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

(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 $|f(x)| \le p(x)$ 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 $|F(x)| \le p(x)$ 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

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

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 $F(x) \le p(x)$ for all $x \in X$. Notice that

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

It follows that $|F(x)| \le p(x)$ 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 $f(ix) = if(x)$ 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

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

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 $\|x^*\| = m$. $\Box$

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 $Y = \text{span}(\{x_0\})$ 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

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

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

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

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

$$x_0 + \lambda v = \frac{1}{2}(x_0 + 2av) + \frac{1}{2}(x_0 + 2ibv) \in K. \quad \square$$

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

$$p^K(x) = \inf\{s \in (0,\infty) : x \in sK\} = \inf\{s \in (0,\infty) : \|x\| < s\} = \|x\| \text{ for all } x \in X.$$

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 $|\lambda| \leq 1$.

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 $x \in X, a \geq 0$,

(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

$$\operatorname{Re}(F(x)) \geq c \ \text{ for all } x \in M \ \text{ and } \ \operatorname{Re}(F(y)) \leq c \ \text{ for all } y \in N.$$

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

$$\operatorname{Re}(F(x)) - \operatorname{Re}(F(y)) \leq 0 \ \text{ for all } x \in M, y \in N,$$

or

$$\operatorname{Re}(F(x)) - \operatorname{Re}(F(y)) \geq 0 \ \text{ for all } x \in M, y \in N.$$

The proof of Lemma 5.34 is quite elementary and is left as an exercise.