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

$$U_\epsilon = B^X_\epsilon(0), \quad V_\delta = B^Y_\delta(0).$$

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

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

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

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

Let us put

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

so that

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

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

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

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

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 $y_1, y_2 \in B^Y_\delta(y_0)$ 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

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

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

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

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

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

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

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

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

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

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

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

It follows that

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

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

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

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 $m \leq M$. Choose c with $m \leq c \leq M$. Then we have

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

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

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

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

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

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

We conclude that

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

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

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

Multiplying through by α and using (ii) we obtain

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

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 $\mathcal{D}(F) = X$. $\square$