0$, then $p(\lambda x) = 0$ (which implies $\lambda x \in K$) for all $\lambda \in \mathbb{K}$. If $p(x) > 0$, then $p(\lambda x) < 1$ (i.e., $\lambda x \in K$) for all $\lambda \in \mathbb{K}$ with $|\lambda| < 1/p(x)$.

To show that $p = p^K$, let $x \in X$ and $s > p(x)$ be given. Then we have $p(s^{-1}x) < 1$. This tells us that $s^{-1}x \in K$, which means that $x \in sK$. We also conclude that $p^K(x) \leq s$. Since

$$p^K(x) \leq s \quad \text{for all } s > p(x),$$

it follows $p^K(x) \leq p(x)$.

To establish the reverse inequality, let $x \in X$ be given and assume that $p(x) > 0$.(If $p(x) = 0$, then trivially, $p(x) \leq p^K(x)$.) Let $t \in \mathbb{R}$ with $0 < t \leq p(x)$ be given. Then $p(t^{-1}x) \geq 1$ which tells us that $t^{-1}x \notin K$. Since this holds for every $t \in (0, p(x)]$, we conclude that $p(x) \leq p^K(x)$. $\square$

Definition 6.9: A family $(p_i|i \in I)$ of seminorms is said to be *separating* provided that

$$\forall x \in X\backslash\{0\},\ \exists i \in I \text{ such that } p_i(x) > 0.$$

Separating families of seminorms can be used to generate Hausdorff topologies on linear spaces such that addition and scalar multiplication are continuous (and having some useful additional properties).

Remark 6.10: Observe that a family $(p_i|i \in I)$ of seminorms is separating if and only if

$$\{x \in X : p_i(x) = 0 \text{ for all } i \in I\} = \{0\}.$$

Remark 6.11: If a finite family $(p_i|i = 1, 2, \cdots, N)$ of seminorms is separating then clearly $p_1 + p_2 + \cdots + p_N$ is a norm on X.

Remark 6.12: Let $(p_i|i \in \mathbb{N})$ be a countable family of seminorms that is separating and define $\rho : X \times X \to \mathbb{R}$ by

$$\rho(x, y) = \sum_{i=1}^{\infty} 2^{-i} \frac{p_i(x-y)}{1 + p_i(x-y)} \quad \text{for all } x, y \in X.$$

Then ρ is a translation invariant metric such that each p_i is continuous from (X, ρ) to $\mathbb{R}$. (Here translation invariance simply means that $\rho(x+z, y+z) = \rho(x, y)$ for all $x, y, z \in X$.) One could replace $(2^{-i}|i \in \mathbb{N})$ with another sequence $(a_i|i \in \mathbb{N})$ with $a_i > 0$ for all $i \in \mathbb{N}$ and $\sum_{i=1}^{\infty} a_i < \infty$. For some purposes, it is more convenient to use the metric given by

$$\hat{\rho}(x, y) = \max\left\{\frac{1}{i}\min\{1, p_i(x-y)\} : i \in \mathbb{N}\right\}.$$

Remark 6.13: Any separating family $(p_i|i \in I)$ of seminorms can be used to construct a Hausdorff topology on X by taking finite intersections of sets of the form

$$\{x \in X : p_i(x - x_0) < \epsilon\}, \quad x_0 \in X, \epsilon > 0$$

as a base.

Example 6.14: Let $\mathbb{K} = \mathbb{R}$ and $X = C(\mathbb{R}; \mathbb{R})$, the set of all continuous functions from $\mathbb{R}$ to $\mathbb{R}$. For each $i \in \mathbb{N}$, put

$$p_i(f) = \max\{|f(x)| : x \in [-i, i]\} \quad \text{for all } f \in X.$$

It is clear that each p_i is a seminorm (but not a norm). Now define

$$\rho(f,g) = \sum_{i=1}^{\infty} 2^{-i} \frac{p_i(f-g)}{1+p_i(f-g)} \quad \text{for all } x, y \in X.$$

Then (X, ρ) is a complete metric space. You should check as an exercise that $f_n \to f$ as $n \to \infty$ in (X, ρ) if and only if for each compact subset $K \subset \mathbb{R}$, f_n converges to f uniformly on K as $n \to \infty$. The topology induced by this metric is a very natural. Since the evaluation mappings $a \to f(a)$ are continuous for every $a \in \mathbb{R}$, it follows from Exercise 14 on Assignment 2, that this topology cannot be induced by a norm.

Remark 6.15: A linear space together with a topology induced by a complete invariant metric (such that scalar multiplication is continuous) is called a Fréchet space. Some authors require the topology to be induced by a metric for which open balls are convex. (It is important to note that two different metrics can induce the same topology and one of the metrics has the properties that all open balls are convex, while the other has the property that no open balls are convex.) We will discuss such spaces in detail later in the course.

Example 6.16: l^p with $0 < p < 1$. Given $p \in (0,1)$, let l^p denote the set of all $x \in \mathbb{K}^{\mathbb{N}}$ such that

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

It can be shown that l^p is a linear space and that the function $\rho : X \times X \to \mathbb{R}$ defined by

$$\rho(x,y) = \sum_{k=1}^{\infty} |x_k - y_k|^p$$

is a complete invariant metric on l^p. Balls in this metric are not convex. Moreover, it can be shown that there is no complete invariant metric having balls which are convex that induces the same topology. This lack of convexity in the topology leads to some unusual properties. In particular the only continuous linear functional on l^p is the zero functional.

Duality and Reflexivity

Let X and Y be NLS over the same field.

Definition 6.17: We say that X and Y are *isomorphic* provided that there is a bijective linear mapping $T : X \to Y$ such that T and T^{-1} are continuous. Any such mapping T is called an *isomorphism* between X and Y.

Definition 6.18: A linear mapping $T : X \to Y$ is said to be an *isometry* provided that $\|Tx\| = \|x\|$ for all $x \in X$.

Remark 6.19:

(a) If T is an isometry, then T is continuous and injective.

(b) If an isometry T is surjective (and hence bijective) then T^{-1} is an isometry.

Definition 6.20: We say that X and Y are *isometrically isomorphic* provided that there is a surjective (linear) isometry $T : X \to Y$.

Let X^* denote the (topological) dual of X and $\langle \cdot, \cdot \rangle : X^* \times X \to \mathbb{K}$ denote the duality pairing, i.e.

$$\langle x^*, x \rangle = x^*(x) \quad \text{for all } x^*, \in X^*, x \in X.$$

Recall that the canonical injection J of X into X^{**} is defined by

$$(J(x))(x^*) = \langle x^*, x \rangle \quad \text{for all } x \in X, x^* \in X^*.$$

By Proposition 5.6, we have

$$\|J(x)\|_{**} = \sup\{|\langle x^*, x \rangle| : \|x^*\|_* \leq 1\} = \|x\| \quad \text{for all } x \in X,$$

and consequently J is an isometry. (This, of course, ensures that J is indeed injective.)

Definition 6.21: We say that X is reflexive if J is surjective.

Remark 6.22: If X is reflexive then X is isometrically isomorphic to X^{**}. In 1951, R.C. James gave an example of a Banach space X such that X is isometrically isomorphic to X^{**} but X is not reflexive. (*Proc. Nat. Acad. Sci.* **37**, 174-177) Consequently, in the definition of reflexivity, it is essential that the canonical injection of X into X^{**} be an isomorphism (rather than just requiring some isometric isomorphism to exist).

Prop. 6.23: If X is reflexive, then X is complete.

Proof: Let $\{x_n\}_{n=1}^{\infty}$ be a Cauchy sequence in X. Then $\{J(x_n)\}_{n=1}^{\infty}$ is a Cauchy sequence in X^{**} since J is an isometry. X^{**} is complete since it is the dual of X^*. We may choose $x^{**} \in X^{**}$ such that $J(x_n) \to x^{**}$ as $n \to \infty$. Put $x = J^{-1}(x^{**})$. Since J is an isometry, we know that

$$\|x_n - x\| = \|J(x_n) - x^{**}\| \to 0 \quad \text{as } n \to \infty. \quad \square$$

Lecture Notes for Week 7 (First Draft)

Duaality and Reflexivity (Continued)

Theorem 7.1: Assume that X is reflexive and let Y be a closed subspace of X. Then Y is reflexive.

Proof: Define the linear mapping $A : X^* \to Y^*$ by

$$(Ax^*)(y) = x^*(y) \ \text{ for all } x^* \in X^*, \ y \in Y.$$

In other words, A is simply the restriction operator for bounded linear functionals on X to the smaller domain Y. Observe that

$$\|Ax^*\|_{Y^*} \le \|x^*\|_{X^*} \ \text{ for all } x^* \in X^*. \tag{1}$$

Now define the linear mapping $B : Y^{**} \to X^{**}$ by

$$(By^{**})(x^*) = y^{**}(Ax^*) \ \text{ for all } y^{**} \in Y^{**}, \ x^* \in X^*.$$

By virtue of (1), we have

$$|(By^{**})(x^*)| \le \|y^{**}\|_{Y^{**}}\|Ax^*\|_{Y^*} \le \|y^{**}\|_{Y^{**}}\|x^*\|_{X^*}.$$

Let J_X, J_Y denote the canonical injections of X into X^{**} and Y into Y^{**}, respectively. (It is probably helpful to draw a diagram showing the sets $X, X^*, X^{**}, Y, Y^*, Y^{**}$ and indicating the mappings A, B, J_X, J_Y.) Let $y^{**} \in Y^{**}$ be given and put

$$x = J_X^{-1}(By^{**}).$$

Let us show that $x \in Y$. (This is a key step in the argument; in particular, it is here that we make use of the fact that Y is closed.) Suppose that $x \notin Y$. Then, since Y is closed, we may choose $x^* \in X^*$ such that $x^*(x) = 1$ and

$$x^*(y) = 0 \ \text{ for all } y \in Y,$$

by Corollary 5.8 (a consequence of the Hahn-Banach Theorem). This implies that $Ax^* = 0$. Therefore, we have

$$\begin{aligned} 0 = y^{**}(Ax^*) &= (By^{**})(x^*) \\ &= (J_X x)(x^*) \\ &= x^*(x). \end{aligned}$$

Since $x^*(x) = 1$, we have a contradiction, from which we conclude that $x \in Y$. Therefore, we have shown that

$$J_x^{-1}[B[Y^{**}]] \subset Y.$$

Let $y_0^{**} \in Y^{**}$ be given. We want to find $y_0 \in Y$ such that $J_Y(y_0) = y_0^{**}$. Let us put

$$y_0 = J_X^{-1}(By_0^{**}),$$

and observe that $y_0 \in Y$. Given $y^* \in Y^*$, let $x^* \in X^*$ be an extension of y^*, so that

$$y^* = Ax^*.$$

(We can choose such an extension by virtue of the Hahn-Banach Theorem.) Then we have

$$\begin{aligned} y_0^{**}(y^*) &= (By_0^{**})(x^*) \\ &= (J_X y_0)(x^*) \\ &= x^*(y_0) \\ &= y^*(y_0). \end{aligned}$$

It follows that $y_0^{**} = J_Y(y_0)$ and consequently Y is reflexive. $\square$

Proposition 7.2: Let X and Y be isomorphic normed linear spaces and assume that Y is reflexive. Then X is reflexive.

Proof: Choose a linear isomorphism $T : X \to Y$, i.e. a linear mapping such that T and T^{-1} are bounded. Define $U : X^* \to Y^*$ by

$$(Ux^*)(y) = x^*(T^{-1}y) \text{ for all } x^* \in X^*,\ y \in Y. \tag{2}$$

I claim that U is a (linear) isomorphism. To see that U is injective, let $x^* \in X^*$ be given and assume that $Ux^* = 0$. Using (2) we find that $x^*(T^{-1}y) = 0$ for all $y \in Y$, which implies that $x^*(x) = 0$ for all $x \in X$ and consequently U is injective. U is bounded because

$$|(Ux^*)(y)| \leq \|x^*\|\|T\|\|y\| \text{ for all } x^* \in X^*,\ y \in Y,$$

and consequently

$$\|Ux^*\| \leq \|T\|\|x^*\| \text{ for all } x^* \in X^*.$$

To see that T is surjective, let $\hat{y}^* \in Y^*$ be given and put define $\hat{x}^* \in X^*$ by

$$\hat{x}^*(x) = \hat{y}^*(Tx) \text{ for all } x \in X.$$

Using (2), we see that

$$U(\hat{x}^*)(y) = \hat{x}^*(T^{-1}y) = \hat{y}^*(TT^{-1}y) = \hat{y}^*(y) \text{ for all } y \in Y.$$

Since X^* and Y^* are complete, it follows the Bounded Inverse Theorem that U^{-1} is bounded.

Now define the linear mapping $V : X^{**} \to Y^{**}$ by

$$(Vx^{**})(y^*) = x^{**}(U^{-1}y^*) \text{ for all } x^{**} \in X^{**},\ y^* \in Y^* \tag{3}$$

I claim that V is also an isomorphism. (The details are similar to those for U. To establish the surjectivity, given $\hat{y}^{**} \in Y^{**}$, put $\hat{x}^{**}(x^*) = \hat{y}^{**}(Ux^*)$ for all $x^* \in X^*$.)

Let $x_0^{**} \in X^{**}$ be given and put

$$y_0 = J_Y^{-1}(VX_0^* *), \tag{4}$$

$$x_0 = T^{-1}y_0. \tag{5}$$

In order to show that J_X is surjective, it suffices to show that

$$x^*(x_0) = x_0^{**}(x^*) \text{ for all } x^* \in X^*.$$

To this end, let $x^* \in X^*$ be given and put

$$y^* = Ux^*. \tag{6}$$

Then we have

$$\begin{aligned} x_0^{**}(x^*) &= (Vx_0^{**}(y^*) \text{ using } (3), (6) \\ &= (J_Y(y_0))(y^*) \text{ using } (4) \\ &= y^*(y_0) \text{ by the definition of } J_Y \\ &= (Ux^*)(Tx_0) \text{ using } (5), (6) \\ &= x^*(x_0) \text{ using } (2). \end{aligned}$$

Remark 7.3: If $T : X \to Y$ is an isometric isomorphism, then the mappings U and V defined by (2) and (3) are also isometric isomorphisms.

Theorem 7.4: Let X Banach space Then X is reflexive if and only if X^* is reflexive.

Remark 7.5: It is important to require that X be a Banach space in Theorem 7.3 because it can happen that an incomplete normed linear space has a reflexive dual space.

Proof of Theorem 7.4: Assume that X is reflexive. Then X^{**} is reflexive by Proposition 7.2. Let $x_0^{***} \in (X^*)^{**}$ be given. We want to produce $x_0^* \in X^*$ such that $x_0^{***}(x^{**}) = x^{**}(x_0^*)$ for all $x^{**} \in X^{**}$. Let us define $x_0^* \in X^*$ by

$$x_0^*(x) = x_0^{***}(J_X(x)) \text{ for all } x \in X.$$

Obsereve that

$$x^{**}(x_0^*) = (J_X(x))(x_0^*) = x_0^*(x) = x_0^{***}(J_X(x)) = x_0^{***}(x^{**}) \text{ for all } x^{**} \in X^{**}.$$

It follows that X^* is reflexive.

To establish the converse, assume that X^* is reflexive. Then by what we just proved above, we know that X^{**} is reflexive. Since X is complete, we see that $J_X[X]$ is complete and hence closed. Since $J_X[X]$ is a closed subspace of X^{**}, it follows from Theorem 7.1 that $J_X[X]$ is reflexive. Since X and $J_X[X]$ are (isometrically) isomorphic, we conclude that X is reflexive. $\square$

Weak and Weak Convergence*

We now introduce and study a very important type of convergence, known as "weak convergence" in a normed linear space X. We also study a related type of convergence in X^*, known as "weak* convergence". We shall refer to convergence in terms of the norm as "strong convergence" in order to minimize the possibility of confusion with weak or weak* convergence.

Let X be a normed linear space with dual space X^*.

Definition 7.6: Let $\{x_n\}_{n=1}^\infty$ be a sequence in X and $x \in X$ be given. We say that $\{x_n\}_{n=1}^\infty$ *converges weakly* to x, or that x is a *weak limit* of $\{x_n\}_{n=1}^\infty$ provided that

$$\forall x^* \in X^*, \quad \langle x^*, x_n \rangle \to \langle x^*, x \rangle \text{ as } n \to \infty.$$

In this case we write $x_n \rightharpoonup x$ (weakly) as $n \to \infty$.

Definition 7.7: A sequence $\{x_n\}_{n=1}^\infty$ in X is said to be *weakly convergent* if there exists $x \in X$ such that $x_n \rightharpoonup x$ (weakly) as $n \to \infty$.

Remark 7.8:

(a) A sequence can have at most one weak limit. Indeed, if $x_n \rightharpoonup x$ (weakly) as $n \to \infty$ and $x_n \rightharpoonup y$ (weakly) as $n \to \infty$ then $\langle x^*, x - y \rangle = 0$ for all $x^* \in X^*$, which implies $x - y = 0$. (Here we are making use of Corollary 5.5.)

(b) If $x_n \to x$ (strongly) as $n \to \infty$ then $x_n \rightharpoonup x$ (weakly) as $n \to \infty$. The converse implication is false in general.

(c) If $x_n \rightharpoonup x$ (weakly) as $n \to \infty$ and $\{x_{n_k}\}_{k=1}^\infty$ is a subsequence of $\{x_n\}_{n=1}^\infty$ then $x_{n_k} \rightharpoonup x$ (weakly) as $k \to \infty$.

Remark 7.9: Let X be a finite-dimensional normed linear space and let $\{x_n\}_{n=1}^\infty$ be a sequence in X and $x \in X$ be given. Then $x_n \rightharpoonup x$ (weakly) as $n \to \infty$ if and only if $x_n \to x$ (strongly) as $n \to \infty$. (We shall give a careful proof of this result later. The key observation is the fact that if X is finite dimensional, then the component mappings for any basis are continuous linear functionals.)

In order to develop interesting examples involving weak convergence, we need characterizations of the dual spaces of some infinite-dimensional normed linear spaces. We shall make use of the following result, which will be proved later.

Proposition 7.10 Let $p \in (1, \infty)$ and put

$$q = \frac{p}{p-1}.$$

(Notice that $p^{-1} + q^{-1} = 1$.) Let $x^* \in (l^p)^*$ be given. Then there is exactly one $y \in l^q$ such that

$$\langle x^*, y \rangle = \sum_{k=1}^{\infty} x_k y_k \quad \text{for all } x \in l^p; \tag{7}$$

moreover, $\|y\|_q = \|x^*\|$.

Remark 7.11:

(a) In some sense, Proposition 7.10 says that when $p \in (1, \infty)$ and $q = p(p-1)^{-1}$ then the dual of l^p is l^q. However, some caution is appropriate here, because, strictly speaking, the elements of $(l^p)^*$ and the elements of l^q are different kinds of objects. What the proposition really says is that $(l^p)^*$ is isometrically isomorphic to l^q. This situation is frequently expressed by saying thet the dual of l^p can be identified with l^q through the pairing defined by (7).

(b) It follows immediately from Proposition 7.10 that l^p is reflexive for $1 < p < \infty$.

Remark 7.12: Let $p \in (1, \infty)$ be given and put $q = p(p-1)^{-1}$. Let $\{x^{(n)}\}_{n=1}^{\infty}$ be a sequence in l^p and let $x \in l^p$ be given. Then $x^{(n)} \rightharpoonup x$ (weakly) in l^p as $n \to \infty$ if and only if

$$\sum_{k=1}^{\infty} x_k^{(n)} y_k \to \sum_{k=1}^{\infty} x_k y_k \quad \text{for all } y \in l^q \text{ as } n \to \infty.$$

Example 7.13: Let $p \in (1, \infty)$ be given and put $q = p(p-1)^{-1}$. Consider the sequence $\{e^{(n)}\}_{n=1}^{\infty}$ in l^p, where

$$e_k^{(n)} = \begin{cases} 1 & \text{if } k = n \\ 0 & \text{if } k \neq n. \end{cases}$$

Let $y \in l^q$ be given. Then we have

$$\sum_{k=1}^{\infty} e_k^{(n)} y_k = y_n \quad \text{for all } n \in \mathbb{N}$$

Since $y_n \to 0$ as $n \to \infty$, we conclude that

$$e^{(n)} \rightharpoonup 0 \text{ (weakly) in } l^p \text{ as } n \to \infty.$$

It is clear that the sequence $\{e^{(n)}\}_{n=1}^{\infty}$ does not converge strongly. (Indeed, $\|e^{(n)} - e^{(m)}\|_p = 2^{\frac{1}{p}}$ when $m \neq n$.) Since $\|e^{(n)}\| = 1$ for all $n \in \mathbb{N}$, we see that the norm of a weak limit need not equal the limit of the norms.

Remark 7.14: Although weak convergence generally does not imply strong convergence in infinite dimensions, there do exist infinite-dimensional Banach spaces having the unusual property that weak convergence of sequences is equivalent to strong convergence. An example of such a space is l^1. We shall see later that if $x \in l^1$ and $\{x^{(n)}\}_{n=1}^{\infty}$ is a sequence in l^1 then $x^{(n)} \to x$ (strongly) as $n \to \infty$ if and only if $x^{(n)} \rightharpoonup x$ (weakly) as $n \to \infty$. This fact comes in very handy for constructing interesting examples (and counterexamples).

Theorem 7.15: Let X be a normed linear space and Y be a closed subspace of X. Let $\{x_n\}_{n=1}^{\infty}$ be a sequence in X and $x \in X$ be given. Assume that $x_n \in Y$ for all $n \in \mathbb{N}$ and that $x_n \rightharpoonup x$ (weakly) as $n \to \infty$. Then

(i) $\{x_n\}_{n=1}^{\infty}$ is bounded,

(ii) $x \in Y$,

(iii) $\|x\| \leq \liminf_{n\to\infty} \|x_n\|$.

Proof: To prove (i), consider the sequence $\{J(x_n)\}_{n=1}^{\infty}$ in X^{**}. For every $x^* \in X^*$ the sequence $\{(J(x_n))(x^*)\}_{n=1}^{\infty}$ is bounded. The Principle of Uniform Boundedness implies that the sequence $\{\|J(x_n)\|\}_{n=1}^{\infty}$ is bounded. (Notice that X^* is complete.) Since $\|x_n\| = \|J(x_n)\|$ for all $n \in \mathbb{N}$, the sequence $\{x_n\}_{n=1}^{\infty}$ is bounded.

To prove (ii), suppose that $x \notin Y$. Then, by Corollary 5.8, we may choose $x^* \in X^*$ such that $\langle x^*, y\rangle = 0$ for all $y \in Y$ and $\langle x^*, x\rangle = 1$. We have $\langle x^*, x_n\rangle = 0$ for all $n \in \mathbb{N}$. This is a contradiction because $\langle x^*, x_n\rangle \to \langle x^*, x\rangle = 1$ as $n \to \infty$. It follows that $x \in Y$.

To prove (iii), let $x^* \in X^*$ be given. Then we have

$$|\langle x^*, x\rangle| = \lim_{n\to\infty} |\langle x^*, x_n\rangle| \leq \liminf_{n\to\infty} \|x^*\|\|x_n\|. \tag{8}$$

Recall that

$$\|x\| = \sup\{|\langle x^*, x\rangle| : x^* \in X^*, \|x^*\| \leq 1\}. \tag{9}$$

Combining (8) and (9) we arrive at

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

Remark 7.16: Using a separation result for convex sets, we shall show later (via a modification of the proof of (ii) above) that if K is closed and convex, $x_n \in K$ for all $n \in \mathbb{N}$ and $x_n \rightharpoonup x$ (weakly) as $n \to \infty$ then $x \in K$. We shall also show that if $F : X \to \mathbb{R}$ is a continuous convex function and $x_n \rightharpoonup x$ (weakly) as $n \to \infty$ then

$$\liminf_{n\to\infty} F(x_n) \geq F(x).$$

These observations are of crucial importance in the calculus of variations.

Since the dual space X^* is a normed linear space in its own right, the notion of weak convergence in the dual space is meaningful. However, it is more natural to use the notion of "weak* convergence" in X^*. Unfortunately, some authors refer to "weak* convergence" as "weak convergence in the dual space". Caution is required regarding this point.

Definition 7.17: Let $\{x_n^*\}_{n=1}^\infty$ be a sequence in X^* and $x^* \in X^*$ be given. We say that $\{x_n^*\}_{n=1}^\infty$ *converges weakly** to x^*, or that x^* is a *weak* limit* of $\{x_n^*\}_{n=1}^\infty$ provided that

$$\forall x \in X, \quad \langle x_n^*, x\rangle \to \langle x^*, x\rangle \text{ as } n \to \infty.$$

In this case we write $x_n^* \overset{*}{\rightharpoonup} x^*$ (weakly*) as $n \to \infty$.

Definition 7.18: A sequence $\{x_n^*\}_{n=1}^\infty$ in X^* is said to be *weakly* convergent* if there exists $x^* \in X^*$ such that $x_n^* \overset{*}{\rightharpoonup} x^*$ (weakly*) as $n \to \infty$.

Remark 7.19: Notice that if X is reflexive then weak* convergence is the same thing as weak convergence in X^*. However most of the time when weak* convergence is used in practice, X is not reflexive. Since $J[X] \subset X^{**}$, it follows that weak convergence in X^* impplies weak* convergence.

Remark 7.20:

(a) A sequence can have at most one weak* limit. Indeed, if $x_n^* \overset{*}{\rightharpoonup} x^*$ (weakly*) as $n \to \infty$ and $x_n^* \overset{*}{\rightharpoonup} y^*$ (weakly*) as $n \to \infty$ then $\langle x^* - y^*, x\rangle = 0$ for all $x \in X$, which implies $x^* - y^* = 0$.

(b) If $x_n^* \to x^*$ (strongly) as $n \to \infty$ then $x_n \overset{*}{\rightharpoonup} x$ (weakly*) as $n \to \infty$. The converse implication is false in general.

(c) If $x_n^* \overset{*}{\rightharpoonup} x^*$ (weakly*) as $n \to \infty$ and $\{x_{n_k}^*\}_{k=1}^\infty$ is a subsequence of $\{x_n\}_{n=1}^\infty$ then $x_{n_k}^* \overset{*}{\rightharpoonup} x^*$ (weakly*) as $k \to \infty$.

(d) If $x_n^* \rightharpoonup x^*$ (weakly) as $n \to \infty$ then $x_n^* \overset{*}{\rightharpoonup} x^*$ (weakly*) as $n \to \infty$. The converse implication is false in general. However, when working in a dual space X^* of a nonreflexive space we shall almost always use weak* convergence rather than weak convergence.

In order to construct interesting examples involving weak* convergence, we need the dual space of a nonreflexive Banach space. We shall make use of the following result which will be proved later.

Proposition 7.21: Let $x^* \in (l^1)^*$ be given. Then there is exactly one $x \in l^\infty$ such that

$$x^*(z) = \sum_{k=1}^{\infty} x_k z_k \quad \text{for all } z \in l^1; \tag{10}$$

moreover $\|x^*\| = \|x\|_\infty$.

Remark 7.22: It follows from Proposition 7.21 that the dual of l^1 can be identified with l^∞ through the isometric isomorphism given in (10). Even though l^∞ is not strictly speaking a dual space, it is standard to talk about weak* convergence in l^∞ and to say that a sequence $\{x^{(n)}\}_{n=1}^\infty$ in l^∞ converges weakly* to $x \in l^\infty$ provided that

$$\sum_{k=1}^{\infty} x_k^{(n)} z_k \to \sum_{k=1}^{\infty} x_k z_k \quad \text{for all } z \in l^1 \text{ as } n \to \infty.$$

Example 7.23:

(a) Consider the sequence $\{e^{(n)}\}_{n=1}^\infty$ in l^∞, where

$$e_k^{(n)} = \begin{cases} 1 & \text{if } k = n \\ 0 & \text{if } k \neq n. \end{cases}$$

Let $z \in l^1$ be given. Then we have

$$\sum_{n=1}^{\infty} e_k^{(n)} z_k = z_n \to 0 \quad \text{as } n \to \infty.$$

We conclude that

$$e^{(n)} \overset{*}{\rightharpoonup} 0 \text{ (weakly}*) \text{ in } l^\infty \text{ as } n \to \infty.$$

(b) Define $x \in l^\infty$ by $x_k = 1$ for all $k \in \mathbb{N}$, and consider the sequence $\{x^{(n)}\}_{n=1}^\infty$ in l^∞ given by

$$x_k^{(n)} = \begin{cases} 1 & \text{if } k \leq n \\ 0 & \text{if } k > n. \end{cases}$$

Observe that $x^{(n)} \in c_0$ for all $n \in \mathbb{N}$. Let $z \in l^1$ be given. Then we have

$$\sum_{k=1}^{\infty} x_k^{(n)} z_k = \sum_{k=1}^{n} z_k \to \sum_{k=1}^{\infty} z_k = \sum_{k=1}^{\infty} x_k z_k \quad \text{as } n \to \infty.$$

We conclude that

$$x^{(n)} \overset{*}{\rightharpoonup} x \text{ (weakly}*\text{) in } l^\infty \text{ as } n \to \infty.$$

The convergence cannot be weak because c_0 is a closed subspace of l^∞ and $x \notin c_0$.

Theorem 7.24: Let X be a Banach space with dual space X^*. Let $\{x_n^*\}_{n=1}^\infty$ be a sequence in X^* and let $x^* \in X^*$ be given. Assume that $x_n^* \overset{*}{\rightharpoonup} x^*$ (weakly*) as $n \to \infty$. Then

(i) $\{x_n^*\}_{n=1}^\infty$ is bounded,

(ii) $\|x^*\| \leq \liminf_{n\to\infty} \|x_n^*\|$.

Remark 7.25:

(a) Completeness of X is essential in order to obtain the boundedness of weakly* convergent sequences. (We need to use the Principle of Uniform Boundedness.)

(b) Weak* limits can "escape" from closed subspaces of X^*. More precisely, it can happen that Z is a closed subspace of X^*, $x_n^* \in Z$ for all $n \in \mathbb{N}$, $x_n^* \rightharpoonup x^*$ (weakly*) as $n \to \infty$, but $x^* \notin Z$. (See Example 7.23(b).)

Proof of Theorem 7.24: For every $x \in X$, the sequence $\{\langle x_n^*, x\rangle\}_{n=1}^\infty$ is bounded. Since X is complete, the Principle of Uniform Boundedness implies that the sequence $\{\|x_n^*\|\}_{n=1}^\infty$ is bounded.

To prove (ii), let $x \in X$ be given. Then we have

$$|\langle x^*, x\rangle| = \lim_{n\to\infty} |\langle x_n^*, x\rangle| \leq \liminf_{n\to\infty} \|x_n^*\|\|x\|. \tag{11}$$

Recall that

$$\|x^*\| = \sup\{|\langle x^*, x\rangle| : x \in X, \|x\| \leq 1\}. \tag{12}$$

Combining (11) and (12) we arrive at

$$\|x^*\| \leq \liminf_{n\to\infty} \|x_n^*\|. \quad \square$$

Theorem 7.26: Let X be a NLS. If X^* is separable then X is separable.

Remark 7.27: The converse implication is false as can be seen from Proposition 7.21.

Proof: Assume that X^* is separable. Choose a dense sequence $\{x_n^*\}_{n=1}^\infty$ in X^*. For each $n \in \mathbb{N}$ choose $x_n \in X$ such that $\|x_n\| = 1$ and

$$|\langle x_n^*, x_n \rangle| \geq \frac{1}{2}\|x_n^*\| \quad \text{for all } n \in \mathbb{N}. \tag{13}$$

Let S be the set of all linear combinations of the x_n with rational coefficients. (By a rational element of $\mathbb{C}$ we mean a number of the form $a + ib$ with $a, b \in \mathbb{Q}$. Let $Y = cl(S)$ and observe that Y is a subspace of X. Suppose that $Y \neq X$. Then we may choose $x^* \in X^*$ such that $\langle x^*, y \rangle = 0$ for all $y \in Y$ and $x^* \neq 0$. Since $\{x_n^*\}_{n=1}^\infty$ is dense in X^* we may choose $\tau : \mathbb{N} \to \mathbb{N}$ such that

$$x_{\tau(n)}^* \to x^* \quad \text{as } n \to \infty.$$

Then we have

$$\|x^* - x_{\tau(n)}^*\| \geq |\langle x^* - x_{\tau(n)}^*, x_{\tau(n)} \rangle| = |\langle x_{\tau(n)}^*, x_{\tau(n)} \rangle| \geq \frac{\|x_{\tau(n)}^*\|}{2} \quad \text{for all } n \in \mathbb{N}$$

by virtue of (13) and the fact that $\langle x^*, x_{\tau(n)} \rangle = 0$ since $x_{\tau(n)}^* \in Y$. Consequently $x_{\tau(n)}^* \to 0$ as $n \to \infty$. This is a contradiction because $x^* \neq 0$. We conclude that $Y = X$. $\square$

Lecture Notes for Week 8 (First Draft)

Weak and Weak Convergence (Continued)*

Theorem 8.1: Let X be a reflexive Banach space and $\{x_n\}_{n=1}^{\infty}$ be a bounded sequence in X. Then $\{x_n\}_{n=1}^{\infty}$ has a weakly convergent subsequence.

Proof: Let $Z = \mathrm{cl}(\mathrm{span}(\{x_n : n \in \mathbb{N}\}))$ and observe that Z is closed and separable. It follows from Theorem 7.1 that Z is reflexive. Since Z^{**} is isomorphic to Z and Z is separable, we conclude that Z^{**} is separable. It follows from Theorem 7.4 that Z^* is separable.

Choose a dense sequence $\{z_n^*\}_{n=1}^{\infty}$ in Z^*. We shall construct a convergent subsequence of $\{x_n\}_{n=1}^{\infty}$ by a diagonalization argument. For consistency in indexing, let us put $x_{n,0} = x_n$ for all $n \in \mathbb{N}$. Then the sequence $\{z_1^*(x_{n,0})\}_{n=1}^{\infty}$ is bounded in $\mathbb{K}$, so we may choose a subsequence $\{x_{n,1}\}_{n=1}^{\infty}$ of $\{x_{n,0}\}_{n=1}^{\infty}$ such that $\{z_1^*(x_{n,1})\}_{n=1}^{\infty}$ is convergent. Then $\{z_2^*(x_{n,1})\}_{n=1}^{\infty}$ is a bounded sequence in $\mathbb{K}$, so we may choose a subsequence $\{x_{n,2}\}_{n=1}^{\infty}$ of $\{x_{n,1}\}_{n=1}^{\infty}$ such that $\{z_2^*(x_{n,2})\}_{n=1}^{\infty}$ is convergent. Proceeding by induction, we obtain for each $k \in \mathbb{N}$ a sequence $\{x_{n,k}\}_{n=1}^{\infty}$ such that

- $\{x_{k+1,n}\}_{n=1}^{\infty}$ is a subsequence of $\{x_{n,k}\}_{n=1}^{\infty}$,
- $\{z_k^*(x_{n,k})\}_{n=1}^{\infty}$ is convergent.

Put $y_n = x_{n,n}$ for all $n \in \mathbb{N}$ and observe that $\{y_n\}_{n=1}^{\infty}$ is a subsequence of $\{x_n\}_{n=1}^{\infty}$. Observe further that for each $k \in \mathbb{N}$ the sequence $\{z_k^*(y_n)\}_{n=1}^{\infty}$ is convergent. Since the sequence $\{y_n\}_{n=1}^{\infty}$ is bounded in Z, and the sequence $\{z_n^*\}_{n=1}^{\infty}$ is dense in Z^*, we conclude that for every $z^* \in Z^*$, the sequence $\{z^*(y_n)\}_{n=1}^{\infty}$ is convergent. A few details have been omitted here, but they can easily be filled in by writing

$$z^*(y_n) - z^*(y_m) = z^*(y_n) - z_k^*(y_n) + z_k^*(y_n) - z_k^*(y_m) + z_k^*(y_m) - z^*(y_m).$$

For each $z^* \in Z^*$, define

$$y^{**}(z^*) = \lim_{n\to\infty} z^*(y_n)$$

and observe that $y^{**} \in Z^{**}$. Since Z is reflexive, we may choose $y \in Z$ such that $J_Z(y) = y^{**}$. Then we have

$$\lim_{n\to\infty} z^*(y_n) = z^*(y) \text{ for all } z^* \in Z^*.$$

Let $x^* \in X^*$ be given and put $z^*(x) = x^*(x)$ for all $x \in Z$. It follows that

$$\langle x^*, y_n\rangle = \langle z^*, y_n\rangle \to \langle z^*, y\rangle = \langle x^*, y\rangle \text{ as } n \to \infty.$$

We conclude that $y_n \rightharpoonup y$ (weakly) as $n \to \infty$. $\square$

Definition 8.2: A sequence $\{x_n\}_{n=1}^{\infty}$ in a NLS X is said to be a *weak Cauchy sequence* provided that for every $x^* \in X^*$, $\{\langle x^*, x_n \rangle\}_{n=1}^{\infty}$ is a Cauchy sequence in $\mathbb{K}$.

Definition 8.3: A NLS is said to be *weakly sequentially complete* provided that every weak Cauchy sequence is weakly convergent.

Remark 8.4: Since $\mathbb{K}$ is complete, a sequence $\{x_n\}_{n=1}^{\infty}$ in a NLS X is a weak Cauchy sequence if and only if the sequence $\{\langle x^*, x_n \rangle\}_{n=1}^{\infty}$ is convergent for each $x^* \in X^*$.

Theorem 8.5: Let X be a reflexive Banach space. Then X is weakly sequentially complete.

Proof: Let $\{x_n\}_{n=1}^{\infty}$ be a weak Cauchy sequence. Then for each $x^* \in X^*$, $\{(J(x_n))(x^*)\}_{n=1}^{\infty}$ is bounded. By the Principle of Uniform Boundedness, $\{\|J(x_n)\|\}_{n=1}^{\infty}$ is bounded, and consequently $\{x_n\}_{n=1}^{\infty}$ is bounded. By Theorem 8.1, we may choose $x \in X$ and a subsequence $\{x_{n_k}\}_{k=1}^{\infty}$ of $\{x_n\}_{n=1}^{\infty}$ such that $x_{n_k} \rightharpoonup x$ (weakly) as $k \to \infty$. Let $x^* \in X^*$ be given. Since $\{\langle x^*, x_n \rangle\}_{n=1}^{\infty}$ is convergent (by completeness of $\mathbb{K}$) and $\langle x^*, x_{n_k} \rangle \to \langle x^*, x \rangle$ as $k \to \infty$, we conclude that

$$\langle x^*, x_n \rangle \to \langle x^*, x \rangle \ \text{ as } n \to \infty,$$

and $\{x_n\}_{n=1}^{\infty}$ is weakly convergent. $\square$

Theorem 8.6 Let X be a separable NLS and $\{x_n^*\}_{n=1}^{\infty}$ be a bounded sequence in X^*. Then $\{x_n^*\}_{n=1}^{\infty}$ has a weakly* convergent subsequence.

Proof: Choose a dense sequence sequence $\{x_n\}_{n=1}^{\infty}$ in X. We shall again use a diagonalization argument to construct the desired subsequence. For consistency in indexing, let us put $x_{n,0}^* = x_n^*$ for all $n \in \mathbb{N}$. Then the sequence $\{\langle x_{n,0}^*, x_1 \rangle\}_{n=1}^{\infty}$ is bounded in $\mathbb{K}$. We may extract a subsequence $\{x_{n,1}^*\}_{n=1}^{\infty}$ of $\{x_{n,0}^*\}_{n=1}^{\infty}$ such that $\{\langle x_{n,1}^*, x_1 \rangle\}_{n=1}^{\infty}$ is convergent. Since $\{\langle x_{n,1}^*, x_2 \rangle\}_{n=1}^{\infty}$ is bounded, we may extract a subsequence $\{x_{n,2}^*\}_{n=1}^{\infty}$ of $\{x_{n,1}^*\}_{n=1}^{\infty}$ such that $\{\langle x_{n,2}^*, x_2 \rangle\}_{n=1}^{\infty}$ is convergent. Proceeding by induction, we obtain, for each $k \in \mathbb{N}$ a sequence $\{x_{n,k}^*\}_{n=1}^{\infty}$ such that

- $\{x_{n,k+1}^*\}_{n=1}^{\infty}$ is a subsequence of $\{x_{n,k}^*\}_{n=1}^{\infty}$,
- $\{\langle x_{n,k}^*, x_k \rangle\}_{n=1}^{\infty}$ is convergent.

Put $y_n^* = x_{n,n}^*$ for all $n \in \mathbb{N}$. Observe that $\{y_n^*\}_{n=1}^{\infty}$ is a subsequence of $\{x_n^*\}_{n=1}^{\infty}$. Observe further that for each $k \in \mathbb{N}$, the sequence $\{\langle y_n^*, x_k \rangle\}_{n=1}^{\infty}$ is convergent. Since $\{y_n^*\}_{n=1}^{\infty}$ is bounded in X^* and $\{x_n\}_{n=1}^{\infty}$ is dense in X we conclude that $\{\langle y_n^*, x \rangle\}_{n=1}^{\infty}$ is convergent for every $x \in X$. Now, define $y^* \in X^*$ by

$$\langle y^*, x \rangle = \lim_{n \to \infty} \langle y_n^*, x \rangle \ \text{ for all } x \in X.$$

It follows that $y_n^* \xrightarrow{*} y^*$ (weakly *) as $n \to \infty$. $\square$

Definition 8.7: A subset S of an NLS X is said to be *total* provided that span(S) is dense in X.

Proposition 8.8: Let X be a NLS, S be a total subset of X^*, $\{x_n\}_{n=1}^{\infty}$ be a sequence in X and $x \in X$ be given. Then $x_n \rightharpoonup x$ (weakly) as $n \to \infty$ if and only if $\{x_n\}_{n=1}^{\infty}$ is bounded and

$$\forall y^* \in S, \quad \langle y^*, x_n \rangle \to \langle y^*, x \rangle \ \text{ as } n \to \infty. \tag{1}$$

Proof: The necessity of the conditions follows from the definition of weak convergence and part (i) of Theorem 7.15. Assume that $\{x_n\}_{n=1}^{\infty}$ is bounded and that (1) holds. Observe that

$$\forall y^* \in \operatorname{span}(S), \quad \langle y^*, x_n \rangle \to \langle y^*, x \rangle \ \text{ as } n \to \infty. \tag{2}$$

Let $x^* \in X^*$ and $\epsilon > 0$ be given. Choose $M > 0$ such that

$$\|x_n - x\| \le M \ \text{ for all } n \in \mathbb{N}. \tag{3}$$

Since span(S) is dense in X^*, we may choose $y^* \in \operatorname{span}(S)$ such that

$$\|x^* - y^*\| < \frac{\epsilon}{2M}. \tag{4}$$

Since (2) holds we may choose $N \in \mathbb{N}$ such that

$$|\langle y^*, x_n - x \rangle| < \frac{\epsilon}{2} \ \text{ for all } n \ge N. \tag{5}$$

Then for all $n \in \mathbb{N}$ with $n \ge N$ we have

$$|\langle x^*, x_n - x \rangle| \le |\langle x^* - y^*, x_n - x \rangle| + |\langle y^*, x_n - x \rangle| < \frac{\epsilon}{2} + \frac{\epsilon}{2},$$

by virtue of (3), (4), and (5). $\square$

Proposition 8.9: Let X be a Banach space, S be a total subset of X, $\{x_n^*\}_{n=1}^{\infty}$ be a sequence in X^* and $x^* \in X^*$ be given. Then $x_n^* \xrightarrow{*} x^*$ (weakly*) as $n \to \infty$ if and only if $\{x_n^*\}_{n=1}^{\infty}$ is bounded and

$$\forall y \in S, \quad \langle x_n^*, y \rangle \to \langle x^*, y \rangle \ \text{ as } n \to \infty. \tag{6}$$

The proof of Proposition 8.9 is very similar to the proof of Proposition 8.8 and will be omitted.

For the sake of completeness, we formulate Remark 7.9 as a propositon and provide a proof.

Proposition 8.10: Let X be a finite-dimensional NLS, $\{x^{(n)}\}_{n=1}^{\infty}$ be a sequence in X and let $x \in X$ be given. Then $x^{(n)} \rightharpoonup x$ (weakly) as $n \to \infty$ if and only if $x^{(n)} \to x$ (strongly) as $n \to \infty$.

Remark 8.11: Since every finite-dimensional NLS is reflexive, it follows from Proposition 8.10 that a sequence in the dual of a finite-dimensional NLS is weakly* convergent if and only if it is strongly convergent.

Proof of Proposition 8.10: We already know that strong convergence implies weak convergence in any NLS. Assume that $x^{(n)} \rightharpoonup x$ (weakly) as $n \to \infty$. Choose a basis $(b_j | j = 1, 2, \cdots, N)$ for X and define $b_k^* \in X^*$ for $k = 1, 2, \cdots, N$ by

$$b_k^*(b_j) = \begin{cases} 1 & \text{if } j = k \\ \\ 0 & \text{if } j \neq k. \end{cases}$$

($(b_j^* | j = 1, 2, \cdots, N)$ is called the dual basis for $(b_j | j = 1, 2, \cdots, N)$.)

Choose $\alpha_j^{(n)}, \ \alpha_j, \, n \in \mathbb{N}, j = 1, 2, \cdots, N$ such that

$$x^{(n)} = \sum_{j=1}^{N} \alpha_j^{(n)} b_j, \quad x = \sum_{j=1}^{N} \alpha_j b_j.$$

Since $\langle b_k^*, x^{(n)} \rangle \to \langle b_k^*, x \rangle$ as $n \to \infty$ for $k = 1, 2, \cdots, N$, we conclude that

$$\alpha_k^{(n)} \to \alpha_k \ \text{ as } n \to \infty \ \text{ for } k = 1, 2, \cdots, N.$$

It follows that

$$\| x^{(n)} - x \| \leq \sum_{k=1}^{N} | \alpha_k^{(n)} - \alpha_k | \cdot \| b_k \| \to 0 \ \text{ as } n \to \infty. \ \square$$

Weak convergence is an extremely important tool in applications. In many situations, an existence theorem is proved by first constructing a sequence of "approximate solutions" and showing that the sequence is bounded in a reflexive Banach space. Then a weakly convergent subsequence can be extracted. The final step is to show that the weak limit of the subsequence is actually a solution of the original problem. A key tool in this process is often the following theorem which says that every closed convex sets is sequentially weakly closed (i.e. contains the limit of each weakly convergent sequence of points from the set).

Theorem 8.12: Let X be a NLS and K be a closed convex subset of X. Let $\{x_n\}_{n=1}^{\infty}$ be a sequence in X and $x \in X$ be given. Assume that $x_n \in K$ for all $n \in \mathbb{N}$ and that $x_n \rightharpoonup x$ (weakly) as $n \to \infty$. Then $x \in K$.

The proof of Theorem 8.12 is based on the following separation lemma for convex sets.

Lemma 8.13: Let X be a NLS, K be a nonempty closed convex subset of X, and $x_0 \in X \backslash K$ be given. Then there exists $x^* \in X^*$ such that

$$\mathrm{Re}(x^*(x_0)) < \inf\{\mathrm{Re}(x^*(y)) : y \in K\}.$$

Proof of Lemma 8.13: Without loss of generality, we may assume that $x_0 = 0$. Since K is closed and $0 \notin K$ we may choose $\delta > 0$ such that $B_\delta(0) \cap K = \emptyset$. By Theorem 6.6, we may choose $x^* \in X^*$ such that $x^* \neq 0$ and

$$\mathrm{Re}(x^*(x)) \leq \mathrm{Re}(x^*(y)) \quad \text{for all } x \in B_\delta(0), y \in K. \tag{7}$$

Let us put

$$\alpha = \inf\{\mathrm{Re}(x^*(y)) : y \in K\}.$$

Then we have

$$\mathrm{Re}(x^*(x)) \leq \alpha \leq \mathrm{Re}(x^*(y)) \quad \text{for all } x \in B_\delta(0), y \in K. \tag{8}$$

Since $\mathrm{Re}(x^*(0)) = 0$, it remains only to show that $\alpha > 0$.

Since the range of a nonzero linear functional is all of $\mathbb{K}$, we may choose $z \in X$ such that $\mathrm{Re}(x^*(z)) > 0$. Let

$$w = \frac{\delta z}{2\|z\|},$$

and notice that $w \in B_\delta(0)$ and $\mathrm{Re}(x^*(w)) > 0$. We conclude from the first inequality in (8) that $\alpha > 0$ and the proof is complete. $\square$

Proof of Theorem 8.12: Suppose that $x \notin K$. Then, by Lemma 8.13, we may choose $x^* \in X^*$ such that

$$\mathrm{Re}(x^*(x)) < \inf\{\mathrm{Re}(x^*(x_n) : n \in \mathbb{N}\} \quad \text{for all } n \in \mathbb{N}. \tag{9}$$

On the other hand, we know that

$$\lim_{n\to\infty} x^*(x_n) = x^*(x) \quad \text{as } n \to \infty, \tag{10}$$

because $x_n \rightharpoonup x$ (weakly) as $n \to \infty$. Equation (10) is not possible because of (9). We conclude that $x \in K$. $\square$

Lemma 8.14: Let X be a reflexive Banach space and K be a nonempty closed convex subset of X. Then there exists $\hat{x} \in K$ such that

$$\|\hat{x}\| \leq \|x\| \quad \text{for all } x \in K.$$

In view of Problem 9 on Assignment 4, we have the following interesting consequence of Lemma 8.14.

Corollary 8.15: $C[0,1]$ (equipped with the maximum norm) is not reflexive.

Proof of Lemma 8.14: Let

$$\gamma = \inf\{\|x\| : x \in K\}.$$

Choose a sequence $\{x_n\}_{n=1}^{\infty}$ such that $x_n \in K$ for all $n \in \mathbb{N}$ and

$$\|x_n\| \to \gamma \quad \text{as } n \to \infty. \tag{11}$$

It follows from (11) that $\{x_n\}_{n=1}^{\infty}$ is bounded. By Theorem 8.1, we may choose $\hat{x} \in X$ and a subsequence $\{x_{n_k}\}_{n=1}^{\infty}$ such that

$$x_{n_k} \rightharpoonup \hat{x} \ \text{ (weakly) as } k \to \infty.$$

Theorem 8.12 tells us that $\hat{x} \in K$. By part (iii) of Theorem 7.15, we have

$$\|\hat{x}\| \le \liminf_{k\to\infty} \|x_{n_k}\| = \gamma \le \|x\| \ \text{ for all } x \in K. \ \square$$

The proof of Lemma 8.14 is a simple instance of the *Direct Method of the Calculus of Variations*. The calculus of variations is concerned with minimizing (or maximizing) functionals (typically defined by integrals) on subsets of infinite-dimensional NLS. The subject has a rich history beginning with Johann Bernoulli's announcement of the *brachsitochrone* problem in 1696. Early emphasis was placed on finding necessary conditions for minima. In the 20^{th} century, starting with the work of Tonelli, it was discovered that methods of functional analysis can be used to prove existence theorems by a direct method. This method makes crucial use of weak convergence and results about convex sets and convex functionals. You will have some homework exercises concerning this topic. The method can be summarized as follows: To prove the existence of a minimizer for a functional $G : K \to \mathbb{R}$, where K is a subset of a reflexive Banach space one carries out the steps below.

- Step 1: Put $\gamma = \inf\{G(y) : y \in K\}$.
- Step 2: Choose a *minimizing sequence*, i.e. a sequence $\{y_n\}_{n=1}^{\infty}$ such that $y_n \in K$ for all $n \in \mathbb{N}$ and $G(y_n) \to \gamma$ as $n \to \infty$.
- Step 3: Show that $\{y_n\}_{n=1}^{\infty}$ is bounded.
- Step 4: Extract a weakly convergent subsequence $\{y_{n_k}\}_{k=1}^{\infty}$ and let y denote the weak limit of this subsequence.
- Step 5: Show that $y \in K$. (This will be automatic if K is closed and convex.)
- Step 6: Show that $\liminf_{k\to\infty} G(y_{n_k}) \ge G(y)$, which implies that $G(y) = \gamma$. (This will be automatic if G is convex and lower semicontinuous in the norm topology.)

As another application of Theorem 8.12, we mention the following result.

Lemma 8.16: Let X be a reflexive Banach space and let $\{K_n\}_{n=1}^{\infty}$ be a sequence of nonempty, closed, bounded, convex subsets of X such that $K_{n+1} \subset K_n$ for all $n \in \mathbb{N}$. Then

$$\bigcap_{n=1}^{\infty} K_n \neq \emptyset.$$

The proof of this lemma is left as an exercise.

Weak and Weak Convergence in Sequence Spaces*

Before giving some simple criteria for weak and weak* convergence in sequence spaces, it is useful to state a result identifying the dual of c_0 with l^1

Proposition 8.17: Let $x^* \in (c_0)^*$ be given. Then there exists exactly one $y \in l^1$ such that

$$x^*(x) = \sum_{k=1}^{\infty} y_k x_k \quad \text{for all } x \in c_0; \tag{12}$$

moreover, $\|y\|_1 = \|x^*\|$.

The proof of Proposition 8.17 will be a homework exercise.

Remark 8.18: It follows from from Proposition 8.17 that the dual space of c_0 can be identified with l^1 through the isometric isomorphism given in (12). Even though l^1 is not strictly speaking a dual space, it is standard to talk about weak* convergence in l^1 and to say that a sequence $\{x^{(n)}\}_{n=1}^{\infty}$ in l^1 converges weakly* to $x \in l^1$ provided that

$$\forall z \in c_0, \quad \sum_{k=1}^{\infty} x_k^{(n)} z_k \to \sum_{k=1}^{\infty} x_k z_k \quad \text{as } n \to \infty.$$

The following simple definition is useful for stating very simple criteria for weak and weak* convergence in sequence spaces.

Definition 8.19: Let $x \in \mathbb{K}^{\mathbb{N}}$ be given and let $\{x^{(n)}\}_{n=1}^{\infty}$ be a sequence of elements of $\mathbb{K}^{\mathbb{N}}$. We say that $x^{(n)} \to x$ *componentwise* or *pointwise* as $n \to \infty$ provided that

$$\forall k \in \mathbb{N}, \text{ we have } x_k^{(n)} \to x_k \text{ as } n \to \infty.$$

Remark 8.20: Consider the sequence $\{e^{(n)}\}_{n=1}^{\infty}$ of elements of $\mathbb{K}^{\mathbb{N}}$ defined by

$$e_k^{(n)} = \begin{cases} 1 & \text{for } k = n \\ \\ 0 & \text{for } k \neq n. \end{cases}$$

(a) The set $\{e^{(m)} : m \in \mathbb{N}\}$ is total in l^p for $1 \le p < \infty$, but it is not total in l^∞.

(b) The set $\{e^{(m)} : m \in \mathbb{N}\}$ is total in c_0, but it is not total in c.

Remark 8.21: Let $\{x^{(n)}\}_{n=1}^\infty$ be a sequence in $\mathbb{K}^\mathbb{N}$ and let $x \in \mathbb{K}^\mathbb{N}$ be given and notice that for every $m \in \mathbb{N}$ we have

$$\sum_{k=1}^\infty x_k^{(n)} e_k^{(m)} = x_m^{(n)}, \quad \sum_{k=1}^\infty x_k e_k^{(m)} = x_m,$$

and consequently $x^{(n)} \to x$ componentwise as $n \to \infty$ if and only if

$$\forall m \in \mathbb{N}, \quad \sum_{k=1}^\infty x_k^{(n)} e_k^{(m)} \to \sum_{k=1}^\infty x_k e_k^{(m)} \quad \text{as } n \to \infty.$$

We can use Definition 8.19 and Remarks 8.20, 8.21 together with Propositions 8.8 and 8.9 to give simple and useful criteria for weak and weak* convergence in sequence spaces.

Proposition 8.22: Let X be one of the spaces l^p, $1 \le p \le \infty$ or c_0 and let $\{x^{(n)}\}_{n=1}^\infty$ be a sequence of elements of X and let $x \in X$ be given.

(a) If $X = c_0$ then $x^{(n)} \rightharpoonup x$ (weakly) as $n \to \infty$ if and only if $\{x^{(n)}\}_{n=1}^\infty$ is bounded in X and $x^{(n)} \to x$ componentwise as $n \to \infty$.

(b) If $X = l^p$ and $1 < p < \infty$ then $x^{(n)} \rightharpoonup x$ (weakly) as $n \to \infty$ if and only if $\{x^{(n)}\}_{n=1}^\infty$ is bounded in X and $x^{(n)} \to x$ componentwise as $n \to \infty$.

(c) If $X = l^1$ or $X = l^\infty$ then $x^{(n)} \overset{*}{\rightharpoonup} x$ as $n \to \infty$ if and only if $\{x^{(n)}\}_{n=1}^\infty$ is bounded in X and $x^{(n)} \to x$ componentwise as $n \to \infty$.

The space l^1 has a very unusual property, namely that it is an infinite-dimensional Banach space in which every weakly convergent sequence is strongly convergent. (Consequently, we cannot extend Proposition 8.22(b) to the case $p = 1$.)

Proposition 8.22: Let $\{x^{(n)}\}_{n=1}^\infty$ be a sequence in l^1 and $x \in l^1$ be given. Then $x^{(n)} \rightharpoonup x$ (weakly) as $n \to \infty$ if and only if $x^{(n)} \to x$ (strongly) as $n \to \infty$.

The proof of Proposition 8.22 will be a homework exercise.

Lecture Notes for Week 9 (Preliminary Draft)

The Dual of l^p, $1 \le p < \infty$

Let $p \in [1, \infty),\ q \in (1, \infty]$ be given and assume that

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Given $y \in l^q$, put

$$L_y(x) = \sum_{k=1}^{\infty} x_k y_k \ \text{ for all } x \in l^p. \tag{1}$$

By Holder's inequality, we have

$$|L_y(x)| \le \|y\|_q \|x\|_p \ \text{ for all } x \in l^p. \tag{2}$$

It follows that $L_y \in (l^p)^*$. Define $T : l^q \to (l^p)^*$ by

$$(Ty)(x) = L_y(x) \ \text{ for all } y \in l^q, x \in l^p. \tag{3}$$

Theorem 9.1: For each $p \in [1, \infty)$, the mapping T defined above is an isometric isomorphism of l^q onto $(l^p)^*$.

Proof: Let $p \in [1, \infty)$ be given and put $q = p(p-1)^{-1}$. We shall employ the standard Scahuder basis $\{e^{(n)}\}_{n=1}^{\infty}$ for l^p, where

$$e_k^{(n)} = \begin{cases} 1 & \text{for } k = n \\ 0 & \text{for } k \ne n. \end{cases}$$

Let $z \in l^q$ be given. We know that

$$\|Tz\|_* \le \|z\|_q, \tag{4}$$

by virtue of (2) and (3), and consequently

$$\|T\| \le 1. \tag{5}$$

Observe that T is injective because if $Tz = 0$ then, for each $n \in \mathbb{N}$, we have

$$0 = (Tz)e^{(n)} = \sum_{k=1}^{\infty} e_k^{(n)} z_k = z_n.$$

It remains to show that T is surjective and that $\|Ty\|_* \ge \|y\|_q$ for all $y \in l^q$.

Let $f \in (l^p)^*$ be given. We want to find $y \in l^q$ such $Ty = f$. To this end put

$$y_k = f(e^{(k)}) \text{ for all } k \in \mathbb{N}. \tag{6}$$

Since $\{e^{(k)}\}_{k=1}^{\infty}$ is a Schauder basis for l^p, we have

$$x = \sum_{k=1}^{\infty} x_k e^{(k)} \text{ for all } x \in l^p. \tag{7}$$

Since f is continuous and the series in (7) converges strongly to x we conclude that

$$f(x) = \sum_{k=1}^{\infty} x_k L(e^{(k)}) = \sum_{k=1}^{\infty} x_k y_k \text{ for all } x \in l^p.$$

To complete the proof, it is sufficient to show that $y \in l^q$ and $\|y\|_q \leq \|f\|_*$. This will imply that T is surjective (hence bijective) and that $\|Ty\|_* \geq \|y\|_q$.

Suppose first that $p = 1$ (so that $q = \infty$). It follows from (2) that

$$|y_k| \leq \|f\|_* \|e^{(k)}\| = \|f\|_* \text{ for all } k \in \mathbb{N},$$

and this tells us that $y \in l^\infty$ and $\|y\|_\infty \leq \|f\|_*$.

Suppose now that $p > 1$. If $y = 0$ we are done, so we may assume that $y \neq 0$. For every $n \in \mathbb{N}$ define $x^{(n)}$ by

$$x_k^{(n)} = \begin{cases} |y_k|^{q-1}\mathrm{sgn}(y_k) & \text{for } k \leq n \\ 0 & \text{for } k > 0. \end{cases} \tag{8}$$

Here, for $\lambda \in \mathbb{K}$,

$$sgn(\lambda) = \begin{cases} \frac{|\lambda|}{\lambda} & \text{for } \lambda \neq 0 \\ 0 & \text{for } \lambda = 0. \end{cases}$$

Observe that

$$\sum_{k=1}^{n} |y_k|^q = f(x^{(n)}) \leq \|f\|_* \|x^{(n)}\|_p \text{ for all } n \in \mathbb{N}. \tag{9}$$

Observe further that

$$\|x^{(n)}\|_p = \left(\sum_{k=1}^{n} |y_k|^q \right)^{1/p} \text{ for all } n \in \mathbb{N}, \tag{10}$$

since $(q-1)p = q$.

For n sufficiently large, using (9), (10), and the fact that $1 - \frac{1}{p} = \frac{1}{q}$, we obtain

$$\left(\sum_{k=1}^{n} |y_k|^q\right)^{1/q} \leq \|f\|_*. \tag{11}$$

Letting $n \to \infty$ in (11) we find that $y \in l^q$ and that

$$\|Ty\|_* \geq \|y\|_q. \quad \square$$

The Dual of c_0

Given $y \in l^1$, put

$$L_y(x) = \sum_{k=1}^{\infty} x_y y_k \text{ for all } x \in c_0.$$

Then we have

$$|L_y(x)| \leq \|y\|_1 \|x\|_\infty \text{ for all} x \in c_0,$$

so that $L_y \in (c_0)^*$. Now define $T : l^1 \to (c_0)^*$ by

$$Ty = L_y \text{ for all } y \in l^1.$$

Theorem 9.2: The mapping T defined above is an isometric isomorphism of l^1 onto $(c_0)^*$.

The proof is very similar to (and simpler than) the proof of Theorem 9.1 and is left as an exercise.

Remark 9.3: It follows from Theorems 9.1 and 9.2 that l^p is reflexive for $1 < p < \infty$ and that the spaces l^1, l^∞, c_0, and c are all nonreflexive. (We haven't said anything yet about the dual of c. However, if c were reflexive then c_0 would have to be reflexive since it is a closed subspace of c_0.)

$L^P - Spaces$

Let Ω be a nonempty open subset of $\mathbb{R}^n$. For $1 \leq p < \infty$ we denote by $L^p(\Omega)$ the set of all (equivalence classes of) measurable functions $f : \Omega \to \mathbb{K}$ such that

$$\int_\Omega |f|^p < \infty.$$

Here, by *measurable* we mean *Lebesgue measurable* and two measurable functions $f_1, f_2 : \Omega \to \mathbb{K}$ belong to the same equivalence class provided that $\mu(\{x \in \Omega : f_1(x) \neq$

$f_2(x)\}) = 0$. Here, $\mu(A)$ denotes the Lebesgue measure of a set A. We follow the standard practice of using the same symbol to denote a measurable function and its equivalence class.

We define $\|\cdot\|_p : L^p(\Omega) \to [0, \infty)$ by

$$\|f\|_p = \left(\int_\Omega |f|^p\right)^{\frac{1}{p}} \quad \text{for all } f \in L^p(\Omega).$$

It is evident that $\|f\|_p = 0$ implies that $f = 0$ *a.e.* and that $\|\alpha f\|_p = |\alpha|\|f\|_p$ for all $\alpha \in \mathbb{K},\ f \in L^p(\Omega)$. The triangle inequality can be verified easily when $p = 1$. It also holds for $p \in (1, \infty)$ as well, but the proof is not completely trivial.

We denote by $L^\infty(\Omega)$ the set of (equivalence classes of) essentially bounded measurable functions $f : \Omega \to \mathbb{K}$. We define $\|\cdot\|_\infty : L^\infty(\Omega) \to [0, \infty)$ by

$$\|f\|_\infty = \operatorname*{ess-sup}_\Omega(|f|) = \inf\{M \in \mathbb{R} : \mu(\{x \in \Omega : |f(x)| > m\}) = 0\}.$$

We make the same conventions concerning equivalence classes of measurable functions as we do for $L^p(\Omega)$ with $p \in [1, \infty)$.

Proposition 9.4 (Minkowski's Inequality): Let $p \in [1, \infty]$ be given. If $f, g \in L^p(\Omega)$ then so is $f + g$ and

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

It follows that $(L^p(\Omega), \|\cdot\|_p)$ is a normed linear space.

Proposition 9.5 (Holder's Inequality): Let $p, q \in [1, \infty]$ with

$$\frac{1}{p} + \frac{1}{q} = 1$$

and $f \in L^p(\Omega)$, $g \in L^q(\Omega)$ be given. Then $fg \in L^1(\Omega)$ and

$$\int_\Omega |fg| \leq \|f\|_p \|g\|_q.$$

Remark 9.6: Assume that $\mu(\Omega) < \infty$ and that $1 \leq p_1 \leq p_2 \leq \infty$. Then $L^{p_2}(\Omega) \subset L^{p_2}(\Omega)$ and there exists $K \in \mathbb{R}$ such that

$$\|u\|_{p_1} \leq \|u\|_{p_2} \quad \text{for all } u \in L^{p_2}(\Omega).$$

Although the idea is very simple (and almost certainly familiar to everyone in the class), we shall give a proof of Remark 9.6 using a simple "trick" involving Holder's inequality. Assume first that $1 \leq p_1 \leq p_2 < \infty$ and let $u \in L^{p_2}(\Omega)$ be given. Then we have

$$\int_\Omega |u|^{p_1} = \int_\Omega (|u|^{p_2})^{\frac{p_1}{p_2}} \cdot 1.$$

Using Holder' Inequality with $p = \frac{p_2}{p_1}$, $q = \frac{p_2}{p_2 - p_1}$, $f = |u|^{p_1}$, and $g = 1$, we find that

$$\|u\|_{p_1} \leq (\mu(\Omega))^{\frac{1}{p_1} - \frac{1}{p_2}} \|u\|_{p_2}.$$

If $u \in L^\infty(\Omega)$ and $p_1 < p_2$, then

$$\int_\Omega |u|^{p_1} \leq \int_\Omega \|u\|_\infty^{p_1} \leq \|u\|_\infty^{p_1} \mu(\Omega),$$

and consequently

$$\|u\|_{p_1} \leq (\mu(\Omega))^{\frac{1}{p_1}}.$$

Theorem 9.7 (Riesz-Fisher): Assume that $1 \leq p \leq \infty$¿ Then $L^p(\Omega)$ is complete.

Theorem 9.8 (Riesz-Representation): Let $p \in [1, \infty)$, $q \in (1, \infty]$ be given and assume that

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Let $u^* \in (L^p(\Omega))^*$ be given. Then there is exactly one $g \in L^q(\Omega)$ such that

$$u^*(f) = \int_\Omega fg;$$

moreover, $\|u^*\| = \|g\|_q$.

Corollary 9.9: $L^p(\Omega)$ is reflexive for $1 < p < \infty$.

Proposition 9.10: $L^p(\Omega)$ is separable for $1 \leq p < \infty$. $L^\infty(\Omega)$ is not separable.

Remark 9.11: It follows from Theorem and Proposition that neither $L^1(\Omega)$ nor $L^\infty(\Omega)$ is reflexive.

Let us denote by $C_c^\infty(\Omega)$ the set of all functions $u : \Omega \to \mathbb{K}$ having continuous derivatives of all orders and such that $\operatorname{spt}(u)$ is a compact subset of Ω. (Here $\operatorname{spt}(u)$ is the closure of $\{x \in \Omega : u(x) \neq 0\}$.)

Lemma 9.12: $C_c^\infty(\Omega)$ is dense in $L^p(\Omega)$ for $1 \leq p < \infty$.

Lemma 9.13 (Clarkson's Inequalities): Let $p, q \in (0, \infty)$ with

$$\frac{1}{p} + \frac{1}{q} = 1$$

and $f, g \in L^p(\Omega)$ be given. Then

(a) for $1 < p \leq 2$ we have

$$\|\frac{f+g}{2}\|_p^q + \|\frac{f-g}{2}\|_p^q \leq \left(\frac{1}{2}\|f\|_p^p + \frac{1}{2}\|g\|_p^p\right)^{q-1},$$

$$\|\frac{f+g}{2}\|_p^p + \|\frac{f-g}{2}\|_p^p \geq \frac{1}{2}\|f\|_p^p + \frac{1}{2}\|g\|_P^P.$$

(b) for $2 \le p < \infty$ we have

$$\|\frac{f+g}{2}\|_p^p + \|\frac{f-g}{2}\|_p^p \le \frac{1}{2}\|f\|_p^p + \frac{1}{2}\|g\|_P^P,$$

$$\|\frac{f+g}{2}\|_p^q + \|\frac{f-g}{2}\|_p^q \ge \left(\frac{1}{2}\|f\|_p^p + \frac{1}{2}\|g\|_p^p\right)^{q-1}.$$

Remark 9.14: An extremely important consequence of Clarkson's Inequalities is that they imply that $L^p(\Omega)$ is *uniformly convex* when $1 < p < \infty$. (We shall study uniformly convex spaces later. They have the useful property that weak convergence, together with convergence of the sequence of norms to the norm of the weak limit, imply strong convergence. They have other important properties as well.)

Theorem 9.15 (Riesz-Kolmogorov): Assume that $1 \le p < \infty$ and let K be a bounded subset of $L^p(\Omega)$. Then K is precompact (i.e. has compact closure) if and only if for every $\epsilon > 0$ there exists a compact set $G \subset \Omega$ and $\delta > 0$ such that (i) and (ii) below hold:

(ii)
$$\int_{\Omega\backslash G} |u|p < \epsilon^p \quad \text{for all } u \in K,$$

(ii)
$$\int_{\Omega} |\tilde{u}(x+h) - \tilde{u}(x)|^p \, dx < \epsilon^p$$

for all $h \in \mathbb{R}^n$ with $|h| < \delta$ and all $u \in K$, where

$$\tilde{u}(x) = \begin{cases} u(x) & \text{for } x \in \Omega \\ 0 & \text{for } x \in \mathbb{R}^n\backslash\Omega. \end{cases}$$

Adjoints

Let X and Y be NLS and $T \in \mathcal{L}(X;Y)$ be given. We shall construct a bounded linear operator $T^* \in \mathcal{L}(Y^*;X^*)$ that is associated with T in a natural way. The operator T^* is called the adjoint of T. Much important information about the operator T can be obtained by studying the adjoint T^*.

If X and Y are finite-dimensional, then any linear operator $T : X \to Y$ can be represented by a matrix. (Any linear operator from Y^* to X^* can also be represented by a matrix). The matrices representing the linear transformations will, of course, depend on a choice of bases for the linear spaces. If the bases are chosen suitably, then the matrix representing T^* will be the transpose of the matrix representing T. Many of the important relations between matrices and their transposes can be generalized to linear operators and their adjoints. In situations involving infinite-dimensional spaces, we will need to make certain adjustments and additional assumptions that are not necessary in finite-dimensional spaces. We begin with a theorem that facilitates the definition of an adjoint. Since certain formulas will involve duality pairings for different spaces, we sometimes include subscripts on the pairings. More precisely, for any NLS Z, the duality pairing from $Z^* \times Z$ to $\mathbb{K}$ will be denoted by $\langle \cdot, \cdot \rangle_{Z^* \times Z}$ in situations where it may help to keep track of where certain object live.

Theorem 9.16: Let X and Y be NLS and let $T \in \mathcal{L}(X;Y)$ be given. Then there is exactly one $T^* \in \mathcal{L}(Y^*;X^*)$ such that

$$\langle T^*y^*, x\rangle_{X^* \times X} = \langle y^*, Tx\rangle_{Y^* \times Y} \quad \text{for all } y^* \in Y^*, x \in X. \tag{12}$$

Moreover $\|T^*\| = \|T\|$.

Definition 9.17: The linear operator T^* in Theorem 9.16 is called the *adjoint* of T.

Remark 9.18:

(a) It is clear that the mapping $T \to T^*$ is a linear mapping from $\mathcal{L}(X;Y)$ to $\mathcal{L}(Y^*;X^*)$, i.e. the mapping that carries an operator to its adjoint is linear.

(b) When there is no danger of confusion, we shall drop the subscripts on the duality pairings in (12).

Proof of Theorem 9.16: For each $y^* \in Y^*$, define

$$T^*y^* = y^* \circ T, \tag{13}$$

where $\circ$ indicates composition of mappings (so that (13) means $(T^*y^*)(x) = y^*(Tx)$ for all $x \in X$). For fixed y^*, T^*y^* is a mapping from X to $\mathbb{K}$. Since the composite of two continuous linear mappings is continuous and linear, it follows that $T^*y^* \in X^*$ for every $y^* \in Y^*$.

Equation (13) clearly implies (12). Let $y^* \in Y^*$ and $x^* \in X^*$ be given. Assume that (12) holds for all $x \in X$ and that

$$\langle x^*, x\rangle_{X^* \times X} = \langle y^*, Tx\rangle_{Y^* \times Y} \quad \text{for all } x \in X.$$

Then

$$\langle x^* - T^*y^*, x\rangle_{X^* \times X} = 0 \quad \text{for all } x \in X,$$

and we conclude that $x^* = T^*y^*$, which proves the required uniqueness.

It is clear that the mapping $y^* \to T^*y^*$ is linear. It remains to show that

$$\sup\{\|T^*y^*\| : y^* \in Y^*, \|y^*\| \leq 1\} = \|T\|,$$

which will establish that T^* is bounded and that $\|T^*\| = \|T\|$. Observe that

$$\begin{aligned}
\|T\| &= \sup\{\|Tx\| : x \in X, \|x\| \leq 1\} \\
&= \sup\{|\langle y^*, Tx\rangle_{Y^*\times Y}| : y^* \in Y^*, x \in X, \|y^*\| \leq 1, \|x\| \leq 1\} \\
&= \sup\{|\langle T^*y^*, x\rangle_{X^*\times X}| : y^* \in Y^*, x \in X, \|y^*\| \leq 1, \|x\| \leq 1\} \\
&= \sup\{\|T^*y^*\| : y^* \in Y^*, \|y^*\| \leq 1\}. \quad \square
\end{aligned}$$

Remark 9.19: Let X be a NLS and let I_X and I_{X^*} denote the identity mappings on X and X^*, respectively. Observe that for all $x \in X, x^* \in X^*$ we have

$$\langle I_{X^*}(x^*), x\rangle = \langle x^*, x\rangle = \langle x^*, I_X(x)\rangle,$$

and consequently

$$(I_X)^* = I_{X^*}.$$

Example 9.20 (The Finite-Dimensional Case) TO BE FILLED IN.

Theorem 9.21: Let X, Y, Z be NLS and let $T \in \mathcal{L}(X;Y)$ and $S \in \mathcal{L}(Y;Z)$ be given. Then $(ST)^* = T^*S^*$.

Proof: Let $x \in X, z^* \in Z^*$ be given. Then we have

$$\begin{aligned}
\langle (ST)^*z^*, x\rangle_{X^*\times X} &= \langle z^*, (ST)x\rangle_{Z^*\times Z} = \langle z^*, S(Tx)\rangle_{Z^*\times Z} \\
&= \langle S^*z^*, Tx\rangle_{Y^*\times Y} = \langle T^*(S^*z^*), x\rangle_{X^*\times X} \\
&= \langle (T^*S^*)z^*, x\rangle_{X^*\times X}.
\end{aligned}$$

We conclude that $(ST)^*z^* = T^*S^*z^*$ for all $z^* \in Z^*$, i.e. $(ST)^* = T^*S^*$. $\square$

Second Adjoints

Let X, Y be NLS and let $T \in \mathcal{L}(X;Y)$ be given. Since $T^* \in \mathcal{L}(Y^*;X^*)$, we can talk about the adjoint T^{**} of T^*. Observe that $T^{**} \in \mathcal{L}(X^{**};Y^{**})$. For matrices, the transpose of the transpose is the same as the original matrix. We want to understand if there is an analogue of this result for adjoints in the general setting. Of course, T^{**} and T are different kinds of objects – the inputs for T^{**} are elements of the second dual of X and the inputs for T are elements of X. We can (and will) identify X with a linear manifold $\hat{X}$ in X^{**}. If X is not reflexive, then $\hat{X} \subsetneq X^{**}$, so that T^{**} cannot be identified with T; however T^{**} can always be identified with an extension of T.

Let

$$\hat{X} = J_X[X], \quad \hat{Y} = J_Y[Y],$$

and define $\hat{T} \in \mathcal{L}(\hat{X}; \hat{Y})$ by

$$\hat{T}\hat{x} = J_Y(T((J_X)^{-1}\hat{x})) \text{ for all } \hat{x} \in \hat{X}.$$

Given a linear manifold $\mathcal{D}(S) \subset X^{**}$, and a linear mapping $S : \mathcal{D}(S) \to Y^{**}$ we say that S is an *extension* of T provided that

$$\mathcal{D}(S) \supset \hat{X} \text{ and } S\hat{x} = \hat{T}\hat{x} \text{ for all } \hat{x} \in \hat{X}.$$

If S is an extension of T and $\mathcal{D}(S) = \hat{X}$ then we write $S = T$. (What we are really doing here is identifying T with $\hat{T}$.)

Lecture Notes for Week 10 (First Draft)

Second Adjoints (Continued)

Theorem 10.1: Let X and Y be normed linear spaces and let $T \in \mathcal{L}(X;Y)$ be given. Then T^{**} is an extension of T. If X is reflexive then $T^{**} = T$.

Proof: What we need to show is that

$$T^{**}(J_X(x)) = J_Y(Tx) \quad \text{for all } x \in X. \tag{1}$$

Let $x \in X$ be given. Then, for all $y^* \in Y^*$, we have

$$\begin{aligned} \langle T^{**}(J_X(x)), y^* \rangle_{Y^{**}\times Y^*} &= \langle J_X(x), T^* y^* \rangle_{X^{**}\times X^*} = \langle T^* y^*, x \rangle_{X^* \times X} \\ &= \langle y^*, Tx \rangle_{Y^* \times Y} = \langle J_Y(Tx), y^* \rangle_{Y^{**}\times Y^*}. \end{aligned}$$

We conclude that (1) holds. $\square$

Adjoints, Null Spaces, Ranges, and Inverses

A square matrix is invertible if and only if the transpose is invertible. In this case the transpose of the inverse is the inverse of the transpose. We now give a generalization of this result to operators between NLS.

Theorem 10.2: Let X be a Banach space, Y be a normed linear space, and $T \in \mathcal{L}(X;Y)$ be given. Then the following two statements are equivalent:

(i) T is bijective and $T^{-1} \in \mathcal{L}(Y;X)$,

(ii) T^* is bijective and $(T^*)^{-1} \in \mathcal{L}(X^*;Y^*)$.

If (i) (and hence also (ii)) holds then $(T^*)^{-1} = (T^{-1})^*$.

Remark 10.3: Completeness of X is not needed for the implication (i) $\Rightarrow$ (ii) in Theorem 10.2.

Proof of Theorem 10.2: Assume that (i) holds. Then we have

$$TT^{-1} = I_Y, \quad T^{-1}T = I_X.$$

Applying Theorem 9.21 (concerning adjoints of products) and Remark 9.19 (concerning adjoints of identity operators), we obtain

$$(T^{-1})^* T^* = (I_Y)^* = I_{Y^*}, \tag{2}$$

$$T^*(T^{-1})^* = (I_X)^* = I_{X^*}. \tag{3}$$

We conclude that T^* is bijective and that $(T^*)^{-1} = (T^{-1})^*$.

Assume that (ii) holds. Then, by what we just proved, T^{**} is bijective and $T^{**} \in \mathcal{L}(Y^{**}; X^{**})$. Since T^{**} is an extension of T, we conclude that T is injective. It remains to show that T is surjective. This is where we use the completeness of X. Let $\hat{X} = J_X[X]$. Since X is complete, we know that $\hat{X}$ is a closed subspace of X^{**}. Since T^{**} is an isomorphism from X^{**} to Y^{**}, we conclude that $T^{**}[\hat{X}]$ is a closed subspace of Y^{**}. Since T^{**} is an extension of T, we conclude that $T[X]$ is a closed subspace of Y.

Suppose that $T[X] \neq Y$. Then, by a corollary to the Hahn-Banach Theorem, we may choose $y^* \in Y^*$ such that $y^* \neq 0$ and $y^*(y) = 0$ for all $y \in T[X]$. This means that $y^*(Tx) = 0$ for all $x \in X$. It follows that $T^*y^* = 0$. Since T^* is injective, we infer that $y^* = 0$, which is impossible. We conclude that T is surjective, and hence bijective. Boundedness of T^{-1} follows from (3) and the boundedness of $(T^{**})^{-1}$. $\square$

Another very useful feature of transposes of matrices is the "orthogonality relationship" between the range of a matrix and the null space of the transpose. This relationship can be generalized to linear operators between infinite-dimensional NLS, but the situation becomes more complicated in infinite dimensions.

Definition 10.4: Let X be a NLS and let $A \subset X$, $B \subset X^*$ be given. The *annihilator* of A is the subset $A^\perp$ of X^* defined by

$$A^\perp = \{x^* \in X^* : x^*(x) = 0 \text{ forall } x \in A\}.$$

The *pre-annihilator* of B is the subset ${}^\perp B$ of X defined by

$${}^\perp B = \{x \in X : x^*(x) = 0 \text{ for all } x^* \in B\}.$$

Remark 10.5: Many authors call both sets "annihilators". I think that it is a good idea to use a different term because a dual space of a NLS is a NLS in its own right and the meaning of the "annihilator of a subset of X^*" could become genuinely ambiguous.

The following remark concerning annihilators and pre-annihilators is easily verified.

Remark 10.6: Let X be a normed linear space and $A \subset X$, $B \subset Y$ be given. Then

(i) $A^\perp$ is a closed subspace of X^*.

(ii) ${}^\perp B$ is a closed subspace of X.

(iii) $A \subset^\perp (A^\perp)$

(iv) $B \subset ({}^{\perp}B)^{\perp}$

It is natural to ask when we might have $A = {}^{\perp}(A^{\perp})$ for a set $A \subset X$. By item (ii) of the remark above, there is no hope for this inequality unless A is a closed subspace.

In fact, we have the following simple, but useful, result.

Proposition 10.7: Let X be a NLS and $A \subset X$. Then ${}^{\perp}(A^{\perp}) = \mathrm{cl}(\mathrm{span}(A))$.

The proof of this Proposition 10.7 is part of Assignment 5.

Remark 10.8: It is possible to have a closed subspace Z of X^* such that $Z \subsetneq ({}^{\perp}Z)^{\perp}$. You are asked to give an example in Assignment 6. What is actually true about this situation is that for any set $B \subset X^*$, $({}^{\perp}B)^{\perp}$ is the weak* closure of span(B).

Theorem 10.9: Let X, Y be NLS and let $T \in \mathcal{L}(X;Y)$ be given. Then we have

(i) $\mathcal{N}(T^*) = (\mathcal{R}(T))^{\perp}$, and

(ii) $\mathcal{N}(T) = {}^{\perp}(\mathcal{R}(T^*))$.

Proof: To prove (i), let $y^* \in Y^*$ be given and observe that

$$\begin{aligned} y^* \in \mathcal{N}(T^*) &\Leftrightarrow T^*y^* = 0 \\ &\Leftrightarrow (T^*y^*)(x) = 0 \text{ for all } x \in X \\ &\Leftrightarrow y^*(Tx) = 0 \text{ for all } x \in X \\ &\Leftrightarrow y^* \in (\mathcal{R}(T))^{\perp}. \end{aligned}$$

To prove (ii), let $x \in X$ be given and observe that

$$\begin{aligned} x \in \mathcal{N}(T) &\Leftrightarrow Tx = 0 \\ &\Leftrightarrow y^*(Tx) = 0 \text{ for all } y^* \in Y^* \\ &\Leftrightarrow (T^*y^*)(x) = 0 \text{ for all } y^* \in Y^* \\ &\Leftrightarrow x \in {}^{\perp}(\mathcal{R}(T^*)). \quad \square \end{aligned}$$

We now investigate the relationship between $\mathcal{R}(T)$ and ${}^{\perp}(\mathcal{N}(T^*))$. In view of Theorem 10.9, it might be tempting to conjecture that these two linear manifolds are the same. However, since the null space of a continuous linear operator is closed, and the range of a continuous linear operator need not be closed, we can see that the linear manifolds in question cannot be equal in general. However, we have the following result.

Theorem 10.10: Let X, Y be normed linear spaces and let $T \in \mathcal{L}(X;Y)$ be given. Then

$$\mathrm{cl}(\mathcal{R}(T)) =^{\perp} (\mathcal{N}(T^*)).$$

The proof follows immediately from Theorem 10.9 (i) and Proposition 10.7.

Finally, we consider the relationship between $\mathcal{R}(T^*)$ and $(\mathcal{N}(T))^{\perp}$. It might be tempting to conjecture that $\mathrm{cl}(\mathcal{R}(T^*))$ is equal to $(\mathcal{N}(T))^{\perp}$; however this is not true in general because, even for a closed subspace Z of X^*, it can happen that the annihilator of the pre-annihilator of Z is strictly larger than Z – and this phenomenon does occur sometimes for closures of ranges of adjoints. However, we do have the following result.

Theorem 10.11: Let X, Y be Banach spaces, let $T \in \mathcal{L}(X;Y)$ be given, and assume that $\mathcal{R}(T)$ is closed. Then $\mathcal{R}(T^*)$ is closed and

$$\mathcal{R}(T^*) = (\mathcal{N}(T))^{\perp}.$$

To prove Theorem 10.11, we shall make use of the following simple lemma, which is based on the open mapping theorem.

Lemma 10.12: Let X, Y be Banach spaces, let $T \in \mathcal{L}(X;Y)$ be given, and assume that $\mathcal{R}(T)$ is closed. Then there exists $K \in \mathbb{R}$ with the following property: For every $y \in \mathcal{R}(T)$, there exists $x \in X$ such that $y = Tx$ and $\|x\| \leq K\|y\|$.

Proof: Since Y is a Banach space and $\mathcal{R}(T)$ is closed, $(\mathcal{R}, \|\cdot\|_Y)$ is a Banach space. Obviously, $T : (X, \|\cdot\|_X) \to (\mathcal{R}(T), \|\cdot\|_Y)$ is linear, continuous, and surjective. By the open mapping theorem, we may choose $\delta > 0$ such that

$$T[B_1^X(0)] \supset \{y \in \mathcal{R}(T) : \|y\| < \delta\}.$$

Let $y \in \mathcal{R}(T)$ be given. If $y = 0$, there is nothing to prove, so we assume that $y \neq 0$. Put

$$z = \frac{\delta y}{2\|y\|}$$

and notice that $z \in \mathcal{R}(T)$ and $\|z\| < \delta$. We may therefore choose $w \in B_1^X(0)$ such that $Tw = z$. Now put

$$x = \frac{2\|y\|}{\delta} w.$$

Observe that

$$Tx = y \quad \text{and} \quad \|x\| \leq \frac{2}{\delta}\|y\|. \quad \square$$

Proof of Theorem 10.11: Let $x^* \in \mathcal{N}(T)^{\perp}$ be given. Notice that if $x_1, x_2 \in X$ satisfy $Tx_1 = Tx_2$ then $x^*(x_1) = x^*(x_2)$. Consequently, we may choose a linear functional $g : \mathcal{R}(T) \to \mathbb{K}$ such such that

$$g(Tx) = x^*(x) \text{ for all } x \in X.$$

Choose K as in Lemma 10.12. Let $y \in \mathcal{R}(T)$ be given. Then we may choose $x \in X$ such that $y = Tx$ and $\|x\| \leq K\|y\|$. It follows that

$$|g(y)| \leq K\|x^*\|\|y\|.$$

By the Hahn-Banach Theorem, we may choose $y^* \in Y^*$ such that

$$y^*(y) = g(y) \text{ for all } y \in \mathcal{R}(T).$$

Then, for all $x \in X$ we have

$$(T^*y^*)(x) = y^*(Tx) = x^*(x),$$

which says that $x^* = T^*y^*$, i.e. $x^* \in \mathcal{R}(T^*)$. We conclude that

$$\mathcal{N}(T)^\perp \subset \mathcal{R}(T^*).$$

Applying the annihilator to the expression in Theorem 10.9 (ii) and using the fact that $({}^\perp B)^\perp \supset B$ for every $B \in X^*$ we find that

$$\mathcal{R}(T^*) \subset \mathcal{N}(T)^\perp.$$

It follows that $\mathcal{R}(T^*) = \mathcal{N}(T)^\perp$. Since $\mathcal{N}(T)^\perp$ is closed, we conclude that $\mathcal{R}(T^*)$ is closed. $\square$

Lemma 10.13: Let X, Y be Banach spaces and $S \in \mathcal{L}(X;Y)$ be given. Assume that there exists $c > 0$ such that

$$\|S^*y^*\| \geq c\|y^*\| \text{ for all } y^* \in Y^*.$$

Then S is surjective.

The proof of Lemma 10.13 is part of Assignment 5.

Lemma 10.14: Let X, Y be Banach spaces and let $T \in \mathcal{L}(X;Y)$ be given. Assume that $\mathcal{R}(T^*)$ is closed. Then $\mathcal{R}(T)$ is closed.

Proof: Let $Z = \text{cl}(\mathcal{R}(T))$ and notice that $(Z, \|\cdot\|_Y)$ is complete. Define $S \in \mathcal{L}(X;Z)$ by

$$Sx = Tx \text{ for all } x \in X.$$

Observe that $\mathcal{R}(S)$ is dense in Z and consequently $(\mathcal{R}(S))^\perp = \{0\}$, because a continuous linear functional that vanishes on a dense set, must be the zero functional. (Here the $^\perp$ operator is the one that takes a set in Z and produces a set in Z^*.) It follows from Theorem 10.9 (i) that

$$\mathcal{N}(S^*) = (\mathcal{R}(S))^\perp = \{0\},$$

and S^* is injective.

Let $z^* \in Z^*$ be given. Choose an extension $y^* \in Y^*$ of z^*. Then, for all $x \in X$ we have

$$(T^* y^*)(x) = y^*(Tx) = z^*(Sx) = (S^* z^*)(x). \tag{4}$$

We conclude that $T^* y^* = S^* z^*$, which tells us that $\mathcal{R}(S^*) \subset \mathcal{R}(T^*)$. Similarly, if we start with $y^* \in Y^*$ and let z^* denote the restriction of y^* to Z, then (4) holds for all $x \in X$ which implies that $\mathcal{R}(T^*) \subset \mathcal{R}(S^*)$, and consequently we have

$$\mathcal{R}(S^*) = \mathcal{R}(T^*).$$

This implies that $\mathcal{R}(S^*)$ is closed. Since S^* is injective and has closed range, the Bounded Inverse Theorem implies the existence of a constant $c > 0$ such that

$$\|S^* z^*\| \geq c\|z^*\| \ \text{ for all } z^* \in Z^*.$$

Lemma 10.13 implies that S is surjective. This tells us that $\mathcal{R}(T) = Z$, and consequently $\mathcal{R}(T)$ is closed. $\square$

The following theorem summarizes several key results that we have established concerning ranges of operators and adjoints.

Theorem 10.15: Let X, Y be Banach spaces and let $T \in \mathcal{L}(X;Y)$ be given. The following four statements are equivalent

(i) $\mathcal{R}(T)$ is closed,

(ii) $\mathcal{R}(T^*)$ is closed,

(iii) $\mathcal{R}(T) = {}^\perp \mathcal{N}(T^*)$,

(iv) $\mathcal{R}(T^*) = \mathcal{N}(T)^\perp$.

Remark 10.16: The implications "(i) $\Leftrightarrow$ (iii)" and "(iv) $\Rightarrow$ (ii) in Theorem 10.15 do not require completeness.

The right and left shift operators can be used to illustrate many important features of adjoints. For simplicity, we look at these operators as acting from l^2 to l^2. (The results can easily be generalized to other l^p spaces.

Example 10.17: Let $X = Y = l^2$. We shall identify $(l^2)^*$ with l^2 as usual. Define $R, L : l^2 \to l^2$ as follows:

$$Rx = (0, x_1, x_2, x_3, \cdots), \quad Lx = (x_2, x_3, x_4, x_5, \cdots) \ \text{ for all } x \in l^2.$$

Observe that R is an isometry (and hence is injective), but is not surjective; L is surjective, but fails to be injective. Since

$$\sum_{k=1}^{\infty}(Rx)_k y_k = \sum_{j=1}^{\infty} x_j (Ly)_j = x_1y_2 + x_2y_3 + x_3y_4 + \cdots \quad \text{for all } x, y \in l^2,$$

conclude that

$$R^* = L, \quad L^* = R.$$

If we think of the elements of l^2 as column vectors, we can associate R and L with infinite matrices $\mathbf{R}$ and $\mathbf{L}$:

$$\mathbf{R} = \begin{pmatrix} 0 & 0 & 0 & \cdot & \cdot & \cdot \\ 1 & 0 & 0 & \cdot & \cdot & \cdot \\ 0 & 1 & 0 & \cdot & \cdot & \cdot \\ 0 & 0 & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix} \quad \mathbf{L} = \begin{pmatrix} 0 & 1 & 0 & \cdot & \cdot & \cdot \\ 0 & 0 & 1 & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix}.$$

Notice that $\mathbf{R}^{\mathrm{T}} = \mathbf{L}$ and $\mathbf{L}^{\mathrm{T}} = \mathbf{R}$.

Example 10.18: Let $X = Y = L^2[0,1]$ and let $k \in C([0,1] \times [0,1])$ be given. As usual, we identify $(L^2[0,1])^*$ with $L^2[0,1]$. For every $f \in L^2[0,1]$, put

$$(Kf)(x) = \int_0^1 k(x,y) f(y)\,dy, \quad x \in [0,1].$$

It is straightforward to show that $K \in \mathcal{L}(L^2[0,1]; L^2[0,1])$. It is also straightforward to show that the adjoint K^* of K is given by

$$(K^*g)(x) = \int_0^1 k(y,x) g(y)\,dy.$$

Linear Mappings, Weak and Weak Convergence*

Let X, Y be normed linear spaces and let $T \in \mathcal{L}(X;Y)$ and $L \in \mathcal{L}(Y^*;X^*)$ be given.

Question 1: Let $\{x_n\}_{n=1}^{\infty}$ be a sequence in X and $x \in X$ be given. Assume that $x_n \rightharpoonup x$ (weakly) as $n \to \infty$. Can we conclude that $Tx_n \rightharpoonup Tx$ (weakly) in Y as $n \to \infty$?

This question can be answered affirmatively by using the adjoint T^*. Let $y^* \in Y^*$ be given. Then we have

$$y^*(Tx_n) = (T^*y^*)(x_n) \to (T^*y^*)(x) = y^*(Tx) \quad \text{as } n \to \infty.$$

We conclude that $Tx_n \rightharpoonup Tx$ (weakly) in Y as $n \to \infty$.

Question 2: Let $\{y_n^*\}_{n=1}^{\infty}$ be a sequence in in Y^* and $y^* \in Y^*$ be given. Assume that $y_n^* \overset{*}{\rightharpoonup} y^*$ (weakly*) in Y^* as $n \to \infty$. Can we conclude that $Ly_n^* \overset{*}{\rightharpoonup} Ly^*$ (weakly*) in X^* as $n \to \infty$?

If $L = U^*$ for some $U \in \mathcal{L}(X;Y)$ then we have

$$(Ly_n^*)(x) = y_n^*(Ux) \to y^*(Ux) = (Ly^*)(x) \quad \text{as } n \to \infty,$$

and consequently $Ly_n^* \overset{*}{\rightharpoonup} Ly^*$ (weakly*) in X^* as $n \to \infty$. However, if L is not the adjoint of some member of $\mathcal{L}(X;Y)$ then the argument given above cannot be used. In fact, bounded linear operators that are not adjoints need not respect weak* convergence. We will see an example at the beginning of class next time.

Lecture Notes for Week 11 (First Draft)

Linear Operators, Weak and Weak Convergence (Continued)*

Last time, it was stated that bounded linear operators that are not adjoints of some other bounded linear operator need not respect weak* convergence. We begin with an example of this phenomenon.

Example 11.1: Let $X = \mathbb{K}$ and $Y = c_0$. We identify X^* with $\mathbb{K}$ and Y^* with l^1. Define $L : l^1 \to \mathbb{K}$ by

$$L(w) = \sum_{n=1}^{\infty} w_n \text{ for all } w \in l^1.$$

Clearly, $L : l^1 \to \mathbb{K}$ is linear and continuous. Let us put $w^{(n)} = (-1)^n e^{(n)}$ for all $n \in \mathbb{N}$. Then we have $w^{(n)} \overset{*}{\rightharpoonup} 0$ (weakly*) as $n \to \infty$, but $L(w^{(n)}) = (-1)^n$ for all $n \in \mathbb{N}$, so that the sequence $\{L^{(n)}\}_{n=1}^{\infty}$ fails to be weakly* convergent in $\mathbb{K}$.

Theorem 11.2: Let X and Y be NLS and let $T : X \to Y$ be a linear mapping. The following 5 statements are equivalent.

(i) T is continuous.

(ii) For every x in X and every sequence $\{x_n\}_{n=1}^{\infty}$ such that $x_n \rightharpoonup x$ (weakly) as $n \to \infty$ we have $Tx_n \rightharpoonup Tx$ (weakly) as $n \to \infty$.

(iii) For every sequence $\{x_n\}_{n=1}^{\infty}$ in X such that $x_n \rightharpoonup 0$ (weakly) as $n \to \infty$ we have $Tx_n \rightharpoonup 0$ (weakly) as $n \to \infty$.

(iv) For every sequence $\{x_n\}_{n=1}^{\infty}$ in X such that $x_n \rightharpoonup 0$ (weakly) as $n \to \infty$, the sequence $\{Tx_n\}_{n=1}^{\infty}$ is bounded.

(v) For every sequence $\{x_n\}_{n=1}^{\infty}$ in X such that $x_n \to 0$ (strongly) as $n \to \infty$, the sequence $\{Tx_n\}_{n=1}^{\infty}$ is bounded.

Proof: We shall prove (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv) $\Rightarrow$ (v) $\Rightarrow$ (i).

Assume that (i) holds and let $x \in X$ and a sequence $\{x_n\}_{n=1}^{\infty}$ in X such that $x_n \rightharpoonup x$ (weakly) as $n \to \infty$ be given. For each $y^* \in Y^*$ we have

$$\langle y^*, Tx_n \rangle = \langle T^* y^*, x_n \rangle \to \langle T^* y^*, x \rangle = \langle y^*, Tx \rangle \text{ as } n \to \infty,$$

i.e. $Tx_n \rightharpoonup Tx$ (weakly) as $n \to \infty$.

The implications (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv) $\Rightarrow$ (v) are clear.

To prove (v) $\Rightarrow$ (i), we shall prove the contrapositive implication, i.e. (not (i)) $\Rightarrow$ (not (v)). Assume that T is not continuous. We shall construct a sequence $\{x_n\}_{n=1}^{\infty}$ in X such that $x_n \to 0$ (strongly) as $n \to \infty$ and $\{\|Tx_n\|\}_{n=1}^{\infty}$ is unbounded. Since T is unbounded, we may choose $w_n \in X$ such that $\|w_n\| = 1$ and $\|Tw_n\| \geq n^2$ for every $n \in \mathbb{N}$. Now put

$$x_n = \frac{w_n}{n} \quad \text{for all } n \in \mathbb{N}.$$

Then we have $x_n \to 0$ as $n \to \infty$ and $\|Tx_n\| \geq n$ for all $n \in \mathbb{N}$ so that the sequence $\{Tx_n\}_{n=1}^{\infty}$ is unbounded. $\square$

We now formally state the result concerning adjoint operators and weak* convergence.

Proposition 11.3: Let X and Y be NLS and let $T \in \mathcal{L}(X;Y)$ be given. Let $y^* \in Y^*$ and a sequence $\{y_n^*\}_{n=1}^{\infty}$ in Y^* be given and assume that $y_n^* \overset{*}{\rightharpoonup} y^*$ (weakly*) as $n \to \infty$. Then we have $T^*y_n^* \overset{*}{\rightharpoonup} T^*y^*$ (weakly*) as $n \to \infty$.

Proof: For every $x \in X$ we have

$$\langle T^*y_n^*, x\rangle = \langle y_n^*, Tx\rangle \to \langle y^*, Tx\rangle = \langle T^*y^*, x\rangle \quad \text{as } n \to \infty,$$

i.e. $T^*y_n^* \overset{*}{\rightharpoonup} T^*y^*$ (weakly*) as $n \to \infty$. $\square$

Compact Linear Operators

Definition 11.4: Let X and Y be normed linear spaces. A linear mapping $T : X \to Y$ is said to be *compact* provided that $\mathrm{cl}(T[B_1(0)])$ is compact. The set of all compact linear mappings from X to Y will be denoted by $\mathcal{C}(X;Y)$.

Remark 11.5: If a linear operator T is compact, then T is bounded. In other words, $\mathcal{C}(X;Y) \subset \mathcal{L}(X;Y)$.

The next result follows easily from the definitions and the fact that compactness can be characterized by sequences in metric spaces.

Proposition 11.6: Let X and Y be NLS and $T : X \to Y$ be a linear mapping. The following 3 statements are equivalent.

(i) T is compact.

(ii) For every bounded set $A \subset X$, $\mathrm{cl}(T[A])$ is compact.

(iii) For every bounded sequence $\{x_n\}_{n=1}^{\infty}$, the sequence $\{Tx_n\}_{n=1}^{\infty}$ has a convergent subsequence.

Definition 11.7: Let X and Y be linear spaces. A linear mapping $T : X \to Y$ is said to be of *finite rank* provided that $\mathcal{R}(T)$ is finite dimensional.

Proposition 11.8: Let X and Y be normed linear spaces and let $T : X \to Y$ be a linear mapping.

(a) If X is finite dimensional, then T is compact.

(b) If T is continuous and of finite rank, then T is compact.

Remark 11.9: Not every compact linear operator is of finite rank. However, in certain important cases, every compact linear operator can be obtained as a limit (in the operator norm) of a sequence of bounded linear operators of finite rank.

One very pleasant feature of compact linear operators is that they map weakly convergent sequences into strongly convergent ones.

Theorem 11.10: Let X and Y be normed linear spaces, $T \in \mathcal{C}(X;Y)$, $x \in X$ and a sequence $\{x_n\}_{n=1}^{\infty}$ be given. Assume that $x_n \rightharpoonup x$ (weakly) as $n \to \infty$. Then $Tx_n \to Tx$ (strongly) as $n \to \infty$.

Proof: Suppose that $\{Tx_n\}_{n=1}^{\infty}$ fails to converge strongly to Tx. Then we may choose $\epsilon > 0$ and a subsequence $\{x_{n_k}\}_{k=1}^{\infty}$ such that

$$\|Tx_{n_k} - Tx\| \geq \epsilon \ \text{ for all } k \in \mathbb{N}. \tag{1}$$

Since $\{x_{n_k}\}_{k=1}^{\infty}$ is bounded and T is compact, we may choose a subsequence $\{x_{n_{k_j}}\}_{j=1}^{\infty}$ and $y \in Y$ such that

$$Tx_{n_{k_j}} \to y \text{ (strongly) as } j \to \infty.$$

By Theorem 11.2,

$$Tx_{n_{k_j}} \rightharpoonup Tx \text{ (weakly) as } j \to \infty.$$

We conclude that $y = Tx$, and consequently

$$Tx_{n_{k_j}} \to Tx \text{ (strongly) as } j \to \infty,$$

which contradicts (1). It follows that $Tx_n \to Tx$ (strongly) as $n \to \infty$. $\square$

It is natural to ask whether or not a linear operator that maps weakly convergent sequences to strongly convergent ones is necessarily compact. Without some additional assumptions on the spaces, the answer is no since the identity operator $I : l^1 \to l^1$ fails to be compact (because the closed unit ball in l^1 is not compact), but sequences in l^1 are weakly convergent if and only if they are strongly convergent. However, if X is reflexive, then linear operators that map weakly convergent sequences to strongly convergent ones are automatically compact.

Theorem 11.11: Let X be a reflexive Banach space, Y be a normed linear space and $T : X \to Y$ be a linear mapping. Assume that for every weakly convergent sequence $\{x_n\}_{n=1}^{\infty}$ in X, the sequence $\{Tx_n\}_{n=1}^{\infty}$ is strongly convergent. Then T is compact.

Proof: Let a bounded sequence $\{x_n\}_{n=1}^{\infty}$ in X be given. Since X is reflexive, we may choose a weakly convergent subsequence $\{x_{n_k}\}_{k=1}^{\infty}$. Then, by our assumption, $\{Tx_{n_k}\}_{k=1}^{\infty}$ is strongly convergent, so T is compact by Proposition 11.6. $\square$

Proposition 11.12: Let X, Y, Z be normed linear spaces and $T \in \mathcal{L}(X; Y)$, $S \in \mathcal{L}(Y; Z)$ be given. If either T or S is compact then ST is compact.

Proof: Assume that S is compact and let $\{x_n\}_{n=1}^{\infty}$ be a bounded sequence in X. Then $\{Tx_n\}_{n=1}^{\infty}$ is a bounded sequence in Y, so we may choose a subsequence $\{Tx_{n_k}\}_{k=1}^{\infty}$ such that $\{STx_{n_k}\}_{k=1}^{\infty}$ is strongly convergent in Z. The case where T is compact is similar. $\square$

Proposition 11.13: Let X and Y be normed linear spaces and let $T \in \mathcal{C}(X; Y)$ be given. Then $\mathcal{R}(T)$ is separable.

Proof: Let $n \in \mathbb{N}$ be given. Since $\mathrm{cl}(T[B_n(0)])$ is compact, we may choose a finite set $D_n = \{y_{n,k} : k = 1, 2, \cdots N_n\} \subset T[B_n(0)]$ such the collection of balls

$$\{B_{\frac{1}{n}}(y_{n,k}) : k = 1, 2, \cdots, N_n\}$$

covers $T[B_n(0)]$. (Indeed the collection of open sets $\{B_{\frac{1}{n}}(y) : y \in T[B_n(0)]\}$ covers the compact set $\mathrm{cl}(T[B_n(0)])$, so we may choose a finite subcollection that covers $\mathrm{cl}(T[B_n(0)]) \supset T[B_n(0)]$.) Put

$$D = \bigcup_{n=1}^{\infty} D_n,$$

and observe that D is countable. It is clear that D is dense in $\mathcal{R}(T)$, because given $y \in \mathcal{R}(T)$ and $\delta > 0$, we may choose $N > \delta^{-1}$ such $y \in T[B_N(0)]$ so that $B_\delta(y) \cap D_N \neq \emptyset$. $\square$

Theorem 11.14: Let X and Y be normed linear spaces and $T \in \mathcal{L}(X; Y)$ be given. Assume that T is compact. Then T^* is compact.

Proof: Let $B^* = \{x^* \in X^* : \|x^*\| < 1\}$. Since X^* is complete, it suffices to show that $T^*[B^*]$ is totally bounded. Put $B = \{x \in X : \|x\| \leq 1\}$. Since T is compact, we know that $T[B]$ is totally bounded. Let $\epsilon > 0$ be given. We may choose $x_1, x_2, \cdots, x_N \in B$ such that for every $x \in B$, there exists $i \in \{1, 2, \cdots, N\}$ such that

$$\|Tx - Tx_i\| < \frac{\epsilon}{3}. \tag{2}$$

Define $L : Y^* \to \mathbb{K}^N$ by

$$Ly^* = (y^*(Tx_1), y^*(Tx_2), \cdots y^*(Tx_n)) \text{ for all } y^* \in Y^*.$$

For definiteness, we equip $\mathbb{K}^N$ with the maximum norm. Observe that L is continuous and has finite rank. It follows that $L[B^*]$ is totally bounded. Therefore we may choose $y_1^*, y_2^*, \cdots y_m^* \in B^*$ such that for every $y^* \in B^*$, there exists $j \in \{1, 2, \cdots, m\}$ satisfying

$$\|Ly^* - Ly_j^*\| < \epsilon.$$

In other words, for every $y^* \in B^*$, there exists $j \in \{1, 2, \cdots, m\}$ such that

$$|y^*(Tx_i) - y_j^*(T(x_i)| < \frac{\epsilon}{3} \text{ for all } i = 1, 2, \cdots, N. \tag{3}$$

Let $y^* \in B^*$ be given and choose $j \in \{1, 2, \cdots, m\}$ such that (3) holds. Let $x \in B$ be given and choose $i \in \{1, 2, \cdots, N\}$ such that (2) holds. Then we have

$$\begin{aligned}
|y^*(Tx) - y_j^*(Tx)| &\leq |y^*(Tx) - y^*(Tx_i)| + |y^*(Tx_i) - y_j^*(Tx)| \\
&\quad + |y_j^*(Tx_i) - y^*(Tx)| \\
&\leq \|y^*\| \cdot \|Tx - Tx_i\| + \epsilon + \|y_j^*\| \cdot \|Tx_i - Tx\| \\
&\leq \epsilon.
\end{aligned}$$

Since the above chain of inequalities holds for all $x \in X$ with $\|x\| \leq 1$, we conclude that for every in B^*, there exists $j \in \{1, 2, \cdots, m\}$ such that

$$\|T^*y^* - T^*y_j^*\| < \epsilon.$$

It follows that $T^*[B^*]$ is totally bounded. $\square$

Theorem 11.15 Let X be normed linear space, Y be a Banach space, and $T \in \mathcal{L}(X; Y)$. Assume that T^* is compact. Then T is compact.

Proof: Let $\{x_n\}_{n=1}^\infty$ be a bounded sequence in X. Then $\{J_X(x_n)\}_{n=1}^\infty$ is a bounded sequence in X^{**}. By Theorem 11.14, T^{**} is compact, so we may choose a subsequence $\{x_{n_k}\}_{k=1}^\infty$ of $\{x_n\}_{n=1}^\infty$ such that $\{T^{**}(J_X(x_{n_k})\}_{k=1}^\infty$ is strongly convergent in Y^{**}. In particular, $\{T^{**}(J_X(x_{n_k})\}_{k=1}^\infty$ is a Cauchy sequence in Y^{**}. Since J_Y is an isometry and

$$J_Y(T(x)) = T^{**}(J_X(x)) \text{ for all } x \in X,$$

it follows that $\{Tx_{n_k}\}_{k=1}^\infty$ is a Cauchy sequence in Y. Since Y is complete, we conclude that $\{Tx_{n_k}\}_{k=1}^\infty$ is strongly convergent in Y. $\square$

Suppose that X is a reflexive Banach space, Y a normed linear space, and $T \in \mathcal{L}(X;Y)$ has a compact adjoint. Can we conclude that T is compact? Since $T^{**} = T$ in this case, it is natural to say that the answer is yes. However, some caution is advised, because $T^{**} : X^{**} \to Y^{**}$ and $T : X \to Y$. If Y fails to be reflexive, then the image under J_Y of Y will not be all of Y^{**} and the mappings T and T^{**} have different codomains. (It could happen, for example, that T is surjective but T^{**} is not.) Examination of the proof of Theorem 11.15 reveals that if we start with a bounded sequence $\{x_n\}_{n=1}^{\infty}$ in X then there will be a subsequence $\{x_{n_k}\}_{k=1}^{\infty}$ such that $\{Tx_{n_k}\}_{k=1}^{\infty}$ is a Cauchy sequence in Y. Since X is reflexive, we can obtain a "candidate" for the limit of this sequence by extracting a weakly convergent subsequence of $\{x_{n_k}\}_{k=1}^{\infty}$ and applying T to the weak limit.

For practice working with second duals, reflexivity, and adjoints, let's formalize the argument above.

Proposition 11.16: Let X be a reflexive Banach space, Y a normed linear space, and $T \in \mathcal{L}(X;Y)$ be given. Assume that T^* is compact. Then T is compact.

Proof: Since X is reflexive, it suffices to show that T maps weakly convergent sequences to strongly convergent ones. Let $\{x_n\}_{n=1}^{\infty}$ be a weakly convergent sequence in X. Choose $x \in X$ such that $x_n \rightharpoonup x$ (weakly) as $n \to \infty$. Since J_X is continuous, we know that $J_X(x_n) \rightharpoonup x$ (weakly) as $n \to \infty$. By Theorem 11.14, $T^{**} : X^{**} \to Y^{**}$ is compact, so we have

$$T^{**}(J_X(x_n)) \to T^{**}(J_X(x)) \text{ (strongly) as } n \to \infty, \text{ i.e.,}$$

$$J_Y(T(x_n)) \to J_Y(Tx) \text{ (strongly) as } n \to \infty.$$

Since J_Y is an isometry, we conclude that $Tx_n \to Tx$ (strongly) as $n \to \infty$ and consequently T is compact. $\square$

We now give two examples below of compact operators on a space of continuous functions. You will encounter compact operators on sequence spaces on Assignment 6.

Example 11.17: Let $X = C[0,1]$, the space of all continuous functions $f : [0,1] \to \mathbb{K}$ equipped with the norm given by

$$\|f\| = \max\{|f(t)| : t \in [0,1]\} \text{ for all } f \in X.$$

(a) Define $T : X \to X$ by

$$(Tf)(t) = \int_0^t f(s)ds \text{ for all } t \in [0,1].$$

To show that T is compact, let $\{f_n\}_{n=1}^{\infty}$ be a bounded sequence in X and choose $M \in \mathbb{R}$ such that $\|f_n\| \le M$ for all $n \in \mathbb{N}$. Then for all $n \in \mathbb{N}$ and all $t, \tau \in [0,1]$, we have

- $\|Tf_n\| \leq M$,
- $|(Tf_n)(t) - (Tf_n)(\tau)| \leq M|t - \tau|$,

i.e. the sequence $\{Tf_n\}_{n=1}^\infty$ is uniformly bounded and uniformly equicontinuous. By the Ascoli-Arzela Theorem, there is a uniformly convergent subsequence $\{Tf_{n_k}\}_{k=1}^\infty$ and we conclude that T is compact.

(b) Assume that $k : [0,1] \times [0,1] \to \mathbb{K}$ is continuous and define $K : X \to X$ by

$$(Kf)(t) = \int_0^1 k(t,s)f(s)ds \text{ for all } f \in X.$$

To show that K is compact, let $\{f_n\}_{n=1}^\infty$ be a bounded sequence in X and choose $M \in \mathbb{R}$ such that $\|f_n\| \leq M$ for all $n \in \mathbb{N}$. Put $\overline{k} = \max\{|k(t,s)| : (t,s) \in [0,1] \times [0,1]\}$. Let $\epsilon > 0$ be given. We may choose $\delta > 0$ such that

$$|k(t,s) - k(\tau,s)| < \frac{\epsilon}{M} \text{ for all } s, \tau, t \in [0,1] \text{ with } |t - \tau| < \delta.$$

Now for all $n \in \mathbb{N}$ and all $t, \tau \in [0,1] \times [0,1]$ with $|t - \tau| < \delta$ we have

- $\|Kf_n\| \leq M\overline{k}$,
- $|(Kf_n)(t) - (Kf_n)(\tau)| < \epsilon$,

i.e. the sequence $\{Kf_n\}_{n=1}^\infty$ is uniformly bounded and uniformly equicontinuous. By the Ascoli-Arzela Theorm, there is a uniformly convergent subsequence $\{Kf_{n_k}\}_{k=1}^\infty$ and we conclude that K is compact.

Continuous and Compact Embeddings of Normed Linear Spaces

Let X and Y be normed linear spaces over the same field. We say that X is *continuously embedded* in Y provided that $X \subset Y$ and the identity mapping $I : X \to Y$ is linear and continuous, i.e. there exists $M \in \mathbb{R}$ such that

$$\|x\|_Y \leq M\|x\|_X \text{ for all } x \in X.$$

The assumption that the identity mapping is linear ensures that the linear structures of the two spaces are compatible. We write

$$X \hookrightarrow Y$$

to indicate that X is continuously embedded in Y.

Remark 11.18: Assume that $X \hookrightarrow Y$ and let $x \in X$ and a sequence $\{x_n\}_{n=1}^\infty$ be given.

(a) If $x_n \to x$ (strongly) in X as $n \to \infty$ then $x_n \to x$ (strongly) in Y as $n \to \infty$.

(b) If $x_n \rightharpoonup x$ (weakly) in X as $n \to \infty$ then $x_n \rightharpoonup x$ (weakly) in Y as $n \to \infty$.

We say that X is *compactly embedded* in Y provided that $X \subset Y$ and the identity mapping $I : X \to Y$ is linear and compact. We write

$$X \hookrightarrow\hookrightarrow Y$$

to indicate that X is compactly embedded in Y.

Remark 11.19: Assume that $X \hookrightarrow\hookrightarrow Y$ and let $x \in X$ and a sequence $\{x_n\}_{n=1}^{\infty}$ be given.

(a) If $\{x_n\}_{n=1}^{\infty}$ is bounded in X then there is a subsequence $\{x_{n_k}\}_{k=1}^{\infty}$ that converges strongly in Y.

(b) If $x_n \rightharpoonup x$ (weakly) in X as $n \to \infty$ then $x_n \to x$ (strongly) in Y as $n \to \infty$.

Before presenting the next group of examples, we introduce a class of spaces of continuous functions on subsets of $\mathbb{R}^n$. Let Ω be a bounded open subset of $\mathbb{R}^n$. By $C(\overline{\Omega})$ we mean the set of all uniformly continuous functions $f : \Omega \to \mathbb{K}$. Unless stated otherwise, we equip this space with $\|\cdot\|_\infty$. (Notice that each function in $C(\overline{\Omega})$ is bounded and has a unique continuous extension to $\overline{\Omega}$.)

Lecture Notes for Week 12 (First Draft)

Continuous and Compact Embeddings (Continued)

Example 12.1 Let $p_1, p_2 \in (1, \infty)$ be given. Then we have

$$l^1 \hookrightarrow l^{p_1} \hookrightarrow l^{p_2} \hookrightarrow c_0 \hookrightarrow c \hookrightarrow l^\infty.$$

None of these embeddings can be compact. Indeed the sequence $\{e^{(n)}\}_{n=1}^\infty$ is bounded in each one of the above spaces and it is also true that when $n \neq m$, we have $\|e^{(n)} - e^{(m)}\| \geq 1$ in all of the spaces (and consequently there can be no strongly convergent subsequence).

Remark 12.2: For continuous and compact embeddings, it is not necessary to have $X \subset Y$. The identity operator can be replaced with any linear injection $I : X \to Y$. Continuity of I is required for a continuous embedding and compactness of I is required for a compact embedding. If there could be any doubt about what the mapping I might be, this should be made clear when the embeddings are mentioned.

Example 12.3 Let Ω be a nonempty bounded subset of $\mathbb{R}^n$ and let $p_1, p_2 \in (1, \infty)$ be given. Then we have

$$C(\overline{\Omega}) \hookrightarrow L^\infty(\Omega) \hookrightarrow L^{p_2}(\Omega) \hookrightarrow L^{p_1}(\Omega) \hookrightarrow L^1(\Omega).$$

It is not too difficult to show that none of these embeddings can be compact. Remark 12.2 applies here because the elements of $C(\overline{\Omega})$ are individual functions and the elements of $L^\infty(\Omega)$ are equivalences classes of functions. In this situation the injection mapping is clear from context.

Example 12.4: Let $\gamma \in (0, 1]$ be given. By $C^{0,\gamma}[0, 1]$, we mean the set of all functions $f : [0, 1] \to \mathbb{K}$ satisfying

$$|f|_{0,\gamma} = \sup\left\{ \frac{|f(t) - f(s)|}{|t - s|^\gamma} : s, t \in [0, 1], s \neq t \right\} < \infty. \tag{1}$$

If $\gamma < 1$, these functions are said to be Holder continuous with exponent γ; if $\gamma = 1$ they are said to be Lipschitz continuous. Notice that $|\cdot|_{0,\gamma}$ is a seminorm, but not a norm, because it vanishes on constant functions. We equip $C^{0,\gamma}$ with the norm defined by

$$\|f\|_{0,\gamma} = \|f\|_\infty + |f|_{0,\gamma} \ \text{ for all } f \in C^{0,\gamma}[0, 1].$$

(An equivalent norm is obtained by replacing $\|f\|_\infty$ with $|f(0)|$ or with $|f(a)|$ for any conveniently chose $a \in [0, 1]$.)

Let $X = (C^{0,\gamma}[0, 1], \|\cdot\|_{0,\gamma})$, $Y = (C[0, 1], \|\cdot\|_\infty)$ and $\{f_n\}_{n=1}^\infty$ be a bounded sequence in X. It is easy to see that $\{f_n\}_{n=1}^\infty$ is uniformly bounded and uniformly equicontinuous. By the Ascoli-Arzela Theorem, there is a subsequence $\{f_{n_k}\}_{k=1}^{\infty}$ that is uniformly convergent. We conclude that $X \hookrightarrow\hookrightarrow Y$. As an exercise for yourself, you should verify that if $0 < \gamma < \alpha \leq 1$ then

$$C^{0,\alpha}[0,1] \hookrightarrow\hookrightarrow C^{0,\gamma}[0,1].$$

Operator Topologies and Sequences of Bounded Linear Operators

In applications, we frequently encounter sequences of bounded linear operators that are generated by some sort of approximation scheme. There are several different (and important) types of convergence associated with such sequences.

Definition 12.5: Let $T \in \mathcal{L}(X;Y)$ and $\{T_n\}_{n=1}^{\infty}$ be a sequence in $\mathcal{L}(X;Y)$. We say that

(a) $T_n \to T$ in the *uniform operator topology* as $n \to \infty$ provided that $\|T_n - T\| \to 0$ as $n \to \infty$.

(b) $T_n \to T$ in the *strong operator topology* as $n \to \infty$ provided that for every $x \in X$ we have $T_n x \to Tx$ (strongly) as $n \to \infty$.

(c) $T_n \to T$ in the *weak operator topology* as $n \to \infty$ provided that for every $x \in X$ we have $T_n x \rightharpoonup Tx$ (weakly) as $n \to \infty$.

It is clear that convergence in the uniform operator topology implies convergence in the strong operator topology implies convergence in the weak operator topology. As the examples below illustrate, the converse implications are false in general.

Example 12.6: Let $X = Y = l^2$.

(a) For every $n \in \mathbb{N}$ define $T_n \in \mathcal{L}(X;Y)$ by

$$(T_n x)_k = \begin{cases} x_k & \text{if } k \leq n \\ \\ 0 & \text{if } k > n. \end{cases}$$

It is easy to see that $T_n \to I$ in the strong operator topology and that the sequence $\{T_n\}_{n=1}^{\infty}$ fails to be convergent in the uniform topology

(b) For every $n \in \mathbb{N}$ define $L_n \in \mathcal{L}(X;Y)$ by

$$(L_n x)_k = \begin{cases} 0 & \text{if } k \leq n \\ \\ x_{k-n} & \text{if } k > n. \end{cases}$$

Let $x \in X$ be given. Observe that $\|L_n x\| = \|x\|$ for all $n \in \mathbb{N}$ and that $L_n x \to 0$ componentwise as $n \to \infty$. It follows that $L_n x \rightharpoonup 0$ (weakly) as $n \to \infty$. The sequence $\{L_n x\}_{n=1}^{\infty}$ fails to be strongly convergent unless $x = 0$; indeed if it were to converge strongly, the limit would have to be equal to the weak limit which is zero, but $\|L_n x\| = \|x\|$ for all $n \in \mathbb{N}$. Therefore $L_n \to 0$ in the weak operator topology as $n \to \infty$ but the sequence $\{L_n\}_{n=1}^{\infty}$ fails to converge in the strong operator topology.

Remark 12.7: Let X be a normed linear, Y be a Banach space, and $\{T_n\}_{n=1}^{\infty}$ be a sequence in $\mathcal{L}(X;Y)$. Recall that if for every $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that $\|T_n - T_m\| < \epsilon$ for all $m, n \geq N$ then there exists $T \in \mathcal{L}(X;Y)$ such $T_n \to T$ in the uniform operator topology as $n \to \infty$. (This is the content of Proposition 2.23.)

We now give an analogous result for Cauchy sequences in the strong operator topology.

Proposition 12.8: Let X, Y be Banach spaces and $\{T_n\}_{n=1}^{\infty}$ be a sequence in $\mathcal{L}(X;Y)$. Assume that for every $x \in X$ the sequence $\{T_n x\}_{n=1}^{\infty}$ is a Cauchy sequence in Y. Then there exists $T \in \mathcal{L}(X;Y)$ such that $T_n \to T$ in the strong operator topology as $n \to \infty$.

Proof: Since Y is complete, the sequence $\{T_n x\}_{n=1}^{\infty}$ is convergent for every $x \in X$. Define $T : X \to Y$ by

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

It is clear that T is linear. Since X is complete, the Principle of Uniform Boundedness implies that we may choose $M \in \mathbb{R}$ such that $\|T_n\| \leq M$ for all $n \in \mathbb{N}$. It follows that

$$\|Tx\| \leq M\|x\| \quad \text{for all } x \in X,$$

and consequently T is bounded. $\square$

An analogue of Proposition 12.8 for the weak operator topology is part of Assignment 6.

It is important to understand if convergence of a sequence of bounded linear operators implies convergence of the sequence of adjoints.

Proposition 12.9: Let X, Y be normed linear spaces and let $T \in \mathcal{L}(X;Y)$ and a sequence $\{T_n\}_{n=1}^{\infty}$ in $\mathcal{L}(X;Y)$ be given.

(a) If $T_n \to T$ in the uniform operator topology (of $\mathcal{L}(X;Y)$) as $n \to \infty$ then $T_n^* \to T^*$ in the uniform operator topology (of $\mathcal{L}(Y^*;X^*)$) as $n \to \infty$.

(b) If $T_n \to T$ in the weak operator topology (of $\mathcal{L}(X;Y)$) as $n \to \infty$ then for every $y^* \in Y^*$ we have $T_n^* y^* \overset{*}{\rightharpoonup} T^* y^*$ (weakly*) in X^* as $n \to \infty$.

(c) If X is reflexive and $T_n \to T$ in the weak operator topology (of $\mathcal{L}(X;Y)$) as $n \to \infty$ then $T_n^* \to T^*$ in the weak operator topology (of $\mathcal{L}(Y^*;X^*)$) as $n \to \infty$.

Proof: To prove (a) we simply observe that $\|T_n^* - T^*\| = \|T_n - T\|$ for all $n \in \mathbb{N}$.

To prove (b), assume that $T_n \to T$ in the weak operator topology as $n \to \infty$ and let $x \in X, y^* \in Y^*$ be given. Then, as $n \to \infty$, we have

$$(T_n^* y^*)(x) = y^*(T_n x) \to y^*(Tx) = (T^* y^*)(x).$$

Part (c) follows immediately from (b) and the observation that if X is reflexive then weak and weak* convergence in X^* are the same thing. $\square$

Remark 12.10: Convergence of a sequence of operators in the strong operator topology does not imply convergence of the sequence of adjoints in the strong operator topology. You are asked to find an example in Assignment 6.

If a sequence of compact operators converges in the uniform operator topology (and the target space is complete), then the limit is a compact operator. This result is very useful in applications.

Theorem 12.11: Let X be a normed linear space and Y be a Banach space. Let $T \in \mathcal{L}(X;Y)$ and a sequence $\{T_n\}_{n=1}^{\infty}$ in $\mathcal{C}(X;Y)$ be given. Assume that $T_n \to T$ in the uniform operator topology as $n \to \infty$. Then $T \in \mathcal{C}(X;Y)$.

Proof: We want to show that $\mathrm{cl}(T[B_1(0)])$ is compact. Since Y is complete, it suffices to show that $T[B_1(0)]$ is totally bounded, i.e. for every $\epsilon > 0$, $T[B_1(0)]$ can be covered by a finite number of balls of radius ϵ. Let $\epsilon > 0$ be given. Since $T_n \to T$ in the uniform operator topology as $n \to \infty$ we may choose $N \in \mathbb{N}$ such that

$$\|T_N - T\| < \frac{\epsilon}{4}. \tag{2}$$

Observe that

$$\begin{aligned} \|Tx - Ty\| &\le \|(T - T_N)x\| + \|T_N x - T_N y\| + \|(T - T_N)y\| \\ &\le 2\|T - T_N\| + \|T_N x - T_N y\| \ \text{ for all } x, y \in B_1(0). \end{aligned} \tag{3}$$

Since T_N is compact, we may cover $\mathrm{cl}(T_N[B_1(0)])$ (and hence also $T_N[B_1(0)]$) by a finite number of balls of radius $\frac{\epsilon}{2}$; choose $x_1, x_2, \cdots, x_k \in B_1(0)$ such the centers of these balls are $T_N(x_1), T_N(x_2), \cdots, T_N(x_k)$:

$$T_N[B_1(0)] \subset \bigcup_{i=1}^{k} B_{\frac{\epsilon}{2}}(T_N(x_i)). \tag{4}$$

Then, by the triangle inequality, (2), (3), and (4), we have

$$T[B_1(0)] \subset \bigcup_{i=1}^{k} B_\epsilon(T(x_i)). \ \square$$

Remark 12.12: We see from part (a) of Example 12.6 that convergence in the strong operator topology does not preserve compactness. Indeed, for each $n \in \mathbb{N}$, the operator T_n is continuous and has finite rank, so it is compact. On ther other hand $T_n \to I$ in the strong operator toplogy as $n \to \infty$ and I is not compact.

Remark 12.13 (Approximation Problem): It follows immediately from Theorem 12.11 that if X and Y are Banach spaces and $\{T_n\}_{n=1}^{\infty}$ is a sequence in $\mathcal{L}(X;Y)$ such that each T_n has finite rank and $T_n \to T$ in the uniform operator topology as $n \to \infty$ then T is compact. An important question that remained open for a long time is the following: If X and Y are Banach spaces and $T \in \mathcal{C}(X;Y)$ does there exist a sequence $\{T_n\}_{n=1}^{\infty}$ in $\mathcal{L}(X;Y)$ such that each T_n has finite rank and $T_n \to T$ in the uniform operator topology as $n \to \infty$. The question was answered negatively by Per Enflo in 1973. We shall see that the answer is yes if Y is a Hilbert space.

Topological Vector Spaces

We now consider vector spaces having a topology that is naturally adapted to the vector space structure, but the topology need not be induced by a norm (or even by a metric). Such spaces are important because there are useful (non-metrizable) topologies (such as weak and weak* topologies) generated by classes of linear functionals on infinite-dimensional normed linear spaces. Moreover the theory of weak differentiation and distributions involves basic spaces of functions having natural topologies that are not induced by metrics. Such spaces also occur naturally in the study of holomorphic functions on open subsets of $\mathbb{C}$.

Definition 12.14: By a *topological vector space* (abbreviated *TVS*) we mean a pair (X, τ) where X is a linear space over $\mathbb{K}$ and τ is a topology on X such that

(i) (X, τ) is a Hausdorff space,

(ii) Addition is continuous from $X \times X$ to X,

(iii) Scalar Multiplication is continuous from $\mathbb{K} \times X$ to $\mathbb{X}$.

Remark 12.15: Many authors do not include item (i) (the Hausdorff property) as part of the definition of TVS. The Hausdorff property is needed for most results of interest and it is satisfied in all of the important examples of TVS that I have encountered. Rudin replaces (i) with the T1 separation axiom and then proves that T1, together with (ii) and (iii), imply the Hausdorff property.

Remark 12.16: In situations when there is no danger of confusion about what the topology could be, we refer to a "topological vector space X" and omit any specific reference to the topology τ.

Proposition 12.17: Let X be a topological vector space and $a \in X$, $\lambda \in \mathbb{K}\backslash\{0\}$ be given. Define the mappings $T_a, M_\lambda : X \to X$ by

$$T_a(x) = a + x, \quad M_\lambda(x) = \lambda x \text{ for all } x \in X.$$

Then T_a and M_λ are homeomorphisms of X onto X.

Proof: The vector space axioms and the definitions of T_a and M_λ imply that T_a and M_λ are bijective and that

$$(T_a)^{-1}(x) = -a + x, \quad (M_\lambda)^{-1}(x) = \lambda^{-1}x \text{ for all } x \in X.$$

Continuity of T_a, $(T_a)^{-1}$, M_λ and $(M_\lambda)^{-1}$ follow from continuity of addition and scalar multilpication. $\Box$

A very important consequence of this proposition is *translation invariance* of the topology.

Corollary 12.18: Let X be a topological vector space and $V \subset X$ be given. The following three statements are equivalent.

(a) V is open.

(b) There exists $x \in X$ such that $x + V$ is open.

(c) $x + V$ is open for every $x \in X$.

In view of Corollary 12.18, the topology in a TVS is completely determined by a local base at any point.

Conventions on Terminology: By a *local base* for a TVS we mean a local base at 0. By a *neighborhood*, we mean an open neighborhood; in particular, a neighborhood of a point x is simply any open set U with $x \in U$.

Another simple, but very useful, consequence of Corollary 12.18 is that if $A, U \subset X$ and U is open then $A + U$ is open.

Proposition 12.9: Let X be a topological vector space and $A, U \subset X$. Assume that U is open. Then $A + U$ is open.

Proof: Observe that

$$A + U = \bigcup_{x \in A} (x + U)$$

and apply Corollary 12.18. $\Box$

The notion of boundedness of a set plays an important role in the theory of topological vector spaces. Some care should be exercised with the terminology, because if the topology is induced by a metric, the sets which are metrically bounded can change when a different metric inducing the same topology is employed. We want important properties of sets to depend only on the topology and the linear structure.

Definition 12.20: Let X be a topological vector space and $E \subset X$ be given. We say that E is *topologically bounded* provided that for every neighborhood V of 0, there exists $t_0 > 0$ such that $E \subset tV$ for all $t > t_0$.

Remark 12.21: Many authors use the term "bounded" in place of "topologically bounded".

The following definition captures many important properties that a given TVS may or may not have.

Definition 12.22: Let X be a topological vector space. We say that X is

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

(b) *locally bounded* if there is a bounded neighborhood of 0.

(c) *locally compact* if there is a neighborhood V of 0 such that $\mathrm{cl}(V)$ is compact.

(d) *metrizable* if the topology is induced by some metric.

(e) *normable* if the topology is induced by a norm.

In order to give an overview of the big picture, the important characterizations of the properties in Definition 2.22 are given in the remark below. We will discuss these statements in detail later.

Remark 12.23: Let X be a topological vector space. Then

(i) X is locally convex if and only if the topology is generated by a separating family of seminorms.

(ii) X is locally compact if and only if X is finite dimensional.

(iii) X is metrizable if and only if there is a countable local base.

(iv) X is normable if and only if X is locally convex and locally bounded.

Lemma 12.24: Let X be a topological vector space and $A, B \subset X$. Then

$$\mathrm{cl}(A) + \mathrm{cl}(B) \subset \mathrm{cl}(A + B).$$

Proof: Consider the function $F : X \times X \to X$ defined by

$$F(x + y) = x + y \quad \text{for all } x, y \in X.$$

Since F is continuous we know that

$$F[\mathrm{cl}(A), \mathrm{cl}(B)] \subset \mathrm{cl}(F[A, B]),$$

which is precisely the desired conclusion. $\square$

Lemma 12.25: Let X be a topological vector space and K be a convex subset of X. Then $\mathrm{int}(K)$ and $\mathrm{cl}(K)$ are convex.

Proof: Let $U = \mathrm{int}(K)$ and $C = \mathrm{cl}(K)$. Since K is convex we know that

$$tK + (1-t)K \subset K \quad \text{for all } t \in (0,1).$$

Let $t \in (0,1)$ be given. Since $U \subset K$ we know that

$$tU + (1-t)U \subset K.$$

Since tU and $(1-t)U$ are open, it follows from Proposition 12.19 that $tU + (1-t)$ is open. Since $tU + (1-t)U$ is open and $tU + (1-t)U \subset K$ we know that

$$tU + (1-t)U \subset \mathrm{int}(K) = U,$$

and consequently U is convex.

To see that C is convex, let $t \in (0,1)$ be given and observe that

$$tC = \mathrm{cl}(tK), \quad (1-t)C = \mathrm{cl}((1-t)K).$$

Using Lemma 2.24, and the fact that $tK + (1-t)K \subset K$ we find that

$$tC + (1-t)C \subset \mathrm{cl}(tK + (1-t)K) \subset \mathrm{cl}(K) = C,$$

and consequently C is convex. $\square$

Lemma 12.26: Let X be a topological vector space. Every neighborhood of 0 contains a balanced neighborhood of 0.

Proof: Let W be a neighborhood of 0. Since scalar multiplication is continuous at $(0,0)$ we may choose a neighborhood V of 0 and $\delta > 0$ such that

$$\alpha V \subset W \quad \text{for all } \alpha \in \mathbb{K} \text{ with } |\alpha| < \delta.$$

Let $A = \{\alpha \in \mathbb{K} : |\alpha| < \delta\}$ and put

$$U = \bigcup_{\alpha \in A} \alpha V.$$

Since αV is open for every $\alpha \in A \backslash \{0\}$ (and since $0 \in \alpha V$ for $\alpha \neq 0$) we know that U is open. Furthermore, since $\gamma\alpha \in A$ for every $\gamma \in \mathbb{K}$ with $|\gamma| \leq 1$ we conclude that U is balanced. Finally, since $\alpha V \subset W$ for all $\alpha \in A$, we see that U is balanced.

Corollary 2.27: Let X be a topological vector space and $E \subset X$. Assume that for every neighborhood U of 0, there exists $t_0 > 0$ such that $E \subset t_0 U$. Then E is topologically bounded.

Proof: Let V be a neighborhood of 0. By Lemma 2.26, we may choose a balanced neighborhood U of 0 such that $U \subset V$. By assumption, we may choose $t_0 > 0$ such that $E \subset t_0 U$. Let $t > t_0$ be given. Then $t^{-1} t_0 U \subset U$ because U is balanced. Therefore, we have

$$E \subset t_0 U = t(t^{-1} t_0 U) \subset tU \subset tV,$$

and consequently E is bounded. $\square$

Lemma 12.28: Let X be a topological vector space. Every convex neighborhood of 0 contains a balanced convex neighborhood of 0.

Proof: Suppose that W is a convex neighborhood of 0. Let $F = \{\alpha \in \mathbb{K} : |\alpha| = 1\}$ and put

$$K = \bigcap_{\alpha \in F} \alpha W.$$

Notice that K is convex and $K \subset W$ (since $1 \in F$). By Lemma 12.26, we may choose a balanced neighborhood U of 0 such that $U \subset W$. Since U is balanced, we know that $\alpha^{-1} U = U$ for all $\alpha \in F$ and we conclude that

$$U \subset K.$$

Let $V = \mathrm{int}(K)$ and observe that $U \subset V$, so $0 \in V$. Furthermore, $V \subset W$ because $K \subset W$. By Lemma 12.25, we know that V is convex. It remains to show that V is balanced. To show that V is balanced, we shall first show that K is balanced. To this end, let $\gamma \in \mathbb{K}$ with $|\gamma| \leq 1$ be given. If $\gamma = 0$ then obviously $\gamma K \subset K$. Assume that $\gamma \neq 0$ and put $\beta = \frac{\gamma}{|\gamma|}$. Then we have

$$\gamma K = |\gamma| \beta K = \bigcap_{\alpha \in F} |\gamma| \beta \alpha W = \bigcap_{\alpha \in F} |\gamma| \alpha W. \tag{5}$$

For every $\alpha \in F$, αW is a convex set containing 0, and consequently $|\gamma| \alpha W \subset \alpha W$. Therefore (5) tells us that $\gamma K \subset K$ and K is balanced.

To conclude that V is balanced, first observe that $0V \subset V$ since $0 \in V$. Let $\gamma \in \mathbb{K}$ with $0 < |\gamma| \leq 1$ be given. Then, by Proposition 12.17, $\mathrm{int}(\gamma K) = \gamma V$ and consequently

$$\gamma V \subset \gamma K \subset K.$$

Since γV is open, we can conclude $\gamma V \subset \mathrm{int}(K) = V$. $\square$

Lemma 2.29: Let X be a topological vector space and V be a neighborhood of 0. Let $\{r_n\}_{n=1}^{\infty}$ be a sequence of strictly positive numbers such that $r_n \to \infty$ as $n \to \infty$.

Then

$$X = \bigcup_{n=1}^{\infty} r_n V.$$

Proof: Let $x \in X$ be given and put

$$A = \{\alpha \in \mathbb{K} : \alpha x \in V\}.$$

Since the mapping $\alpha \to \alpha x$ is continuous and V is an open set containing 0 we know that A is an open subset of $\mathbb{K}$ and that $0 \in A$. Therefore we may choose $N \in \mathbb{N}$ such that $(r_N)^{-1} \in A$. It follows that $x \in r_N V$. $\square$

Lemma 2.30: Let X be a topological vector space. Suppose that V is a bounded neighborhood of 0 and let $\{\delta_n\}_{n=1}^{\infty}$ be a sequence of strictly positive real numbers such that $\delta_n \to 0$ as $n \to \infty$. Then $\{\delta_n V : n \in \mathbb{N}\}$ is a local base.

Proof: Let U be a neighborhood of 0. We need to shoe that there exists $n \in \mathbb{N}$ such that $\delta_n V \subset W$. Since V is bounded, we may choose $t_0 > 0$ such that $V \subset tU$ for all $t > t_0$. Choose $N \in \mathbb{N}$ such that $\delta_n t_0 < 1$. Then we have $\delta_N V \subset U$. $\square$

Remark 2.31: In view of Remark 12.23 (iii) and Lemma 2.30, any topological vector space having a bounded neighborhood of 0 is metrizable. In other words, if X fails to be metrizable, then X cannot have a bounded neighborhood of 0.

Lecture Notes for Week 13 (First Draft)

Linear Mappings Between TVS

Definition 13.1: Let X and Y be topological vector spaces. A function $F : X \to Y$ is said to be *uniformly continuous* provided that for every neighborhood V of 0 in Y there is a neighborhood U of 0 in X such that

$$F(x) - F(y) \in V \quad \text{for all } x, y \in X \text{ with } x - y \in U.$$

Definition 13.2: Let X and Y be topological vector spaces and $T : X \to Y$ be a linear mapping. We say that T is bounded provided that $T[E]$ is topologically bounded for every topologically bounded set $E \subset X$.

Proposition 13.3: Let X and Y be topological vector spaces and assume that $T : X \to Y$ is linear. The following three statements are equivalent,

(i) T is uniformly continuous.

(ii) T is continuous on X.

(iii) T is continuous at 0.

Proof: It is clear that (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii). We need to show that (iii) $\Rightarrow$ (i).

Assume that T is continuous at 0. Let V be a neighborhood of 0 in Y. Since T is continuous at 0 and $T0 = 0$ we may choose a neighborhood U of 0 in V such that $T[U] \subset V$. Let $x, y \in X$ with $x - y \in U$ be given. Then $T(x - y) \in V$. Since $T(x - y) = Tx - Ty$, we conclude that $Tx - Ty \in V$. $\Box$

Proposition 13.4: Let X and Y be topological vector spaces and assume that $T : X \to Y$ is linear. If T is continuous then T is bounded.

Remark 13.5: The converse of Proposition 13.4 is false in general, but is true if X is metrizable.

Proof of Proposition 13.4: Assume that T is continuous and let E be a topologically bounded subset of X. Let V be a neighborhood of 0 in Y. Since T is continuous we may choose a neighborhood U of 0 in X such that $T[U] \subset V$. Since E is topologically bounded, we may choose $t_0 > 0$ such that $E \subset tU$ for all $t > t_0$. Then for all $t > t_0$ we have

$$T[E] \subset T[tU] = tT[U] \subset tV,$$

and consequently $T[E]$ is topologically bounded. $\Box$

Theorem 13.6: Let X be a topological vector space and assume that $f : X \to \mathbb{K}$ is linear and nontrivial (i.e. $f(x) \neq 0$ for some $x \in X$). The following four statements are equivalent.

(a) f is continuous,

(b) $\mathcal{N}(f)$ is closed,

(c) $\mathcal{N}(f)$ is not dense in X,

(d) There is a neighborhood W of 0 such that f is bounded on W.

Proof: We shall show that

$$(a) \Rightarrow (b) \Rightarrow (c) \Rightarrow (d) \Rightarrow (a).$$

Assume that f is continuous. Since $\{0\}$ is a closed subset of $\mathbb{K}$ and f is continuous, we conclude that $\mathcal{N}(f) = \{x \in X : f(x) \in \{0\}\}$ is closed, and consequently (a) $\Rightarrow$ (b).

Since f is not the zero functional, we know that $\mathcal{N}(f) \neq X$, and consequently (b) $\Rightarrow$ (c).

Assume that (c) holds. Then

$$\operatorname{int}((\mathcal{N}(f))^c) \neq \emptyset.$$

Therefore we may choose $x \in (\mathcal{N}(f))^c$ and a neighborhood U of 0 such that

$$(x + U) \cap \mathcal{N}(f) = \emptyset.$$

By Lemma 12.26, we may choose a balanced neighborhood W of 0 with $W \subset U$. Observe that

$$(x + W) \cap \mathcal{N}(f) = \emptyset. \tag{1}$$

Then $f[W]$ is a balanced subset of $\mathbb{K}$. This implies that either $f[W]$ is bounded or $f[W] = K$. Suppose that $f[W] = \mathbb{K}$. Then we may choose $y \in W$ such that $f(y) = -f(x)$. This implies that $x + y \in \mathcal{N}(f)$, which is impossible because of (1). We conclude that $f[W]$ is bounded and (d) holds.

Assume that there is a neighborhood W of 0 such that f is bounded on W. We shall show that f is continuous at 0. Choose $M > 0$ such that $|f(x)| < M$ for all $x \in W$. Let $\epsilon > 0$ be given and put $V = \epsilon M^{-1} W$. Then we have $|f(x)| < \epsilon$ for all $x \in V$ and f is continuous at 0. By Proposition 13.3, f is continuous. $\square$

Metrization

Remark 13.7: Let X be a metrizable topological vector space and let ρ be a metric that generates the topology. Put $V_n = \{x \in X : \rho(x, 0) < \frac{1}{n}\}$ for every $n \in \mathbb{N}$. Then $\{V_n : n \in \mathbb{N}\}$ is a countable local base.

Theorem 13.8: Let X be a topological vector space and assume that X has a countable base. Then there is a metric $\rho : X \times X \to [0, \infty)$ such that

(a) ρ induces the topology on X,

(b) ρ is translation invariant,

(c) Each open ball centered at 0 is balanced.

If X is locally convex, then there exists a metric $\rho : X \times X \to [0, \infty)$ such that (a), (b), (c) holds and

(d) Each open ball is convex.

Lemma 13.9: Let X be a topological vector space and V be a neighborhood of 0. Then there exists a balanced neighborhood U of 0 such that $U + U \subset V$.

Proof: Let V be a neighborhood of 0. Since addition is continuous at $(0, 0)$ we may choose neighborhoods V_1 and V_2 of 0 such that $V_1 + V_2 \subset V$. By Lemma 12.26, we may choose balanced neighborhoods U_1 and U_2 of 0 such that $U_1 \subset V_1$ and $U_2 \subset V_2$. Let us put $U = U_1 \cap U_2$. Then U is a balanced neighborhood of 0 and $U + U \subset V$. □

Proof of Theorem 13.8: We shall prove the existence of a metric ρ satisfying (a), (b), and (c). It is possible to construct the metric in such a way that if X is locally convex, then (d) also holds. We shall not do so, because when X is locally convex, there is a different construction of ρ satisfying (a), (b), (c), and (d). This construction will be part of Assignment 7. Choose a local base $\{V_n : n \in \mathbb{N}\}$ such that each V_n is balanced and

$$V_{n+1} + V_{n+1} \subset V_n. \tag{2}$$

Let $\mathcal{F}$ denote the set of all nonempty finite subsets of $\mathbb{N}$. Given $F \in \mathcal{F}$ and $n \in \mathbb{N}$, we write $n < F$ to indicate that $n < k$ for all $k \in F$.

For each $F \in \mathcal{F}$, put

$$V^F = \sum_{n \in F} V_N,$$

$$q^F = \sum_{n \in F} 2^{-n}.$$

Notice that each V^F is balanced. Using an induction argument and (2), we can show that for all $F \in \mathcal{F}$ and $n \in \mathbb{N}$, we have

$$q^F < 2^{-n} \Rightarrow n < F \Rightarrow V^F \subset V_n.$$

Define the function $f : X \to \mathbb{R}$ by

$$f(x) = \begin{cases} \inf\{q^F : x \in V^F\} & \text{if } x \in \bigcup\limits_{F \in \mathcal{F}} V^F \\ \\ 1 & \text{if } x \notin \bigcup\limits_{F \in \mathcal{F}} V^F. \end{cases}$$

Notice that $f(x) \in [0,1]$ for all $x \in X$. Furthermore, since

$$\forall x \in X,\ F \in \mathcal{F}\ x \in F \Leftrightarrow -x \in F,$$

it follows that $f(-x) = f(x)$ for all $x \in X$. I claim that the function $\rho : X \times X \to [0,\infty)$ defined by

$$\rho(x,y) = f(x-y)$$

does the trick.

Clearly $0 \le \rho(x,y) = \rho(y,x) < \infty$ for all $x, y \in X$. Since $0 \in V^F$ for all $F \in \mathcal{F}$, we know that $f(0) = 0$. Since X has the Hausdorff property, if $x \in X \backslash \{0\}$, we may choose N sufficiently large so that $x \notin V^F$ whenever $N < F$ and consequently

$$f(x) > 0 \ \text{ for all } x \in X \backslash \{0\}.$$

We conclude that

$$\forall x, y \in X,\ \rho(x,y) = 0 \ \Leftrightarrow x = y.$$

To establish the triangle inequality for ρ, it suffices to show that

$$f(x+y) \le f(x) + f(y) \ \text{ for all } x, y \in X. \tag{3}$$

Let $x, y \in X$ be given. If $f(x) + f(y) \ge 1$ then we get $f(x+y) \le f(x) + f(y)$ "for free" because $f(z) \le 1$ for all $z \in X$. Assume that $f(x) + f(y) < 1$. Then we may choose $\epsilon > 0$ such that

$$f(x) + f(y) + 2\epsilon < 1.$$

We may also choose $F, G \in \mathcal{F}$ such

$$x \in V^F, \quad q^F < f(x) + \epsilon, \quad y \in V^G, \quad q^G < f(y) + \epsilon.$$

There is exactly one $H \in \mathcal{F}$ such that

$$q^H = q^F + q^G.$$

By virtue of (2) we have

$$V^F + V^G \subset V^H.$$

It follows that $x + y \in V^H$ and consequently

$$f(x+y) \le q^H = q^F + q^G < f(x) + f(y) + 2\epsilon.$$

Since $\epsilon > 0$ was arbitrary, we conclude that (3) holds.

We have established that ρ is a metric. It is immediate that ρ is translation invariant because

$$\rho(x+z, y+z) = f(x+z-y-z) = f(x-y) = \rho(x,y) \text{ for all } x, y, z \in X.$$

For each $\delta > 0$ put

$$B_\delta = \{x \in X : \rho(x,0) < \delta\}.$$

It is straightforward to show that $\{V_\delta : \delta > 0\}$ is a local base for X. It follows that the topology of X is induced by ρ.

To see that each open ball centered at 0 is balanced, it suffices to show that

$$f(\alpha x) \leq f(x) \text{ for all } x \in X,\ \alpha \in \mathbb{K} \text{ with } |\alpha| \leq 1. \tag{4}$$

Since each V^F is balanced, we know that if $x \in V^F$ and $\alpha \in \mathbb{K}$ with $|\alpha| \leq 1$ then $\alpha x \in V^F$ and consequently $f(\alpha x) \leq f(x)$ and (4) is satisfied.

Normability

Theorem 13.10: Let X be a topological vector space. Then X is normable if and only if 0 has a bounded convex neighborhood.

Proof: If X is normable, then we may choose a norm $\|\cdot\|$ that induces the topology and it is straightforward to check that $\{x \in X : \|x\| < 1\}$ is a bounded convex neighborhood of 0.

Assume that X has a bounded convex neighborhood U of zero. By Lemma 12.28, we may choose a balanced convex neighborhood W of 0 such that $W \subset U$. Continuity of scalar multiplication implies that W is absorbing. Let us define the function $\|\cdot\| : X \to [0,\infty)$ by

$$\|x\| = p^U(x), \quad \text{for all } x \in X,$$

where p^U is the Minkowski functional for U. From previous considerations, we know that $\|\cdot\|$ is a seminorm. To show that it is a norm, we need to show that

$$\forall x \in X, \quad \|x\| = 0 \Rightarrow x = 0.$$

Lecture Notes for Week 14 (First Draft)

Local Convexity and Seminorms

Let X be a linear space over $\mathbb{K}$. There is a natural correspondence between topologies τ on X such that (X, τ) is a locally convex topological vector space and separating families of seminorms on X.

Theorem 14.1: Let X be a locally convex topologivcal vector space and assume that $\mathcal{B}$ is a local base such that each $V \in \mathcal{B}$ is balanced and convex. (Each $V \in \mathcal{B}$ is automatically absorbing.) Then $\{p^V : V \in \mathcal{B}\}$ is a separating family of continuous seminorms.

Proof: We know from Lemma 5.32 that each p^V is a seminorm. The Hausdorff property implies that the family $\{p^V \; V \in \mathcal{B}\}$ is separating. Let $V \in \mathcal{B}$ and $x \in V$ be given. Since V is open and scalar multiplication is continuous, we may choose $t > 1$ such that $tx \in V$. It follows that $p^V(x) < 1$. In particular, we have

$$p^V(x) < 1 \;\text{ for all } x \in V,\; V \in \mathcal{B}.$$

Let $V \in \mathcal{B}$ and $\epsilon > 0$ be given. Then for all $x, y \in X$ with $x - y \in \epsilon V$, we have

$$|p^V(x) - p^V(y)| \leq p^V(x - y) < \epsilon,$$

by virtue of Proposition 6.8(c). $\square$

Theorem 14.2: Let X be a linear space over $\mathbb{K}$ and $\mathcal{P}$ be a separating family of seminorms. For each $p \in \mathcal{P}$ and $n \in \mathbb{N}$ put

$$V(p, n) = \{x \in X : p(x) < \frac{1}{n}\}.$$

Let $\mathcal{B}$ be the collection of finite intersections of sets of the form $V(p, n)$, $p \in \mathcal{P}$, $n \in \mathbb{N}$. Then each member of $\mathcal{B}$ is balanced and convex. Moreover $\mathcal{B}$ is a local base for a topology τ such that (X, τ) is a topological vector space and each $p \in \mathcal{P}$ is continuous. Furthermore, for every $E \subset X$ we have that E is topologically bounded if and only if

$$\forall p \in \mathcal{P}, \;\; p \text{ is toplogically on } E.$$

Proof: Declare $S \subset X$ to be open if and only if S can be expressed as a union of translates of members of $\mathcal{B}$. This collection of open sets is a translation invariant topology on X. To see that the Hausdorff property holds, let $x, y \in X$ with $x \neq y$ be given. Put $z = x - y$. Since $z \neq 0$ we may choose $p \in \mathcal{P}$ such that $p(z) > 0$. Put $r = p(z)$ and choose $N \in \mathbb{N}$ such that $\frac{1}{N} < \frac{r}{2}$. Then $x + V(p, N)$ is a neighborhood of x and $y + V(p, N)$ is a neighborhood of y such that

$$(x + V(p, N)) \cap (y + V(p, N)) = \emptyset.$$

To see that addition is continuous, let U be a neighborhood of 0. Then we may choose $p_1, p_2, \cdots, p_k \in \mathcal{P}$ and $n_1, n_2, \cdots, n_k \in \mathbb{N}$ such that

$$\bigcap_{i=1}^{k} V(p_i, n_i) \subset U.$$

Now, put

$$W = \bigcap_{i=1}^{k} V(p_i, 2n_i).$$

Using the subadditivity of seminorms, we conclude that $W + W \subset U$, and consequently addition is continuous at $(0, 0)$. Since the topology is translation invariant, we conclude that addition is continuous on $X \times X$.

To see that scalar multiplication is continuous, let $x \in X$, $\alpha \in \mathbb{K}$, and a neighborhood U of 0 be given. Define W as above and choose $s > 0$ such that $x \in sW$. (Any $s > 2n_i p_i(x)$ for all $i \in \{1, 2, \cdots, k\}$ will do.) Now choose $N \in \mathbb{N}$ with

$$N > \frac{(1 + |\alpha| s)}{s},$$

and put

$$\hat{W} = \bigcap_{i=1}^{k} V(p_i, 2Nn_i).$$

Observe that

$$\hat{W} \subset \frac{(1 + |\alpha| s)}{s} W.$$

Then for all $y \in x + \hat{W}$ and $\beta \in \mathbb{K}$ with $|\beta - \alpha| < \frac{1}{s}$ we have

$$\begin{aligned} \beta y - \alpha x &= \beta(y - x) + (\beta - \alpha)x \\ &\in |\beta| \hat{W} + |\beta - \alpha| sW \\ &\in W + W \subset U. \end{aligned}$$

It follows that scalar multiplication is continuous.

Let $p \in \mathcal{P}$ be given. From the construction of $\mathcal{B}$ we see that p is continuous at 0. It follows from Proposition 6.8 that p is continuous on X.

Let $E \subset X$ be given. Assume that E is topologically bounded. Let $p \in \mathcal{P}$ be given. Since $V(p, 1)$ is a neighborhood of zero, we may choose $k \in \mathbb{N}$ such that $E \subset kV(p, 1)$. This implies that $p(x) \leq k$ for all $x \in E$.

Assume now that p is bounded on E for each $p \in \mathcal{P}$. Let U be a neighborhood of 0. Then we may choose $p_1, p_2, \cdots, p_k \in \mathcal{P}$ and $n_1, n_2, \cdots, n_k \in \mathbb{N}$ such that

$$\bigcap_{i=1}^{k} V(p_i, n_i) \subset U.$$

For each $i \in \{1, 2, \cdots, k\}$, we may choose $M_i \in \mathbb{R}$ such that

$$p_i(x) < M_i \quad \text{for all } x \in E.$$

Then for all $t > \max\{M_i n_i : i = 1, 2, \cdots, k\}$ we have

$$E \subset tU. \quad \square$$

Remark 14.3: It is not too difficult to show that if X is a locally convex topological vector space and one takes any local base $\mathcal{B}$ consisting of balanced convex sets and then applies the construction of Theorem 14.2 to the family $\{p^V : V \in \mathcal{B}\}$ of seminorms, the topology so obtained coincides with the original topology of X.

Cauchy Sequences and Completeness

Definition 14.4: Let X be a topological vector space. A sequence $\{x_n\}_{n=1}^{\infty}$ is said to be a *Cauchy sequence* provided that for every neighborhood U of 0 there exists $N \in \mathbb{N}$ such that

$$x_n - x_m \in U \quad \text{for all } m, n \in \mathbb{N} \text{ with } m, n \geq N.$$

Remark 14.5: Let X be a metrizable topological vector space and let ρ be a translation invariant metric that induces the topology. Let $\{x_n\}_{n=1}^{\infty}$ be a sequence in X. Then $\{x_n\}_{n=1}^{\infty}$ is a Cauchy sequence in the sense of Definition 14.4 if and only if it is a Cauchy sequence in the metric space (X, ρ). In particular, all translation invariant metrics generating the topology of X have the same Cauchy sequences.

Definition 14.6: By a *Fréchet space* we mean a topological vector space whose topology is induced by a complete translation invariant metric.

Remark 14.7: It is important to keep in mind that in a Fréchet space topological boundedness of a set is not the same as boundedness in the metric sense.

Remark 14.8: Many authors require Fréchet spaces to be locally convex as part of the definition. The term *F-space* is also used in some places to indicate a topological vector space whose topology is induced by a complete translation invariant metric. Some authors require F-spaces to be locally convex, while others do not. In any event you should check carefully the convention being used with regard to local convexity whenever you encounter the term "Fréchet space" or "F-space" in the literature.

The Open Mapping Theorem and Closed Graph Theorem extend to Fréchet spaces (no local convexity needed). The statements of these results are recorded below. A good reference for the proofs is Rudin.

Theorem 14.9 (Open Mapping Theorem): Let X and Y be Fréchet spaces and assume that $T : X \to Y$ is continuous, linear, and surjective. Then T is an open mapping, i.e. $T[U]$ is open in Y for every open set $U \subset X$.

Theorem 14.10 (Closed Graph Theorem): Let X and Y be Fréchet spaces. Assume that $T : X \to Y$ is linear and that $\mathrm{Gr}(T)$ is closed. Then T is continuous.

A version of the Banach-Steinhaus Theorem (Principle of Uniform Boundedness) holds for topological vector spaces. In order to state such a result, we need to extend the definitions of sets of the first and second categories to topological vector spaces and to make a definition of equicontinuos family of linear mappings.

Definition 14.11: Let X be a topological vector space and $S \subset X$. We say that S is

(a) *nowhere dense* provided $\mathrm{int}(\mathrm{cl}(S)) = \emptyset$.

(b) of the *first category* if S can be expressed as a countable union of nowhere dense sets.

(c) of the *second category* if S it is not of the first category.

Definition 14.12: Let X and Y be topological vector spaces and $(T_i | i \in J)$ be a family of linear mappings from X to Y. We say that $(T_i | i \in J)$ is *equicontinuous* provided that for every neighborhood V of 0 in Y there is a neighborhood U of 0 in X such that $T_i[U] \subset V$ for all $i \in J$.

An good reference for the proof of Theorem 14.13 below is Rudin.

Theorem 14.13 (Banach-Steinhaus Theorem): Let X be a Fréchet space and Y be a topological vector space. Let $(T_i | i \in I)$ be a family of continuous linear mappings from X to Y and assume that

$$\forall x \in X, \quad \{T_i x : i \in J\} \text{ is topologically bounded in } Y.$$

Then $(T_i | i \in J)$ is equicontinuous.

A simple, but very useful, consequence of Theorem 14.13 is given below.

Theorem 14.14: Let X be a Fréchet space and Y be a topological vector space, and $\{T_n\}_{n=1}^{\infty}$ be a sequence of continuous linear amppings from X to Y. Assume that $\{T_n x\}_{n=1}^{\infty}$ is convergent for every $x \in X$ and put

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

Then T is continuous.

Proof: Let W be a neighborhood of 0 in Y. It is straightforward to show that we may choose a neighborhood V of 0 in Y such that $\mathrm{cl}(V) \subset W$. Observe that $(T_n | n \in \mathbb{N})$ is equicontinuous. Therefore, we may choose a neighborhood U of 0 in X such that $T_n[U] \subset V$ for all $n \in \mathbb{N}$. It follows that $T[U] \subset \mathrm{cl}(V) \subset W$ and consequently T is continuous. $\square$

Let X be a normed linear space with dual space X^*. For each $x^* \in X^*$, put

$$p_{x^*}(x) = |x^*(x)| \text{ for all } x \in X,$$

and for each $x \in X$, put

$$q_x(x^*) = |x^*(x)| \text{ for all } x^* \in X^*.$$

Then $\{p_{x^*} : x^* \in X^*\}$ is a separting family of seminorms on X and $\{q_x : x \in X\}$ is a separating family of seminorms on X^*.

Notation: Given a normed linear space Y, we write $\overline{B}_1^Y(0)$ for the closed unit ball in Y, i.e.

$$\overline{B}_1^Y(0) = \{y \in Y : \|y\|_Y \leq 1\}.$$

Definition 14.15 (Weak Topology): The family of seminorms $\{p_{x^*} : x^* \in X^*\}$ generates a topology, denoted by $\sigma(X, X^*)$ and called the *weak topology* of X, such that $(X, \sigma(X, X^*))$ is a locally convex topological vector space and for each $x^* \in X^*$, $p_{x^*} : X \to \mathbb{R}$ is continuous. Some authors refer to $\sigma(X, X^*)$ as the X^*-topology of X.

Definition 14.16 (Weak* Topology): The family of seminorms $\{q_x : x \in X\}$ generates a topology, denoted by $\sigma(X^*, X)$ and called the *weak* topology* of X^*, such that $(X, \sigma(X^*, X))$ is a locally convex topological vector space and for each $x \in X$, $q_x : X^* \to \mathbb{R}$ is continuous. Some authors refer to $\sigma(X^*, X)$ as the X-topology of X^*

Remark 14.17: Let X be a normed linear space. Then the weak topology on X is weaker than the norm topology (i.e., every weakly open set is open).

Remark 14.18: Let X be a normed linear space. The weak* topology on X^* is weaker than the weak topology on X (i.e., $\sigma(X^*, X)$ is weaker than $\sigma(X^*, X^{**})$) and the weak topology on X^* is weaker than the norm topology on X^*. Moreover, the weak and weak* topologies on X^* coincide if and only if X is reflexive.

In situations when weak or weak* topologies are relevant, we shall refer to the norm topologies on X and X^* as the *strong topologies* on X and X^*, respectively.

Remark 14.19: Let X be a normed linear space and let $f : X \to \mathbb{K}$ be a linear functional. Then f is strongly continuous if and only if it is weakly continuous.

You should verify Remarks 14.17 – 14.19 as an exercise for yourself.

Proposition 14.20: Let X be a normed linear space and assume that $g : X^* \to \mathbb{K}$ is linear. Then g is weakly* continuous if and only if there exists $x \in X$ such that

$$g(x^*) = x^*(x) \text{ for all } x^* \in X^*. \tag{1}$$

Proof: It follows immediately from the definition of weak* topology that if g has the form (1) then g is weakly* continuous.

Suppose that g is weakly* continuous. Then the set U^* defined by

$$U^* = \{x^* \in X^* : |g(x^*)| < 1\}$$

is a weak* neighborhood of 0 in X^*. By the definition of weak* topology, we may choose $\epsilon > 0$ and $x_1, x_2, \cdots, x_N \in X$ such that

$$\{x^* \in X^* : |x^*(x_i)| < \epsilon,\ i = 1, 2, \cdots, N\} \subset U^*.$$

Observe that

$$\bigcap_{i=1}^{N} \{x^* \in X^* : x^*(x_i) = 0\} \subset \mathcal{N}(g).$$

It follows from a standard result in linear algebra that g is a linear combination of the linear functionals

$$x^* \mapsto x^*(x_i), \quad i = 1, 2, \cdots, N.$$

We conclude that there exists $x \in X$ such that (1) holds. $\square$

Theorem 14.21 (Alaoglu's Theorem): Let X be a normed linear space. Then $\overline{B}_1^{X^*}(0)$ is $\sigma(X^*, X)$-compact (i.e., the closed unit ball in X^* is weakly* compact).

Proof: Let us write B^* for $\overline{B}_1^{X^*}(0)$ and B for $\overline{B}_1^{X}(0)$. For each $x \in B$, put

$$D_x = \{\alpha \in \mathbb{K} : |\alpha| \le 1\}.$$

$$D = \prod_{x \in B} D_x,$$

equipped with the product topology. Then D is compact by Tychonov's Theorem.

Define $\phi : B^* \to D$ by

$$(\phi(x^*))(x) = x^*(x) \ \text{ for all } x^* \in B^*,\ x \in B.$$

Let $x^*, y^* \in B^*$ be given and suppose that $\phi(x^*) = \phi(y^*)$. Then $x^*(x) = y^*(x)$ for all $x \in B$, which implies that $x^* = y^*$. It follows that ϕ is injective.

Let $x^* \in B^*$ be given and let $(x^*_\lambda)_{\lambda \in J}$ be a net in B^*. (Here, J is a directed set. If you are not familiar with nets, See pages 171 – 172 of Carothers or Appendix A in Conway for a nice quick introduction.) Then for each $x \in B$ we have

$$(\phi(x^*_\lambda))(x) = x^*_\lambda(x),$$

so that

$$\phi(x^*_\lambda) \to \phi(x^*) \Leftrightarrow x^*_\lambda \to x^*.$$

It follows that $\phi : B^* \to \phi[B^*]$ is a homeomporphism.

We shall now show that $\phi[B^*]$ is closed in D. Let $(x^*_\lambda)_{\lambda \in J}$ be a net in B^*, $z \in D$, and assume that $(\phi(x^*_\lambda))_{\lambda \in J}$ converges to z in D. Then

$$z(x) = \lim x^*_\lambda(x) \text{ exists for every } x \in X.$$

There is exactly one linear mapping $F : X \to \mathbb{K}$ such that

$$F(x) = z(x) \text{ for all } x \in B.$$

Notice that $|F(x)| \leq 1$ for all $x \in B$. We conclude that $F \in B^*$ and $\phi(F) = z$. It follows that $\phi[B^*]$ is closed. Since every closed subset of a compact set is compact, we conclude that $\phi[B^*]$ is compact. Since ϕ is a homeomorphism, we conclude that B^* is compact. $\square$

Remark 14.22: Alaglou's Theorem is one of the most important results in functional analysis. In fact, we could now talk about "Four Basic Principles of Linear Functional Analysis": Hahn Banach Theorem, Open Mapping Theorem (Closed Graph Theorem), Banach-Steinhaus Theorem, and Alaoglu's Theorem.

Remark 14.23: The use of nets is very convenient for many arguments involving weak or weak* topologies. However, one must keep in mind that arguments involving subnets can be tricky.

Theorem 14.24: Let X be a separable normed linear space and K^* be a weakly* compact subset of X^*. Then $(K^*, \sigma(X^*, X))$ is metrizable.

Proof: Choose a sequence $\{x_n\}_{n=1}^\infty$ that is dense in X and define $\rho : K^* \times K^* \to \mathbb{R}$ by

$$\rho(x^*, y^*) = \sum_{n=1}^\infty 2^{-n} \frac{|(x^* - y^*)(x_n)|}{1 + |(x^* - y^*)(x_n)|} \text{ for all } x^*, y^* \in K^*.$$

Let us denote the topology induced by this metric on K^* by τ. It is straightforward to verify that the identity mapping from $(K^*, \sigma(X^*, X))$ to (K^*, τ) is continuous. Since every continuous injection from a compact space to a Hausdorff space has a continuous inverse, we conclude that $(K^*, \sigma(X^*, X)) = (K^*, \tau)$. $\square$.

Lemma 14.25: Let X be a normed linear space. Then $J : (X, \sigma(X, X^*)) \to (J[X], \sigma(X^{**}, X^*))$ is a homeomorphism.

Proof: We know from previous considerations that $J : X \to J[X]$ is bijective. To show that J is continuous, it suffices to show that J is continuous at 0. Let V^{**} be a neighborhood of 0 in $(X^{**}, \sigma(X^{**}, X^*))$. Then we may choose $\epsilon > 0$ and $x_1^*, x_2^*, \cdots, x_k^* \in X^*$ such that

$$\{x^{**} \in X^{**} : |x^{**}(x_i^*)| < \epsilon,\ i = 1, 2, \cdots, k\} \subset V^{**}.$$

Let

$$W = \{x \in X : |x_i^*(x)| < \epsilon,\ i = 1, 2, \cdots, k\}.$$

Since $J(x)(x_i^*) = x_i^*(x)$ for all $x \in X$ and $i = 1, 2, \cdots, k$ we have $J[W] \subset V^{**}$ and J is continuous.

To show that $J^{-1} : (J[X], \sigma(X^{**}, X^*)) \to (X, \sigma(X, X^*))$ is continuous, let U be a neighborhood of 0 in $(X, \sigma(X, X^*)$. We may choose $\delta > 0$ and $y_1^*, y_2^*, \cdots, y_m^* \in X^*$ such that

$$\{x \in X : |y_i^*(x)| < \delta,\ i = 1, 2, \cdots, m\} \subset U.$$

Put

$$W^{**} = \{x^{**} \in X^{**} : |x^{**}(y_i^*)| < \delta,\ i = 1, 2, \cdots, m\}.$$

It is straightforward to check that if $z^{**} \in J[X] \cap W^{**}$ then $J^{-1}(z^{**}) \in U$, and consequently J^{-1} is $\sigma(X^{**}, X^*)$-continuous at 0. Since J^{-1} is linear, we conclude that $J^{-1} : (J[X], \sigma(X^{**}, X^*)) \to (X, \sigma(X, X^*))$ is continuous. $\square$

Theorem 14.26 (Goldstine's Theorem): Let X be a normed linear space. Then $J[\overline{B}_1^X(0)]$ is $\sigma(X^{**}, X^*)$-dense in $\overline{B}_1^{X^{**}}(0)$.

Proof: Put

$$B = \overline{B}_1^X(0), \quad B^{**} = \overline{B}_1^{X^{**}}(0),$$

and let $\hat{B}^{**}$ denote the $\sigma(X^{**}, X^*)$-closure of $J[B]$. By Alaoglu's Theorem, B^{**} is $\sigma(X^{**}, X^*)$-compact and consequently it is $\sigma(X^{**}, X^*)$-closed as well. Since $J[B] \subset B^{**}$, we conclude that $\hat{B}^{**} \subset B^{**}$. Since $J[B]$ is convex, it follows from that $\hat{B}^{**}$ is convex. We need to show that $\hat{B}^{**} = B^{**}$.

Suppose that $\hat{B}^{**} \neq B^{**}$. Then we may choose

$$x_0^{**} \in B^{**} \backslash \hat{B}^{**}.$$

A minor modification of the proof of Lemma 8.13 shows that we can choose a linear functional $f : X^{**} \to \mathbb{K}$, $\alpha \in \mathbb{R}$, and $\epsilon > 0$ such that f is $\sigma(X^{**}, X^*)$-continuous and

$$\mathrm{Re}(f(x^{**})) \le \alpha < \alpha + \epsilon \le \mathrm{Re}(f(x_0^{**})) \text{ for all } x^{**} \in \hat{B}^{**}.$$

By Proposition 14.20, we may choose $x_0^* \in X^*$ such that

$$f(x^{**}) = x^{**}(x_0^*) \text{ for all } x^{**} \in X^{**}.$$

Since $J[B] \subset \hat{B}^{**}$, we conclude that

$$\mathrm{Re}((x_0^*)(x)) \le \alpha < \alpha + \epsilon \le \mathrm{Re}(x^{**}(x_0^*) \text{ for all } x \in B.$$

Since $0 \in B$, we know that $\alpha > 0$. Let us put $y_0^* = \alpha^{-1} x_0^*$. Then we have

$$\mathrm{Re}(y_0^*(x)) \le 1 < 1 + \frac{\epsilon}{\alpha} \le \mathrm{Re}(x_0^{**}(y_0^*)) \text{ for all } x \in B. \tag{2}$$

Since $\gamma x \in B$ for all $x \in B$, $\gamma \in \mathbb{K}$ with $|\gamma\| \le 1$, the left-most inequality in (2) tells us that

$$|y_0^*(x)| \le 1 \text{ for all } x \in B,$$

and consequently $\|y_0^*\| \leq 1$ from which we conclude that

$$|x_0^{**}(y_0^*)| \leq 1. \tag{3}$$

Since $\alpha, \epsilon > 0$, the right-most inequality in (2) is inconsistent with (3). $\square$

Theorem 4.27: Let X be a normed linear space. Then X is reflexive if and only if $\overline{B}_1^X(0)$ is $\sigma(X, X^*)$-compact (i.e., if and only if the closed unit ball is weakly compact).

Proof: Assume that X is reflexive. Then, since $J : X \to X^{**}$ is a surjective isometry, we know that

$$J[\overline{B}_1^X(0)] = \overline{B}_1^{X^{**}}(0).$$

We also know that $J : (X, \sigma(X, X^*)) \to (X^{**}, \sigma(X^{**}, X^*))$ is a homeomorphism. By Alaoglu's Theorem, $\overline{B}_1^{X^{**}}(0)$ is $\sigma(X^{**}, X^*)$-compact. It follows that $\overline{B}_1^X(0)$ is $\sigma(X, X^*)$-compact.

Assume now that $\overline{B}_1^X(0)$ is $\sigma(X, X^*)$-compact. Then $J[\overline{B}_1^X(0)]$ is $\sigma(X^{**}, X^*)$-compact and hence $\sigma(X^{**}, X^*)$-closed. It follows from Goldstine's Theorem that $J[\overline{B}_1^X(0)] = \overline{B}_1^{X^{**}}(0)$. By linearity of J, we conclude that J is surjective and X is reflexive. $\square$

Theorem 4.28: Let X be a separable normed linear space and assume that $K \subset X$ is weakly compact. Then $(K, \sigma(X, X^*))$ is metrizable.

Sketch of Proof: We can construct a sequence $\{x_n^*\}_{n=1}^{\infty}$ in $X^* \backslash \{0\}$ such that

$$\bigcap_{n=1}^{\infty} \mathcal{N}(x_n^*) = \{0\}.$$

The metric $\rho : K \times K \to \mathbb{R}$ defined by

$$\rho(x, y) = \sum_{n=1}^{\infty} 2^{-n} \frac{|x_n^*(x - y)|}{1 + |x_n^*(x - y)|} \quad \text{for all } x, y \in K,$$

does the job. $\square$

Theorem 4.29: Let X be a normed linear space and assume that $C \subset X$ is convex. Then the weak and strong closures of C coincide.

Proof: Let us denote the strong closure of C by $\overline{C}$ and the weak closure of C by $\hat{C}$. Since every weakly set is closed, we know that

$$\overline{C} \subset \hat{C}.$$

To establish the reverse inclusion, let $x_0 \in \overline{C}$ be given. Using a Hahn-Banach argument, we may choose $x^* \in X^*$, $\alpha \in \mathbb{R}$, and $\epsilon > 0$ such that

$$\mathrm{Re}((x^*)(x)) \leq \alpha \leq \alpha < \alpha + \epsilon \leq \mathrm{Re}(x^*(x_0)) \quad \text{for all } x \in \overline{C}.$$

The set $\{x \in X : \mathrm{Re}(x^*(x)) < \alpha + \frac{\epsilon}{2}\}$ is a weak neighborhood of x_0 that has empty intersection with C. It follows that

$$\hat{C} \subset \overline{C}. \quad \square$$

Proposition 4.30:

(a) Let X be a normed linear space and let $E \subset X$. Then E is topologically bounded in $(X, \sigma(X, X^*))$ (i.e., E is weakly bounded) if and only if E is norm bounded.

(b) Let X be a Banach space and let $E^* \subset X^*$. Then E^* is topologically bounded in $(X, \sigma(X^*, X))$ (i.e., E^* is weakly* bounded) if and only if E^* is norm bounded.

The proof of Proposition 4.30 is part of Assignment 8.

Inner Products Spaces

We now study normed linear spaces with additional structure due to the presence of an inner product. An inner product is a natural generalization of the dot product in Euclidean space. It gives rise to a natural notion of orthogonality that leads to a very rich geometric structure.

Definition 14.31: Let X be a linear space over $\mathbb{K}$. By an *inner product* on X we mean a function $(\cdot,\cdot) : X \times X \to \mathbb{K}$ such that

(a) $\forall x, y, z \in X, \quad (x+y, z) = (x, z) + (y, z)$,

(b) $\forall x, y \in X, \alpha \in \mathbb{K}, \quad (\alpha x, y) = \alpha(x, y)$,

(c) $\forall x, y \in \mathbb{K}, \quad (x, y) = \overline{(y, x)}$,

(d) $\forall x \in X, \quad (x, x) \geq 0$ (notice that (c) implies $(x, x) \in \mathbb{R}$),

(e) $\forall x \in X, \quad (x, x) = 0 \Leftrightarrow x = 0$.

A linear space, together with an inner product, is called an *inner-product space.*

Remark 14.32: In the above definition, the bar in (c) denotes the complex conjugate. Of course, if $\mathbb{K} = \mathbb{R}$ then $\overline{\alpha} = \alpha$ for all $\alpha \in \mathbb{K}$. It follows from (b) and (c) that $(x, \alpha y) = \overline{\alpha}(x, y)$ for all $x, y \in X, \alpha \in \mathbb{K}$. It follows form (a) and (c) that $(x, y+z) = (x, y) + (x, z)$ for all $x, y, z \in X$. In other words, an inner product is linear in the first argument and conjugate linear in the second argument. Some authors (especially in the physics literature) define inner products to be conjugate linear in the first argument and linear in the second argument. This leads to a number of small changes in the formulas that follow.

Example 14.33:

(a) $X = l^2$, $(x, y) = \sum_{k=1}^{\infty} x_k \overline{y_k}$.

(b) Let X be the set of all real or complex-valued continuous functions on $[0, 1]$. We define an inner product $(\cdot, \cdot) : X \times X \to \mathbb{K}$ by

$$(f, g) = \int_0^1 f(t)\overline{g}(t)dt.$$

Proposition 14.34 (Cauchy-Schwarz Inequality): Let X be an inner-product space and let $x, y \in X$ be given. Then

$$|(x, y)|^2 \leq (x, x)(y, y).$$

Proof: If $y = 0$ there is nothing to prove, so we assume that $y \neq 0$. Let $\lambda \in \mathbb{K}$ be given and observe that

$$0 \leq (x - \lambda y, x - \lambda y) = (x, x) - 2\text{Re}[\overline{\lambda}(x, y)] + |\lambda|^2 (y, y). \tag{4}$$

Rearranging (4) we obtain

$$2\text{Re}[\overline{\lambda}(x, y)] \leq (x, x) + |\lambda|^2 (y, y) \tag{5}$$

Let us put

$$\lambda = \frac{(x, y)}{(y, y)}. \tag{6}$$

Substituting (6) into (5) we find that

$$2\frac{|(x, y)|^2}{(y, y)} \leq (x, x) + \frac{|(x, y)|^2}{(y, y)}. \tag{7}$$

The desired conclusion follows from multiplying both sides of (7) by (y, y). $\square$

Definition 14.35: Let X be an inner product space with inner product $(\cdot, \cdot)$. We define the *norm associated with* $(\cdot, \cdot)$ by

$$\|x\| = \sqrt{(x, x)} \quad \text{for all } x \in X. \tag{8}$$

Remark 14.36: Using the norm associated with the inner product, the Cauchy-Schwarz inequality can be rewritten as

$$|(x, y)| \leq \|x\|\|y\| \quad \text{for all } x, y \in X. \tag{9}$$

To see that the function $\|\cdot\|$ defined by (8) is, in fact, a norm, we need to verify the triangle inequality. (The other properties of a norm are obviously satisfied. To this end, let $x, y \in X$ be given Observe that

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\mathrm{Re}(x,y). \tag{10}$$

Since $\mathrm{Re}(x,y) \leq |(x,y)|$, the Cauchy-Schwarz inequality (9) implies that

$$\|x+y\|^2 \leq \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| = (\|x\| + \|y\|)^2,$$

and we conclude that $\|x+y\| \leq \|x\| + \|y\|$ for all $x, y \in X$.

When we speak of the norm in an inner-product space, we always mean the norm associated with the inner product, unless stated otherwise. Topological and metric concepts, such as continuity and distance, are understood to be defined in terms of the norm associated with the inner product and the metric associated with that norm.

Remark 14.37: Equation (10) is called the *polar identity*. Some authors use the term "polar identity" for an identity that follows from (10).

Definition 14.38: Let X be an inner-product space and let $x, y \in X$ and $A, B \subset X$ be given.

(a) We say that x is *orthogonal* to y and write $x \perp y$ provided that $(x,y) = 0$.

(b) We say that x is *orthogonal* to A and write $x \perp A$ provided that $(x,y) = 0$ for all $y \in A$.

(c) We say that A is *orthogonal* to B and write $A \perp B$ provided that $(x,y) = 0$ for all $x \in A, y \in B$.

(d) The *orthogonal complement* of A is defined by $A^\perp = \{y \in X : (x,y) = 0 \text{ for all } x \in A\}$.

The following well-known result is an immediate consequence of the polar identity (10).

Proposition 14.39 (Pythagorean Theorem): Let X be an inner product space and let $x, y \in X$ with $x \perp y$ be given. Then

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2.$$

If we apply the polar identity to compute $\|x+y\|^2 + \|x-y\|^2$, the "cross terms" cancel and we obtain the following result.

Proposition 14.40 (Parallelogram Law): Let X be an inner product space and let $x, y \in X$ be given. Then we have

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2). \tag{11}$$

The parallelogram law characterizes norms that come from inner products.

Proposition 14.41: Let $(X, \|\cdot\|)$ be a normed linear space and assume that (11) holds for all $x, y \in X$. Then there is an inner product $(\cdot,\cdot) : X \times X \to \mathbb{K}$ such that $\|x\| = \sqrt{(x,x)}$ for all $x \in X$.

Proof: **Case 1**: $\mathbb{K} = \mathbb{R}$

Define $(\cdot,\cdot) : X \times X \to \mathbb{R}$ by

$$(x, y) = \frac{1}{4}\{\|x+y\|^2 - \|x-y\|^2\} \text{ for all } x, y \in X.$$

It is straightforward, although a bit tedious, to check that $(\cdot,\cdot)$ is an inner product.

Case 2: $\mathbb{K} = \mathbb{C}$

Define $(\cdot,\cdot) : X \times X \to \mathbb{C}$ by

$$(x, y) = \frac{1}{4}\{\|x+y\|^2 - \|x-y\|^2 + i[\|x+iy\|^2 - \|x-iy\|^2]\} \text{ for all } x, y \in X.$$

It is straightforward, but rather tedious, to check that $(\cdot,\cdot)$ is an inner product. $\square$

Remark 14.42: Although most of the results described here are valid in both real and complex inner-product spaces, you should be aware that certain results of interest are valid in real spaces, but not complex ones, or vice versa. Two such examples are given below. Here, X is an inner-product space of dimension ≥ 2.

- If $x, y \in X$ satisfy $\|x+y\|^2 = \|x\|^2 + \|y\|^2$, can we conclude that $x \perp y$? The answer is yes if $\mathbb{K} = \mathbb{R}$, but no if $\mathbb{K} = \mathbb{C}$.
- If $A \in \mathcal{L}(X; X)$ is such that $(Ax, x) = 0$ for all $x \in X$, can we conclude that $A = 0$? The answer is yes if $\mathbb{K} = \mathbb{C}$, but no if $\mathbb{K} = \mathbb{R}$.

Most of the deep results about inner-product spaces, require the space to be complete.

Definition 14.43: An inner-product space that is complete with respect to the norm associated with the inner product is called a *Hilbert space*.

Lecture Notes for Week 15 (Preliminary Draft)

Orthogonal Projections

Theorem 15.1: Let K be a nonempty closed convex subset of a Hilbert space X and let $x \in X$ be given. Then there exists exactly one point $y_0 \in K$ such that

$$\|x - y_0\| = \inf\{\|x - y\| : y \in K\}.$$

Proof: Without loss of generality, we may assume that $x = 0$. (Indeed, the set $-x + K$ is nonempty closed and convex.) Let us put

$$\gamma = \inf\{\|y\| : y \in K\}.$$

Choose a sequence $\{y_n\}_{n=1}^{\infty}$ such that $y_n \in K$ for all $n \in \mathbb{N}$ and such that $\|y_n\| \to \gamma$ as $n \to \infty$. By the parallelogram law, we have

$$\|\frac{1}{2}(y_n - y_m)\|^2 = \frac{1}{2}(\|y_n\|^2 + \|y_m\|^2) - \|\frac{1}{2}(y_n + y_m)\|^2. \tag{1}$$

Since K is convex, $\frac{1}{2}(y_n + y_m) \in K$ and consequently

$$\|\frac{1}{2}(y_n + y_m)\|^2 \geq \gamma^2. \tag{2}$$

Let $\epsilon > 0$ be given. Choose $N \in \mathbb{N}$ such that

$$\|y_n\|^2 < \gamma^2 + \frac{1}{4}\epsilon^2 \text{ for all } n \geq N. \tag{3}$$

Combining (1), (2), and (3) we find that

$$\|y_n - y_m\|^2 < \epsilon^2 \text{ for all } m, n \geq N,$$

and consequently $\{y_n\}_{n=1}^{\infty}$ is a Cauchy sequence.

Since X is complete, we may choose $y_0 \in X$ such that $y_n \to y_0$ as $n \to \infty$. We deduce that $y_0 \in K$ because K is closed. Since the mapping $x \mapsto \|x\|$ is continuous, we infer that $\|y_0\| = \gamma$.

To prove that there is only one such point y_0, suppose that $\hat{y}_0 \in K$ satisfies $\|\hat{y}_0\| = \gamma$. Then, since K is convex, we have $\frac{1}{2}(y_0 + \hat{y}_0) \in K$. This implies that

$$\gamma \leq \|\frac{1}{2}(y_0 + \hat{y}_0)\| \leq \frac{1}{2}(\|y_0\| + \|\hat{y}_0\|) = \gamma,$$

and consequently

$$\|\frac{1}{2}(y_0 + \hat{y}_0)\| = \gamma. \tag{4}$$

By the parallelogram law, we have

$$\|\frac{1}{2}(y_0 + \hat{y}_0)\|^2 + \|\frac{1}{2}(y_0 - \hat{y}_0)\|^2 = \gamma^2. \tag{5}$$

Combining (4) and (5) we deduce that $\|y_0 - \hat{y}_0\| = 0$. $\square$

Theorem 15.2: Let X be a Hilbert space and M be a closed subspace of X. Let $x \in X$ be given and let y_0 be the unique element of K such that $\|x - y_0\| \leq \|x - y\|$ for all $y \in M$. Then $x - y_0 \in M^\perp$. Conversely, if $\hat{y}_0 \in M$ and $x - \hat{y}_0 \in M^\perp$ then $\|x - \hat{y}_0\| \leq \|x - y\|$ for all $y \in M$.

Proof: Let $y \in M$ and $\alpha \in \mathbb{K}$ be given. Then $y_0 + \alpha y \in M$ so that

$$\begin{aligned}\|x - y_0\|^2 &\leq \|x - (y_0 + \alpha y)\|^2 = \|(x - y_0) - \alpha y\|^2 \\ &\leq \|x - y_0\|^2 + |\alpha|^2\|y\|^2 - 2\mathrm{Re}(x - y_0, \alpha y).\end{aligned}$$

It follows that

$$2\mathrm{Re}(x - y_0, \alpha y) \leq |\alpha|^2\|y\|^2. \tag{6}$$

Putting $\alpha = t > 0$ in (6) yields

$$2t\mathrm{Re}(x - y_0, y) \leq t^2\|y\|^2. \tag{7}$$

Dividing (7) by t and letting $t \downarrow 0$ we arrive at

$$\mathrm{Re}(x - y_0, y) \leq 0.$$

Since $-y \in M$ we also have $\mathrm{Re}(x - y_0, y) \geq 0$ and consequently

$$\mathrm{Re}(x - y_0, y) = 0. \tag{8}$$

If $\mathbb{K} = \mathbb{R}$, then $(x - y_0, y) = 0$ and $x - y_0 \in M^\perp$.

Suppose that $\mathbb{K} = \mathbb{C}$. Then we can replace y with iy in (8) and use the fact that

$$\mathrm{Re}(u, iv) = \mathrm{Im}(u, v) \quad \text{for all } u, v \in X$$

to deduce that

$$\mathrm{Im}(x - y_0, y) = 0.$$

We conclude that $(x - y_0, y) = 0$ and that $x - y_0 \in M^\perp$.

To prove the second claim in the theorem, suppose that $\hat{y}_0 \in M$ and $x - \hat{y}_0 \in M^\perp$. Then, for all $y \in M$ we have

$$\begin{aligned}\|x - y\|^2 &= \|(x - \hat{y}_0) + (\hat{y}_0 - y)\|^2 \\ &= \|x - \hat{y}_0\|^2 + \|\hat{y}_0 - y\|^2 \geq \|x - \hat{y}_0\|^2,\end{aligned}$$

since $\hat{y}_0 - y \in M$. $\square$

Remark 15.3: Dealing with inequalities involving inner products and norms that hold for all elements of a certain subspace is, of course, very important. For this reason, we point out a slightly different way of obtaining the first claim in Theorem 15.2. If $y = 0$, then trivially $(x - y_0, y) = 0$, so we may assume that $y \neq 0$. If we put

$$\alpha = \frac{(x - y_0, y)}{\|y\|^2}$$

in (6) then we get

$$\frac{|(x - y_0, y)|^2}{\|y\|^2} \leq 0,$$

which implies that $(x - y_0, y) = 0$. This argument is a bit shorter than the one given above, but probably a bit trickier to discover.

An important consequence of Theorem 15.2 is that every element of X can be written in precisely one way as the sum of an element of M and an element of $M^\perp$. This result is known as the Projection Theorem.

Corollary 15.4 (Projection Theorem): Let X be a Hilbert space, M be a closed subspace of X and $x \in X$ be given. Then there exists exactly one pair $(y, z) \in M \times M^\perp$ such that $x = y + z$.

For a fixed closed subspace M, the mapping that carries $x \in X$ to the unique point y_0 in M that minimizes the distance to x is linear. (You should verify this fact as a simple exercise for yourself.) We refer to this mapping as the orthogonal projection onto M and denote it by P_M. Notice that $P_{M^\perp} = I - P_M$.

Definition 15.5: Let M be a closed subspace of a Hilbert space X. For each $x \in X$ we define $P_M x$ to be the unique element of M such that

$$\|x - P_M x\| \leq \|x - y\| \text{ for all } y \in M.$$

Duality in Hilbert Spaces

Lemma 15.6: Let X be a Hilbert space and M be a closed subspace of X. Assume that $M \neq X$. Then there exists $z \in M^\perp$ such that $z \neq 0$.

Proof: Choose $x \in X \backslash M$. By the Projection Theorem, we may choose $y \in M, z \in M^\perp$ such that $x = y + z$. Since $x \notin M$, we know that $x - y \neq 0$ and consequently $z \neq 0$. $\square$

One of the most important results concerning Hilbert spaces, is that a Hilbert space can be identified with its own dual.

Theorem 15.7 (Riesz Representation Theorem): Let X be a Hilbert space and $x^* \in X^*$ be given. Then there exists exactly one $y \in X$ such that

$$x^*(x) = (x, y) \quad \text{for all } x \in X.$$

Moreover $\|x^*\|_* = \|y\|$.

Proof: Put $M = \mathcal{N}(x^*)$ and observe that M is a closed subspace of X. If $M = X$ we are done because $x^* = 0$ and $y = 0$ is the unique element of X that does the job. Assume that $M \neq X$. Then, by Lemma 15.6, we may choose $y_0 \in M^\perp \backslash \{0\}$. Let us put $\alpha = x^*(y_0)$.

Let $x \in X$ be given and observe that

$$x - \frac{x^*(x)}{\alpha} y_0 \in \mathcal{N}(x^*) = M.$$

Since $y_0 \in M^\perp$, we conclude that

$$(x - \frac{x^*(x)}{\alpha} y_0, y_0) = 0,$$

and consequently

$$x^*(x)(\alpha^{-1} y_0, y_0) = (x, y_0). \tag{9}$$

Let us put

$$y = \frac{\overline{\alpha} y_0}{(y_0, y_0)}.$$

Substitution into (9) gives

$$x^*(x) = (x, y).$$

Suppose that $z \in X$ is such that $(x, y) = (x, z)$ for all $x \in X$. Then $(x, y - z) = 0$ for all $x \in X$, so that $(y - z, y - z) = 0$ and $z = y$.

It remains only to show that $\|x^*\|_* = \|y\|$. Observe that

$$\begin{aligned} \|x^*\|_* &= \sup\{|x^*(x)| : x \in X, \|x\| \leq 1\} \\ &= \sup\{|(x, y)| : x \in X, \|x\| \leq 1\} \\ &\leq \|y\|, \end{aligned}$$

by virtue of the Cauchy-Schwarz inequality. To establish the reverse inequality, we assume that $y \neq 0$. (If $y = 0$ there is nothing to prove.) Then we have

$$\|y\|^2 = (y, y) = x^*(y) \leq \|x^*\|_* \|y\|,$$

which gives $\|y\| \leq \|x^*\|_*$. $\Box$

A very important consequence of the Riesz Representation Theorem is that **every Hilbert space is reflexive**. (You should verify the details of a proof for yourself).

Another important consequence of the Riesz Representation Theorem is the following simple remark concerning weak convergence.

Remark 15.8: Let X be a Hilbert space, $x \in X$ and $\{x_n\}_{n=1}^{\infty}$ be a sequence in X. The following three statements are equivalent.

(i) $x_n \rightharpoonup x$ (weakly) as $n \to \infty$.

(ii) $\forall y \in X, \quad (x_n, y) \to (x, y)$ as $n \to \infty$.

(iii) $\forall y \in X, \quad (y, x_n) \to (y, x)$ as $n \to \infty$.

In applications, one frequently constructs a sequence of approximate solutions to some problem and shows that this sequence (or a subsequence) is weakly convergent. In order to pass to the limit and obtain a solution to the original problem, it is very helpful if it can be shown the weakly convergent sequence is actually strongly convergent. The following proposition gives a simple criterion that can sometimes be used to show that a weakly convergent sequence is actually strongly convergent.

Proposition 15.9: Let X be a Hilbert space, $x \in X$ and $\{x_n\}_{n=1}^{\infty}$ be a sequence in X. The following two statements are equivalent.

(i) $x_n \to x$ (strongly) as $n \to \infty$.

(ii) $x_n \rightharpoonup x$ (weakly) as $n \to \infty$ and $\|x_n\| \to \|x\|$ as $n \to \infty$.

Proof: The implication (i) $\Rightarrow$ (ii) is clear. Assume that (ii) holds. Then for all $n \in \mathbb{N}$ we have

$$\|x_n - x\|^2 = (x_n - x, x_n - x) = \|x_n|^2 + \|x\|^2 - 2\text{Re}(x_n, x). \tag{10}$$

As $n \to \infty$, $\|x_n\|^2 \to \|x\|^2$ and $\text{Re}(x_n, x) \to \text{Re}(x, x) = \|x\|^2$, and consequently, the right-hand side of (10) tends to 0. $\square$

Remark 15.10: Proposition 15.9 is valid for *uniformly convex* Banach spaces. (See Assignment 6 for the definition of uniform convexity.)

Hilbert Adjoints

Let X be a Hilbert space and $A \in \mathcal{L}(X; X)$ be given. Since X^* can be identified with X via the Riesz Representation Theorem, it is useful to modify the definition of adjoint so that the domain of the adjoint is the space X itself, rather than the dual space X^*. We shall temporarily use the notation A_H^* for this new type of adjoint, so that we can define it in terms of A^* and deduce its properties from previous results in a simple and clear-cut manner. After we have developed the basic properties of Hilbert adjoints, we shall drop the subscript "H".

We begin by defining the mapping $R : X \to X^*$ by

$$(R(y))(x) = (x, y) \quad \text{for all } x, y \in X \tag{11}$$

Observe that R is *conjugate linear* (i.e. $R(y+z) = R(y) + R(z)$ and $R(\alpha y) = \overline{\alpha} R(y)$ for all $y, z \in X, \alpha \in X$) and isometric (i.e. $\|R(y)\|_* = \|y\|$ for all $y \in X$).

Definition 15.11: We define the *Hilbert adjoint* A_H^* of A by the formula

$$A_H^* = R^{-1} A^* R. \tag{12}$$

It is straightforward to check that $A_H^* \in \mathcal{L}(X; X)$ and that

$$(Ax, y) = (x, A_H^*, y) \quad \text{and} \quad (A_H^* x, y) = (x, Ay) \quad \text{for all } x, y \in X. \tag{13}$$

Some of the basic properties of Hilbert adjoints are summarized in the following theorem.

Theorem 15.12: Let X be a Hilbert space and let $A, B \in \mathcal{L}(X; X)$ and $\alpha \in \mathbb{K}$ be given. Then:

(a) $(\alpha A + B)_H^* = \overline{\alpha} A_H^* + B_H^*$,

(b) $(AB)_H^* = B_H^* A_H^*$,

(c) $(A_H^*)_H^* = A$,

(d) A is bijective if and only if A_H^* is bijective; in this case we have $(A_H^*)^{-1} = (A^{-1})_H^*$,

(e) $\|A_H^*\| = \|A\|$,

(f) If there exists $c > 0$ such that $\|A_H^* x\| \geq c\|x\|$ for all $x \in X$ then A is surjective,

(g) $\mathcal{R}(A)$ is closed if and only if $\mathcal{R}(A_H^*)$ is closed,

(h) A is compact if and only if A_H^* is compact.

Proposition 15.13: Let X be a Hilbert space and $A \in \mathcal{L}(X; X)$. Then

$$\|A\|^2 = \|A_H^* A\|. \tag{14}$$

Proof: Let $x \in X$ with $\|x\| \leq 1$ be given. Then we have

$$\|Ax\|^2 = (Ax, Ax) = (A_H^* Ax, x) \leq \|A_H^* Ax\| \|x\| \leq \|A_H^* A\|.$$

Taking the supremum over all $x \in X$ with $\|x\| \leq 1$ we find that $\|A\|^2 \leq \|A_H^* A\|$. On the other hand, we have $\|A_H^* A\| \leq \|A_H^*\| \|A\| = \|A\|^2$. □

Proposition 15.14 Let X be a Hilbert space $A \in \mathcal{L}(X; X)$ and $\{A_n\}_{n=1}^{\infty}$ be a sequence in $\mathcal{L}(X; Y)$.

(a) If $A_n \to A$ in the uniform operator topology as $n \to \infty$ then $(A_n)_H^* \to A_H^*$ in the uniform operator topology.

(b) If $A_n \to A$ in the weak operator topology as $n \to \infty$ then $(A_n)_H^* \to A_H^*$ in the weak operator topology.

In order to translate the results concerning the relationship between the ranges and null spaces of operators and their adjoints over to Hilbert adjoints, we need to examine the relationship between annihilators, pre-annihilators, and orthogonal complements. For this purpose, we temporarily put a subscripted "H" on the orthogonal complement of a set.

Let X be a Hilbert space, $M \subset X$ and $Z \subset X^*$. We put

$$M_H^{\perp} = \{y \in X : (x, y) = 0 \text{ for all } x \in X\},$$

$$M^{\perp} = \{x^* \in X^* : x^*(x) = 0 \text{ for all } x \in X\},$$
$$^{\perp}Z = \{x \in X : x^*(x) = 0 \text{ for all } x^* \in Z\}.$$

With these definitions we have

$$M^{\perp} = R[M_H^{\perp}],$$

$$^{\perp}Z = (R^{-1}[Z])_H^{\perp},$$

where R is defined by (11).

Theorem 15.15: Let X be a Hilbert space and $A \in \mathcal{L}(X; X)$ be given. Then

(a) $\mathcal{N}(A) = (\mathcal{R}(A_H^*))_H^{\perp}$,

(b) $\mathcal{N}(A_H^*) = (\mathcal{R}(A))_H^{\perp}$,

(c) $\text{cl}(\mathcal{R}(A)) = (\mathcal{N}(A_H^*))_H^{\perp}$,

(d) $\mathrm{cl}(\mathcal{R}(A_H^*)) = (\mathcal{N}(A))_H^\perp$.

Orthonormal Families and Bases

Throughout this section, X is a given Hilbert space. Families of nonzero vectors that are pairwise orthogonal have many nice properties. It is convenient to normalize the vectors in such a family so that they each have norm one.

Definition 15.16: A family $(e_i|i \in I)$ of elements of X is said to be *orthonormal* provided that $e_i \perp e_j$ for all $i, j \in I$ with $i \neq j$ and $\|e_i\| = 1$ for all $i \in I$.

Definition 15.17: An orthonormal family $(e_i|i \in I)$ is said to be *maximal* provided that

$$\forall x \in X, \quad (x \perp e_i \text{ for all } i \in I) \Rightarrow x = 0.$$

A maximal orthonormal family is called an *orthonormal basis*.

Remark 15.18: We say that a set $\mathcal{E} \subset X$ is orthonormal, or is an orthonormal basis, provided that the family $(e|e \in \mathcal{E})$ has the desired property. Notice that an orthonormal set $\mathcal{E}$ is an orthonormal basis if and only if it is a maximal orthonormal set with respect to set inclusion.

Proposition 15.19: Let $\mathcal{O}$ be an orthonormal subset of X. Then there is an orthonormal basis $\mathcal{E} \subset X$ such that $\mathcal{O} \subset \mathcal{E}$.

The proof of Proposition 15.19 is a straightforward application of Zorn's Lemma. It is not difficult to show that all orthonormal bases for X have the same cardinality.

A major advantage of orthonormal families is that there is a simple formula for the coefficients appearing in linear combinations. More specifically, suppose that $(e_i|i \in I)$ is an orthonormal family, F is a finite subset of I, and that

$$x = \sum_{i \in F} \alpha_i e_i. \tag{15}$$

For each $j \in F$, we take the inner product with e_j in (15) to obtain

$$(x, e_j) = \alpha_j. \tag{16}$$

It follows easily that every orthonormal family is linearly independent. Indeed, suppose that (15) holds with $x = 0$. Then (16) implies that $\alpha_j = 0$ for all $j \in F$.

Proposition 15.20: Let $(e_i|i = 1, 2, \cdots, N)$ be an orthonormal family in X and let $M = \mathrm{span}(e_1, e_2, \cdots, e_n)$. Then

$$P_M x = \sum_{k=1}^{N} (x, e_k) e_k \quad \text{for all } x \in X. \tag{17}$$

Here P_M is the orthogonal projection onto M.

Proof: Let us put

$$Qx = \sum_{k=1}^{N}(x, e_k)e_k \quad \text{for all } x \in X.$$

We need to show that $Q = P_M$. Clearly $Qx \in M$ for all $x \in X$. Therefore, it suffices to show that $x - Qx \in M^{\perp}$ for all $x \in X$. Let $x \in X$ be given. For each $j = 1, 2, \cdots, N$ we have

$$(x - Qx, e_j) = (x, e_j) - (x, e_j) = 0,$$

and consequently $(x - Qx, y) = 0$ for all $y \in M$. □

Proposition 15.21 (Gram-Schmidt): Let $(w_i|i \in \mathbb{N})$ be a linearly independent family of elements of X. Then there is an orthonormal family $(e_i|i \in \mathbb{N})$ such that

$$\mathrm{span(e_1, e_2, \cdots, e_n)} = \mathrm{span}(w_1, w_2, \cdots, w_n) \quad \text{for all } n \in \mathbb{N}.$$

Proposition 15.22 (Bessel's Inequality): Let $\{e_k\}_{k=1}^{\infty}$ be an orthonormal sequence. Then

$$\sum_{k=1}^{\infty} |(x, e_k)|^2 \leq \|x\|^2 \quad \text{for all } x \in X. \tag{18}$$

Proof: Let $n \in \mathbb{N}$ be given and put

$$y_n = x - \sum_{k=1}^{n}(x, e_k)e_k.$$

Observe that $(y, e_k) = 0$ for all $k = 1, 2, \cdots, n$. By the Pythagorean Theorem, we have

$$\begin{aligned} \|x\|^2 &= \|y_n\|^2 + \left\| \sum_{k=1}^{n}(x, e_k)e_k \right\|^2 \\ &= \|y_n\|^2 + \sum_{k=1}^{n}|(x, y_k)|^2 \\ &\geq \sum_{k=1}^{n}|(x, e_k)|^2. \end{aligned}$$

Letting $n \to \infty$ gives the desired result. □

Corollary 15.23: Let $(e_i|i \in I)$ be an orthonormal family and let $x \in X$ be given. Then $\{i \in I : (x, e_i) \neq 0\}$ is countable.

The proof of the corollary follows from the observation that for each $n \in \mathbb{N}$, $\{i \in I : |(x, e_i)| \geq \frac{1}{n}\}$ is finite, by virtue of Bessel's Inequality.

Corollary 15.24: Let $(e_i|i \in I)$ be an orthonormal family and let $x \in X$ be given. Then

$$\sum_{i \in I} |(x, e_i)|^2 < \infty.$$

There is, of course, no difficulty in interpreting the sum in the last corollary because all of the terms are nonnegative (and only countably many are nonzero). Now it is natural to ask if we can make sense of the sum

$$\sum_{i \in I} (x, e_i) e_i, \tag{19}$$

when $(e_i|i \in I)$ is an orthonormal family and $x \in X$. The answer is yes, but we must be a bit careful in assigning a precise meaning to the sum. If I is finite, there is nothing to worry about. If I is infinite, then there are only countably many nonzero terms in the sum, but we need to worry about whether the order in which the terms are arranged could matter. We begin by giving a result when the index set is $\mathbb{N}$.

Theorem 15.25: Let X be a Hilbert space, $(e_i|i \in \mathbb{N})$ be an orthonormal family, and put

$$M = \text{span}(\text{cl}(e_i|i \in I)).$$

Then for every $x \in X$ we have

$$\sum_{i=1}^{\infty} (x, e_i) e_i = P_M x.$$

Definition 15.26: Let Y be a normed linear space and $(y_i|i \in I)$ be a family of elements of Y. We say that $(y_i|i \in I)$ is *summable* with sum $S \in Y$, and we write

$$\sum_{i \in I} y_i = S,$$

provided that for every $\epsilon > 0$ there exists a finite set $F \subset I$ such that for every finite set J with $F \subset J \subset I$ we have

$$\left\| \sum_{i \in J} y_i - S \right\| < \epsilon.$$

Remark 15.27: A family $(a_i|i \in I)$ of real numbers is summable if and only if it is absolutely summable; in this case there can be only a countable of nonzero terms.

Theorem 15.28: Let X be a Hilbert space and $(e_i|i \in I)$ be an orthonormal family. The following 5 statements are equivalent.

(i) $(e_i|i \in I)$ is an orthonormal basis,

(ii) $\mathrm{cl}(\mathrm{span}(e_i|i \in I)) = X$,

(iii) $\forall x \in X, \quad x = \sum_{i \in I}(x, e_i)e_i$,

(iv) $\forall x, y \in X, \quad (x, y) = \sum_{i \in I}(x, e_i)\overline{(y, e_i)}$,

(v) $\forall x \in X, \quad \|x\|^2 = \sum_{i \in I} |(x, e_i)|^2$

Proposition 15.29: Let X be a Hilbert space. X is separable if and only if it has a countable orthonormal basis.

Theorem 15.30: Let X be a Hilbert space and $A \in \mathcal{C}(X; X)$ be given. Then there is a sequence $\{A_n\}_{n=1}^{\infty}$ in $\mathcal{L}(X; X)$ such that each A_n has finite rank and $A_n \to A$ in the uniform operator topology as $n \to \infty$.

Sketch of Proof: Let $Y = \mathrm{cl}(\mathcal{R}(A))$. If Y is finite dimensional, then we may take $A_n = A$ for all $n \in \mathbb{N}$. Assume that Y is infinite dimensional. By Theorem $\mathcal{R}(A)$ is separable, and consequently Y is also separable. By Prop. 41.5, we may choose an orthonormal basis $(e_i|i \in \mathbb{N})$ for Y. For each $n \in \mathbb{N}$, let us put

$$M_n = \mathrm{span}(e_1, e_2, \cdots, e_n), \quad A_n = P_{M_n}A.$$

Clearly, each A_n has finite rank. Using the fact that A is compact, it is not too difficult to show that

$$\|A_n - A\| \to 0 \quad \text{as } n \to \infty.$$

You should check the details as an exercise for yourself. $\square$

Remark 15.31: Theorem 15.30 is valid for any Banach space X having a Schauder basis.

Spectral Theory for Compact Self-Adjoint Operators

Let X be a Hilbert space.

Definition 15.32: Let $A \in \mathcal{L}(X; X)$ be given.

(a) The *resolvent set* of A, denoted $\rho(A)$ is the set of all $\alpha \in \mathbb{K}$ such that $(\alpha I - A)$ is bijective.

(b) The *spectrum* of A, denoted $\sigma(A)$ is defined by $\sigma(A) = \mathbb{K} \backslash \rho(A)$.

(c) Let $\lambda \in \mathbb{K}$ be given. λ is said to be an *eigenvalue* of A provided that $\mathcal{N}(\lambda I - A) \neq \{0\}$. The set of all eigenvalues of A is called the *point spectrum* of A and is denoted by $\sigma_p(A)$.

(d) Let $\lambda \in \sigma_p(A)$ be given. The nonzero members of $\mathcal{N}(A)$ are called *eigenvectors* of A.

Proposition 15.33: Let $A \in \mathcal{L}(X;X)$ and $\lambda \in \mathbb{K}$ be given. If $\lambda > \|A\|$ then $\lambda \in \rho(A)$.

Proposition 15.34: Let $A \in \mathcal{C}(X;X)$ be given and $\lambda \in \mathbb{K}\backslash\{0\}$ be given. Then $\mathcal{N}(\lambda I - A)$ is finite dimensional.

Proposition 15.35: Let $A \in \mathcal{C}(X;X)$ be given. Then $\sigma(A) \subset \sigma_p(A) \cup \{0\}$. Moreover $\sigma_p(A)$ is countable and 0 is the only possible accumulation point.

Theorem 15.36: Let $A \in \mathcal{C}(X;X)$ be given and assume that A is self-adjoint, i.e. $A_H^* = A$. Then

(a) $\sigma(A) \subset \mathbb{R}$

(b) there exists $\lambda \in \sigma_p(A)$ such that $|\lambda| = \|A\|$.

(c) $\mathcal{N}(\lambda - a) \perp \mathcal{N}(\mu - A)$ for all $\lambda, \mu \in \mathbb{K}$ with $\lambda \neq \mu$.

(d) there is an orthonormal basis $(e_i | i \in I)$ such e_i is an eigenvector of A for every $i \in I$. Moreover for every such basis we have

$$\forall x \in X, \quad Ax = \sum_{i \in I} \lambda_i (x, e_i) e_i, \quad \text{where } Ae_i = \lambda_i e_i \text{ for all } i \in I.$$

Zorn's Lemma

In order to prove the Hahn-Banach theorem, we shall make use of a result known as *Zorn's Lemma*.

Zorn's lemma is equivalent to the *axiom of choice* in the sense that the Zermelo-Fraenkel axioms of set theory, together with the axiom of choice imply Zorn's lemma and the Zermelo-Fraenkel axioms, together with Zorn's lemma imply the axiom of choice. Zorn's lemma is more convenient to use than the axiom of choice in a number of important situations. In particular, Zorn's lemma is well adapted to proving the existence of a Hamel basis in an arbitrary linear space. It is also useful for proving Tychonoff's theorem in topology, the existence of maximal ideals in rings, the existence of algebraic closures of fields, and the existence of maximal extensions of solutions of differential equations.

Before we can give a precise statement of the lemma, we need some definitions concerning *partially ordered sets*.

Definition Z.1: A *partially ordered set* is a pair $(\mathcal{M}, \leq)$ where $\mathcal{M}$ is a set and $\leq$ is a binary relation on $\mathcal{M}$ satisfying the following three conditions:

(i) $\forall x \in \mathcal{M}$, we have $x \leq x$,

(ii) $\forall x, y \in \mathcal{M}$, if $x \leq y$ and $y \leq x$ then $x = y$,

(iii) $\forall x, y, z \in \mathcal{M}$, if $x \leq y$ and $y \leq z$ then $x \leq z$.

The relation $\leq$ is called a *partial order* on $\mathcal{M}$.

Definition Z.2: Let $(\mathcal{M}, \leq)$ be a partially ordered set and let $x, y \in \mathcal{M}$ be given. We say that x *precedes* y (or equivalently that y *follows* x) provided $x \leq y$. We say that x and y are *comparable* provided that x precedes y or y precedes x. If neither x precedes y nor y precedes x we say that x and y are *incomparable*.

An obvious example of a partially ordered set is $(\mathbb{R}, \leq)$, the real numbers together with the usual notion of less than or equal to. Here we have the additional property that for all $x, y \in \mathbb{R}$, x and y are comparable. It is easy to produce a partially ordered set with incomparable elements.

Example Z.3: Let S be any set and let $\mathcal{M} = \mathcal{P}(S)$, the *power set* of S (i.e., the collection of all subsets of S). Define the relation $\leq$ on $\mathcal{M}$ by $A \leq B$ if and only if $A \subset B$. It is easy to see that $(\mathcal{M}, \leq)$ is a partially ordered set. If S contains two or more elements, then there are incomparable elements of $\mathcal{M}$.

Definition: Let $(\mathcal{M}, \leq)$ be a partially ordered set. A set $\mathcal{C} \subset \mathcal{M}$ is said to be a *chain* (or a *totally ordered subset*) provided that for all $x, y \in \mathcal{C}$, x and y are comparable (i.e., for all $x, y \in \mathcal{C}$, we have $x \leq y$ or $y \leq x$).

Definition Z.4: Let $(\mathcal{M}, \leq)$ be a partially ordered set and $\mathcal{A} \subset \mathcal{M}$. An element $x \in \mathcal{M}$ is called an *upper bound* for $\mathcal{A}$ provided that $y \leq x$ for all $y \in \mathcal{A}$.

Definition Z.5: Let $(\mathcal{M}, \leq)$ be a partially ordered set. An element $z \in \mathcal{M}$ is called a *maximal element* provided that for all $x \in \mathcal{M}$ if $z \leq x$ then $z = x$. In other words, a maximal element is not followed by an element other than itself.

Remark Z.6: A maximal element of $(\mathcal{M}, \leq)$ need not be an upper bound for $\mathcal{M}$.

Lemma Z.7 (Zorn's Lemma): Let $(\mathcal{M}, \leq)$ be a partially ordered set and assume that every chain has an upper bound. Then $(\mathcal{M}, \leq)$ has a maximal element.