AN INTRODUCTION TO FUNCTIONAL ANALYSIS

DILEEP MENON

Abstract. In this paper we discuss some fundamental results in real and functional analysis including the Riesz representation theorem, the Hahn-Banach theorem, and the Baire category theorem. We also discuss applications of these theorems to other topics in analysis.

Contents

1. Introduction

The goal of this paper is to develop the foundations of functional analysis. In doing so, we first review some basic concepts from real analysis. After a brief review of Banach space and Hilbert space theory, we introduce L^p spaces, which will serve as our primary working example throughout the paper. We circumvent a detailed discussion of measure theory and simply assume the reader is comfortable with Lebesgue integration. In section 4, we introduce *bounded linear operators* and *dual spaces* which are the building blocks in the theory that follows. We conclude this section with the Riesz representation theorem which states that the dual of any Hilbert space is isometrically isomorphic with itself. The Riesz representation theorem is undoubtedly one of the most important results in real analysis, because it allows us to explicitly describe certain dual spaces. In section 5, we prove the Hahn-Banach theorem using Zorn's Lemma and discuss some corollaries. We continue by proving two geometric versions of the Hahn-Banach theorem. In section 6, we prove the Baire category theorem in order to prove the uniform boundedness principle. In the final section, we introduce the notions of weak and weak-* convergence. We will apply the Hahn-Banach theorem and the uniform boundedness principle to deduce

Date: August 7, 2010.

properties of these weaker topologies. The weak topologies are fundamental in the study of functional analysis and partial differential equations.

2. Some Basic Definitions

In this section we define Banach spaces and Hilbert spaces. We also note that every Hilbert space is a Banach space under an appropriate norm, but the converse is not always true.

Definition 2.1. A ***normed linear space*** is a vector space X over $\mathbb{R}$ and a function $\|\cdot\| : X \to \mathbb{R}$ satisfying:

(i) $\|x\| \geq 0$ for all $x \in X$,
(ii) $\|x\| = 0$ if and only if $x = 0$,
(iii) $\|\alpha x\| = |\alpha|\|x\|$ for all $x \in X$, $\alpha \in \mathbb{R}$,
(iv) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in X$.

A metric space X is ***complete*** if every Cauchy sequence in X converges. A ***Banach space*** is a complete normed linear space.

Definition 2.2. A subset Y of a Banach space X is ***dense*** in X if for every $\epsilon > 0$ and $x \in X$ there exists a $y \in Y$ such that $\|x - y\| < \epsilon$ (or equivalently if $\overline{Y} = X$). A Banach space X is ***separable*** if it has a countable dense subset.

We can generalize the notion of length through the norm, and similarly we can generalize the notion of the angle between two elements by introducing the inner product.

Definition 2.3. An ***inner product*** over $\mathbb{R}$, on a general vector space X, is a map $(\cdot,\cdot) : X \times X \to \mathbb{R}$, such that

(i) $(\lambda x + \mu y, z) = \lambda(x, z) + \mu(y, z)$
(ii) $(x, y) = (y, x)$ for all $x, y \in X$, and
(iii) $(x, x) \geq 0$ for all $x \in X$, with equality if and only if $x = 0$.

The natural norm associated with the inner product is given by,

$$\|x\| = (x, x)^{1/2}. \tag{2.4}$$

Definition 2.5. A ***Hilbert Space*** is a complete inner product space.

We now state a very useful inequality that provides the triangle inequality for the norm defined in (2.4).

Theorem 2.6. (**Cauchy-Schwarz Inequality**) *Let $x, y \in X$, then*

$$|(x, y)| \leq \|x\|\|y\|.$$

By defining a norm through (2.4), any Hilbert space is also a Banach space. However, the converse is not necessarily true. For example, the L^p spaces, $p \neq 2$, are Banach spaces, but not Hilbert spaces. We will discuss L^p spaces in section 3.

3. Preliminaries

In this section we wish to review the theory of L^p spaces of Lebesgue integrable functions and some properties of Hilbert spaces.

We begin by stating the dominated convergence theorem, a fundamental convergence result from the theory of Lebesgue integration:

Theorem 3.1. (**Dominated Convergence Theorem**) *Let Ω be an open subset of $\mathbb{R}^n$ and $\{f_n\}$ be a sequence of measurable functions converging pointwise almost everywhere to a limit on Ω. If there is a function $g \in L^1(\Omega)$ such that $|f_n(x)| \leq g(x)$ for every n and almost every $x \in \Omega$, then*

$$\lim_{n\to\infty} \int_\Omega f_n(x)\, dx = \int_\Omega \left(\lim_{n\to\infty} f_n(x) \right)\, dx.$$

Now we define the norm on $L^p(\Omega)$, $1 \leq p < \infty$, which makes use of the Lebesgue integral:

Definition 3.2. The L^p***-norm*** is given by the integral (to be understood in the Lebesgue sense):

$$\|f\|_{L^p} = \left(\int_\Omega |f(x)|^p dx \right)^{1/p}, \tag{3.3}$$

and we define the space

$$L^p(\Omega) = \{f : \|f\|_{L^p} < \infty\}$$

with $1 \leq p < \infty$.

Next we record Hölder's inequality which is used to prove the triangle inequality for the L^p-norm:

Lemma 3.4. (**Hölder's Inequality**) *Let $p, q > 1$ with $p^{-1} + q^{-1} = 1$ and $1 < p < \infty$, and suppose that $f \in L^p(\Omega)$ and $g \in L^q(\Omega)$. Then, $fg \in L^1(\Omega)$, with*

$$\|fg\|_{L^1} \leq \|f\|_{L^p}\|g\|_{L^q}. \tag{3.5}$$

Now we state the triangle inequality for the L^p spaces, which is called Minkowski's inequality.

Lemma 3.6. (**Minkowski's Inequality**) *If $f, g \in L^p(\Omega)$, $1 \leq p < \infty$, then $f + g \in L^p(\Omega)$, with*

$$\|f + g\|_{L^p} \leq \|f\|_{L^p} + \|g\|_{L^q}. \tag{3.7}$$

Although we do not prove it here, one can show that all L^p spaces are complete and thus Banach spaces.

One "Lebesgue" space that does not arise naturally from the integration theory is $L^\infty(\Omega)$, the space of "essentially bounded" functions.

Definition 3.8. $L^\infty(\Omega)$ is the space of all functions f such that the essential supremum of f, given by

$$\|f\|_\infty = \operatorname{ess\,sup}_\Omega |f(x)|, \tag{3.9}$$

is finite. The ***essential supremum*** of f is the smallest value that bounds f almost anywhere:

$$\operatorname{ess\,sup}_\Omega |f(x)| = \inf \left\{ \sup_{x\in S} |f(x)| : S \subset \overline{\Omega}, \text{ with } \Omega \setminus S \text{ of measure zero} \right\}.$$

In particular $|f(x)| \leq \|f\|_\infty$ almost everywhere, and it follows that if $f \in C_b^0(\Omega)$, the space of bounded continuous functions on Ω, then the essential supremum of f is the same as its supremum.

Remark 3.10. L^∞ is a Banach space.

Now we define a related family of spaces, the l^p spaces of sequences.

Definition 3.11. For $1 \leq p < \infty$, l^p is the space of all infinite sequences $\{x_n\}_{n=1}^\infty$ such that the l^p-norm

$$\|x\|_{l^p} = \left(\sum_{j=1}^{\infty} |x_j|^p \right)^{1/p} \tag{3.12}$$

is finite. l^∞ consists of all bounded sequences, with norm

$$\|x\|_{l^\infty} = \sup_{j \in \mathbb{Z}^+} |x_j|. \tag{3.13}$$

At this point we wish to clarify the relationship between L^p and l^p spaces. First, consider the characteristic function:

$$\chi_{[n,n+1]}(x) = \begin{cases} 1 & : x \in [n, n+1] \\ 0 & : x \notin [n, n+1] \end{cases}$$

Given $\{\alpha_n\} \in l^p$, let f be defined such that

$$f(x) = \sum_{n=1}^{\infty} \alpha_n \chi_{[n,n+1]}(x).$$

Then the L^p norm of f is

$$\begin{aligned} \|f(x)\|_{L^p} &= \left(\int \left(\sum_{n=1}^{\infty} \alpha_n \chi_{[n,n+1]}(x) \right)^p \right)^{1/p} \\ &= \left(\sum_{n=1}^{\infty} (\alpha_n)^p \right)^{1/p} = \|\alpha_n\|_{l^p} < \infty. \end{aligned}$$

which is the l^p-norm of the sequence $\{\alpha_n\}_{n=1}^\infty$ with $f \in L^p$.

We end this section by stating, without proof, the orthogonal decomposition of Hilbert spaces onto a linear subspace. We will use this result in the proof of the Riesz representation theorem.

Definition 3.14. If M is a subset of a Hilbert space H, then the ***orthogonal complement*** of H, $M^\perp$, is given by

$$M^\perp = \{u \in H : (u, v) = 0 \text{ for all } v \in M\}.$$

Proposition 3.15. *If M is a closed linear subspace of H, then every $x \in H$ has a unique decomposition as*

$$x = u + v, \qquad u \in M, v \in M^\perp.$$

4. Dual Spaces and the Riesz Representation Theorem

In this section we prove some useful results about bounded linear operators and introduce dual spaces. We then prove our first major result, the Riesz representation theorem.

Definition 4.1. An operator A on a vector space V is ***linear*** if

$$A(x + \lambda y) = Ax + \lambda Ay$$

for all $x, y \in V$ and $\lambda \in \mathbb{R}$ (or $\mathbb{C}$).

Definition 4.2. A linear operator A from a normed vector space $(X, \|\cdot\|_X)$ into another normed space $(Y, \|\cdot\|_Y)$ is ***bounded*** if there exists a constant M such that

$$\|Ax\|_Y \leq M\|x\|_X \qquad \text{for all} \qquad x \in X. \tag{4.3}$$

Let $\mathcal{L}(X, Y)$ denote the space of all bounded linear maps from X into Y.

Definition 4.4. The ***operator norm*** of an operator $A : X \to Y$ is the smallest value of M such that (4.3) holds:

$$\|A\|_{\mathcal{L}(X,Y)} = \inf\{M : \|Ax\|_Y \leq M\|x\|_X \text{ for all } x \in X\}. \tag{4.5}$$

Equivalently,

$$\|A\|_{\mathcal{L}(X,Y)} = \sup_{x \neq 0} \frac{\|Ax\|_Y}{\|x\|_X} = \sup_{\|x\|_X = 1} \|Ax\|_Y. \tag{4.6}$$

The space $\mathcal{L}(X, Y)$ is a Banach space whenever Y is a Banach space; this does not depend on whether or not the space X is complete.

Proposition 4.7. *Let X be a normed vector space and Y a Banach space. Then $\mathcal{L}(X, Y)$ is a Banach space.*

Proof. Let $\{A_n\}$ be a Cauchy sequence in $\mathcal{L}(X, Y)$. We want to show that $A_n \to A$ for some $A \in \mathcal{L}(X, Y)$. Since A_n is Cauchy, given $\epsilon > 0$ there exists an N such that

$$\|A_n - A_m\|_{\mathcal{L}(X,Y)} \leq \epsilon \qquad \text{for all} \qquad n, m \geq N. \tag{4.8}$$

We now show that for every fixed $x \in X$ the sequence $\{A_n x\}$ is Cauchy in Y. This follows since

$$\|A_n x - A_m x\|_Y = \|(A_n - A_m)x\|_Y \leq \|A_n - A_m\|_{\mathcal{L}(X,Y)}\|x\|_X, \tag{4.9}$$

and $\{A_n\}$ is Cauchy in $\mathcal{L}(X, Y)$. Since Y is complete, it follows that $A_n x \to y$, where y depends on x. Let us define a mapping $A : X \to Y$ by $Ax = y$. First we show that A is linear since

$$A(x + \lambda y) = \lim_{n\to\infty} A_n(x + \lambda y) = \lim_{n\to\infty} A_n x + \lambda \lim_{n\to\infty} A_n y = Ax + \lambda Ay.$$

In order to show that A is bounded, take $n, m \geq N$ (from (4.8)) in (4.9) and let $m \to \infty$. Since $A_m x \to Ax$ this shows that

$$\|A_n x - Ax\|_Y \leq \epsilon\|x\|_X. \tag{4.10}$$

Since (4.10) holds for every x it follows that

$$\|A_n - A\|_{\mathcal{L}(X,Y)} \leq \epsilon, \tag{4.11}$$

and so $A_n - A \in \mathcal{L}(X, Y)$. Since $\mathcal{L}(X, Y)$ is a vector space and $A_n \in \mathcal{L}(X, Y)$ it follows that $A \in \mathcal{L}(X, Y)$, and $A_n \to A$ in $\mathcal{L}(X, Y)$. □

An extremely useful property of linear operators is that boundedness is equivalent to continuity.

Proposition 4.12. *Let $L : X \to Y$ be a linear map. Then L is continuous if and only if it is bounded.*

Proof. First let us suppose L is bounded, then

$$\|L(x_n - x)\|_Y \leq \|L\|_{\mathcal{L}(X,Y)}\|x_n - x\|_X,$$

which implies continuity since given any $\epsilon > 0$ there exists $\delta > 0$ such that $\|x_n - x\|_X < \delta$ implies $\|L(x_n - x)\|_Y < \epsilon$.

Now let us suppose L is continuous but unbounded; this implies that for every n there exists a y_n such that $\|Ly_n\|_Y > n^2\|y_n\|_X$. Now let

$$x_n = \frac{y_n}{(n\|y_n\|_X)} \to 0,$$

but $\|lx_n\|_Y > n$, so that L is not continuous at the origin, which is a contradiction. Thus continuity implies boundedness. □

Definition 4.13. A bounded linear map from a Banach space X into $\mathbb{R}$ (an element of $\mathcal{L}(X, \mathbb{R})$) is called a ***linear functional*** on X. The space $\mathcal{L}(X, \mathbb{R})$ of all linear functionals on X is denoted by X^* and is called the ***dual space*** of X.

Before we start working with dual spaces, we will provide an example using L^p spaces.

Example 4.14. *The Dual Space of* L^p, $1 < p < \infty$:

Let $p, q > 1$ with $p^{-1} + q^{-1} = 1$. Then if $f \in L^q(\Omega)$, we can define a linear functional L_f on L^p by

$$L_f(g) = \int_\Omega f(x)g(x)dx, \tag{4.15}$$

and this is well defined because Hölder's inequality gives us

$$|L_f(g)| \leq \|f\|_{L^q}\|g\|_{L^p} < \infty.$$

Clearly, we have $\|L_f\|_{(L^p)^*} \leq \|f\|_{L^q}$. If we set

$$g(x) = |f(x)|^{q-2}f(x)$$

then

$$\|g(x)\|_{L^p} = \left(\int_\Omega |f(x)|^{(q-1)p}dx\right)^{1/p} = \left(\int_\Omega |f(x)|^q dx\right)^{1/p} = \|f\|_{L^q}^{q/p},$$

and

$$|L_f(g)| = \left|\int_\Omega |f(x)|^q dx\right| = \|f\|_{L^q}^q.$$

Therefore,

$$|L_f(g)| = \|f\|_{L^q}\|g\|_{L^p},$$

showing that we have

$$\|L_f(g)\|_{(L^p)^*} = \|f\|_{L^q}.$$

We have shown that the map $f \mapsto L_f$ is an isometry from L^q ***into*** $(L^p)^*$. Furthermore, one can show that every element of $(L^p)^*$ can be realized as L_f for some $f \in L^q$. (In other words that the map $f \mapsto L_f$ is onto). It follows that L^q and

$(L^p)^*$ are isometrically isomorphic, denoted as $L^q \simeq (L^p)^*$. We conclude that every element of $(L^p)^*$ can be identified with an element of L^q.

It is important to note that $(L^\infty)^* \not\simeq L^1$. We will prove this later using the Hahn-Banach theorem.

Now we show that for any Hilbert Space H, the dual of H is identifiable with H itself. In particular, we show that any linear functional l on H can be represented as an inner product with some appropriate element x_l of H itself:

Theorem 4.16. (**Riesz Representation Theorem**). *For any Hilbert space H, $H^* \simeq H$. In particular, for every $x \in H$,*

$$l_x(y) \equiv (x, y) \tag{4.17}$$

is bounded and has norm $\|l_x\|_{H^*} = \|x\|$. Furthermore, for every bounded linear functional $l \in H^*$ there exists a unique $x_l \in H$ such that

$$l(y) = (x_l, y) \qquad \textit{for all} \qquad y \in H, \tag{4.18}$$

and $\|x_l\|_H = \|l\|_{H^*}$. It follows that $l \mapsto x_l$ is continuous.

Proof. By using equation (4.17) and the Cauchy-Schwarz inequality we get

$$|l_x(y)| = |(x, y)| \leq \|x\|\|y\|$$

so that $l_x \in H^*$ with $\|l_x\|_{H^*} \leq \|x\|$, thus l_x is bounded. By choosing $y = x$ we see that in fact $\|l_x\|_{H^*} = \|x\|$.

Now suppose that $l \in H^*$. Since l is bounded it is also continuous by Prop. 4.12. We know that if l is continuous then the pre-image of a closed set is closed. Then the kernel of l, $K = \{y \in H : l(y) = 0\}$, is a closed subspace of H. We next observe that the subspace $K^\perp$ of vectors orthogonal to K is a one-dimensional subspace of H. To see this, for $u, v \in K^\perp$ we have

$$l[l(u)v - l(v)u] = 0$$

and so $[l(u)v - l(v)u] \in K$ by definition. Since u and v are orthogonal to K, so is $l(u)v - l(v)u$. Thus $l(u)v - l(v)u$ is in K and $K^\perp$ and hence is equal to 0. This shows that u and v are proportional.

Now choose a unit vector $z \in K^\perp$. By Prop. 3.15, we can decompose every $y \in H$ as $y = (z, y)z + w$, where $w \in K$. Then $l(y) = (z, y)l(z)$, and so if we set $x_l = l(z)z$, we have

$$(x_l, y) = (l(z)z, y) = l(z)(z, y) = l(y).$$

Again using Cauchy-Schwarz we see that $\|l\|_{H^*} \leq \|x_l\|_H$ and in fact, $\|l\|_{H^*} = \|x_l\|_H$. To show that x_l is unique suppose $l(y) = (x_1, y) = (x_2, y)$, with $x_1, x_2 \in K^\perp$. This implies that $x_1 = \lambda x_2$ for some $\lambda \in \mathbb{R}$. Therefore, $x_1 = x_2$. It also follows that $l \mapsto x_l$ is continuous since l is bounded. □

5. The Hahn-Banach Theorem

5.1. The Analytic Hahn-Banach Theorem.

Definition 5.1. A set S is called ***partially ordered*** if there is a relation $x \prec y$ defined for certain pairs of elements (x, y) of S such that

(i) $x \prec x$ for all $x \in S$,

(ii) $x \prec y$, $y \prec x$ implies $x = y$,
(iii) $x \prec y$, $y \prec z$ implies $x \prec z$.

Definition 5.2. The set S is called ***totally ordered*** if for each pair (x, y) of elements of S, one has either $x \prec y$ or $y \prec x$ (or both).

Definition 5.3. Let T be a subset of a partially ordered set S. Then $x_0 \in S$ is an ***upper bound*** of T if $x \prec x_0$ for all $x \in T$. An element x_0 is said to be ***maximal*** for S if $x_0 \prec x$ implies $x = x_0$.

Lemma 5.4. (**Zorn's Lemma**) *If S is a partially ordered set such that each totally ordered subset has an upper bound in S, then S has a maximal element.*

Now we prove that any linear functional defined on a linear subspace of a Banach space X can be extended to a bounded functional on all of X:

Theorem 5.5. (**Hahn-Banach Theorem**) *Let E be a vector space and $p : E \to \mathbb{R}$ be a function such that*

(i) $p(\lambda x) = \lambda p(x)$ for all $x \in E$ and $\lambda \in \mathbb{R}$, $\lambda > 0$ (homogeneity),
(ii) $p(x + y) \leq p(x) + p(y)$ for all $x, y \in E$ (sublinearity).

Suppose G is a subspace of E and g is a linear functional defined on G which satisfies $g(x) \leq p(x)$ for all $x \in G$. Then, there exists a linear functional $f : E \to \mathbb{R}$ with $f(x) = g(x)$ for all $x \in G$ and $f(y) \leq p(y)$ for all $y \in E$.

Proof. Let P denote the set of all linear maps $h : D(h) \to \mathbb{R}$, where $D(h)$ is a subspace of E, such that

$$h|_G = g, \text{ and } h(x) \leq p(x) \text{ for all } x \in D(h).$$

Clearly P is nonempty since $g \in P$.

Let us define a partial order on P such that

$$h_1 \leq h_2 \quad \text{if and only if} \quad D(h_1) \subset D(h_2) \quad \text{and} \quad h_2(x) = h_1(x) \text{ for all } x \in D(h_1).$$

Now we will show that every totally ordered subset Q of P has an upper bound. Let

$$Q = (h_i)_{i \in I}.$$

If

$$D(h) = \bigcup_{i \in I} D(h_i) \quad \text{and} \quad h(x) = h_i(x) \qquad \text{for all} \qquad x \in D(h_i)$$

then $h \in P$ and $h_i \leq h$ for all $i \in I$, and therefore, h is an upper bound. By Zorn's Lemma P admits a maximal element, f.

Then by construction $f : D(f) \to \mathbb{R}$, $f(x) = g(x)$ for all $x \in G$, and $f \leq p$ on $D(f)$.

If $D(f) = E$ then the proof is complete.

If $D(f) \neq E$ then there exists $x_0 \in E \setminus D(f)$. We will show that f can be further extended which contradicts the maximality of f. Let us construct

$$\tilde{f} : D(f) + \mathbb{R}x_0 = D(\tilde{f}) \to \mathbb{R}.$$

Define $\tilde{f}(x_0) = \alpha$ and $\tilde{f}(x + tx_0) = f(x) + t\alpha$. Clearly $\tilde{f}$ is a linear functional that extends g.

Now we want to show $f(x) + t\alpha \leq p(x + tx_0)$ for all $t \in \mathbb{R}$ and $x \in D(f)$. Since p is positively homogeneous of degree 1, let $t = \pm 1$, so

$$f(x) - \alpha \leq p(x - x_0)$$
$$f(y) + \alpha \leq p(y + x_0),$$

for all $x, y \in D(f)$. Then

$$f(x) - p(x - x_0) \leq \alpha \leq p(y + x_0) - f(y),$$

for all $x, y \in D(f)$. Now it suffices to show that

$$f(x) - p(x - x_0) \leq p(y + x_0) - f(y),$$

and we know that

$$f(x) + f(y) = f(x + y) \leq p(x + y) \leq p(x - x_0) + p(y + x_0),$$

since $f \leq p$ on $D(f)$. Thus we have shown that f can be further extended, which contradicts the maximality of f in E. Hence $D(f) = E$ and the theorem is proved. □

The Hahn-Banach theorem is an extremely powerful tool in functional analysis; here we state just a few of its immediate consequences:

Remark 5.6. Note that if we let $p(x) = \|g\|_{G^*}\|x\|$, then by the Hahn-Banach theorem there exists $f \in E^*$ such that $f(x) = g(x)$ for all $x \in G$ and $f(y) \leq \|g\|_{G^*}\|y\|$ for all $y \in E$. It follows that

$$\|f\|_{E^*} = \sup_{x \in E} \frac{|f(x)|}{\|x\|} = \|g\|_{G^*}.$$

Corollary 5.7. *Let X be a Banach space.*

(i) If $x \in X$, then there exists $f \in X^*$ such that $\|f\|_{X^*} = 1$, $f(x) = \|x\|_X$.
(ii) For all $x \in X$,

$$\|x\|_X = \max_{f \in X^*, \|f\| \leq 1} f(x).$$

Proof. *(i)* Let G be a subspace of E such that $G = \mathbb{R}x$. Define $g : G \to \mathbb{R}$ on G by

$$g(\alpha x) = \alpha\|x\| \qquad \text{for all} \qquad \alpha \in \mathbb{R}.$$

Then g is linear, and

$$|g(\alpha x)| = |\alpha| \cdot \|x\| = \|\alpha x\|.$$

So g is bounded on G and,

$$\|g\|_{G^*} = \sup_{x \in G, x \neq 0} \frac{g(x)}{\|x\|} = 1.$$

Then by the Hahn-Banach theorem, there is a bounded linear functional f on X such that $\|f\|_{X^*} = 1$ and $f(x) = \|x\|_X$.

(ii) It follows from Remark 5.6 that

$$|f(x)| \leq \|f\|\|x\|.$$

Therefore,

$$\|x\| \geq \sup_{f \in X^*, f \neq 0} \frac{|f(x)|}{\|f\|}.$$

However, by *(i)* for each $x \in X$ there is an $f \in X^*$ such that $\|f\| = 1$ and $f(x) = \|x\|$. So,

$$\|x\| = \max_{f\in X^*, f\neq 0} \frac{|f(x)|}{\|f\|}.$$

□

Corollary 5.8. *Let X be a Banach space. If $x, y \in X$ and $f(x) = f(y)$ for every $f \in X^*$ then $x = y$.*

Proof. If $x = y = 0$ then the proof is done. Therefore, without loss of generality assume that $x \neq 0$. Assume $x \neq y$ and let Y be the linear subspace spanned by x and y. If x and y are linearly dependent, set

$$g(\alpha x) = \alpha|x|,$$

otherwise set

$$g(\alpha x + \beta y) = \alpha|x| \qquad \text{for all} \qquad \beta \in \mathbb{R}.$$

In both cases $g(x) \neq g(y)$. Then by using the Hahn-Banach theorem, we can extend g to an element $f \in X^*$. Clearly, $f(x) \neq f(y)$. □

We use the Hahn-Banach theorem in the following example to show that L^1 is not the whole dual space of L^∞.

Example 5.9. We want to show that not every element of the dual of L^∞ is given by

$$\Lambda(g) = \int_{-1}^{1} f(x)g(x)\, dx$$

for some $f \in L^1$.

Proof. Let $C[-1, 1]$ be the space of continuous functions on the closed interval $[-1, 1]$. Clearly, $C[-1, 1]$ is a subspace of $L^\infty[-1, 1]$. Also, $C[-1, 1]$ is a Banach space with norm

$$\|\cdot\|_\infty = \max_{x\in[-1,1]} |f(x)|.$$

Let $\Lambda : C[-1, 1] \to \mathbb{R}$ be such that $\Lambda(f) = f(0) \leq \|f\|_\infty$ and assume Λ is linear. By the Hahn-Banach theorem, there exists $\overline{\Lambda} : L^\infty[-1, 1] \to \mathbb{R}$ such that

$$\overline{\Lambda}(f) = f(0) \qquad \text{for all} \qquad f \in C[-1, 1],$$

and $\overline{\Lambda} \in (L^\infty)^*$. We want to show that there does not exist $g \in L^1$ such that

$$\overline{\Lambda}(f) = \int_{-1}^{1} f(x)g(x)dx \qquad \text{for all} \qquad f \in L^\infty[-1, 1].$$

To show this suppose there exists such a function $g \in L^1[-1, 1]$. Consider the sequence of functions $\{f_n\}$ such that

$$f_n(x) = \max\{1 - n|x|, 0\}.$$

$f_n(x)$ converges to 0 pointwise and clearly, $|f(x)| \leq 1$.

We know $\overline{\Lambda}$ is well defined since

$$\int |f_n(x)g(x)|dx \leq \int |g(x)|dx < \infty.$$

Hence, by the dominated convergence theorem we have

$$\lim_{n\to\infty} \overline{\Lambda}(f_n) = \lim_{n\to\infty} \int f_n(x)g(x)dx = \int \lim_{n\to\infty} f_n(x)g(x)dx = 0.$$

However, $f_n(0) = 1$ for all n, by definition, which leads to a contradiction. This shows that there does not exist such a function $g \in L^1[-1,1]$. □

5.2. **The Geometric Hahn-Banach Theorems.** Here we apply the analytic Hahn-Banach theorem to prove two versions of the geometric Hahn-Banach theorem. The geometric versions of the Hahn-Banach theorem are often more useful when trying to prove results in analysis.

Definition 5.10. Let $f : E \to \mathbb{R}$ be a nonzero linear functional, and $\alpha \in \mathbb{R}$. We then define a ***hyperplane*** H on E by the set $H = \{x \in E : f(x) = \alpha\}$.

Definition 5.11. Let E be a normed vector space and let A and B be subsets of E. We say that a hyperplane $H = \{x \in E : f(x) = \alpha\}$ ***separates*** A and B if $f(x) \le \alpha$ for all $x \in A$ and $f(y) \ge \alpha$ for all $y \in B$. H ***strictly separates*** A and B if $f(x) < \alpha$ for all $x \in A$ and $f(y) > \alpha$ for all $y \in B$.

We state and prove two lemmas that will allow us to prove one version of the Geometric Hahn-Banach theorem.

Lemma 5.12. *Let C be an open convex subset of a normed vector space E such that $0 \in C$. Then there exists $p : E \to \mathbb{R}$ such that*
(i) $p(\beta x) = \beta p(x)$ *for all* $\beta \in \mathbb{R}$ *and* $x \in E$,
(ii) $p(x+y) \le p(x) + p(y)$ *for all* $x, y \in E$,
(iii) $0 \le p(x) \le M\|x\|$ *for all* $x \in E$,
(iv) $C = \{x \in E : p(x) < 1\}$.

Proof. For all $x \in E$ define $p(x) = \inf\{\alpha > 0 : \alpha^{-1}x \in C\}$.
(i) Clearly, $p(\beta x) = \inf\{\alpha > 0 : \alpha^{-1}\beta x \in C\} = \beta p(x)$.

(iii) Since C is open, there exists an $r > 0$ such that $B(0,r) \subset C$, which implies $p(x) \le \frac{1}{r}\|x\|$ for all $x \in E$. Set $M = \frac{1}{r}$.

(iv) Since C is open, $x \in C$ implies $(1+\epsilon)x \in C$, $0 < \epsilon \ll 1$. By definition it follows that $p(x) \le \frac{1}{1+\epsilon} < 1$, which implies that $C \subseteq \{p(x) < 1\}$. Now suppose that $p(x) < 1$, then there exists $\alpha \in (0,1)$ such that $\alpha^{-1}x \in C$. Since C is convex and $0, \alpha^{-1}x \in C$ this implies $x \in C$. This shows that $C = \{x \in E : p(x) < 1\}$.

(ii) For every $x \in E$ and $\epsilon > 0$,

$$\frac{x}{p(x)+\epsilon} \in C,$$

because

$$p\left(\frac{x}{p(x)+\epsilon}\right) = \frac{p(x)}{p(x)+\epsilon} < 1.$$

Again, since C is convex, for all $t \in [0,1]$,

$$\frac{tx}{p(x)+\epsilon} + \frac{(1-t)y}{p(y)+\epsilon} \in C.$$

Now, set

$$t = \frac{p(x) + \epsilon}{p(x) + p(y) + 2\epsilon}$$

It follows that

$$\frac{x+y}{p(x)+p(y)+2\epsilon} \in C, \quad \text{and} \quad p\left(\frac{x+y}{p(x)+p(y)+2\epsilon}\right) < 1.$$

This gives

$$p(x+y) \leq p(x) + p(y) + 2\epsilon.$$

□

Lemma 5.13. *Let C be a nonempty open convex subset of a normed vector space E and $x_0 \in E \setminus C$. Then there exists $f \in E^*$ such that $f(x) < f(x_0)$ for all $x \in C$.*

Proof. Without loss of generality, assume $0 \in C$. Let G be a subspace of E such that $G = \mathbb{R}x_0$ and $g : G \to \mathbb{R}$ be a linear functional such that $g(tx_0) = t$.

Define $p : E \to \mathbb{R}$ as in Lemma 5.12 above. Note that $x_0 \notin C$, then $g \leq p$ on G since $1 \leq p(x_0)$. Then by the Hahn-Banach theorem, there exists a bounded linear functional $f : E \to \mathbb{R}$ such that $f(x) = g(x)$ for all $x \in G$ and $f \leq p$ on E.

It follows that $f(x) \leq M\|x\|$ for all $x \in E$, and $f(x) \leq p(x) < 1$ for all $x \in C$. Set $f(x_0) = 1$, then $f(x) < f(x_0)$ for all $x \in C$. □

Now we are well equipped to prove the geometric Hahn-Banach theorems:

Theorem 5.14. (**Geometric Hahn-Banach 1**). *Let A and B be nonempty disjoint convex subsets of a normed vector space E. If A is open, then there exists a hyperplane H separating A and B.*

Proof. Let $C = A - B = \{x - y : x \in A, y \in B\}$. Since A and B are disjoint $0 \notin C$. It is easy to see that C is also a nonempty open convex subset of E.

By Lemma 5.13, if $x_0 \in E \setminus C$, then there exists $f \in E^*$ such that $f(z) < f(x_0)$ for all $z \in C$. Let $x_0 = 0$ and $f(0) = 0$. Then, $f(z) < 0$ for all $z \in C$.

It follows that

$$\begin{aligned} f(x-y) &< 0, \\ f(x) &\leq f(y) \qquad \text{for all} \qquad x \in A, y \in B. \end{aligned}$$

So there exists an $\alpha \in \mathbb{R}$ such that

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

Thus the hyperplane $H = \{x \in E : f(x) = \alpha\}$ separates A and B. □

Theorem 5.15. (**Geometric Hahn-Banach 2**). *Let A and B be nonempty disjoint convex subsets of a normed vector space E. If A is closed and B is compact, then there exists a hyperplane H that strictly separates A and B.*

Proof. Define $A_\epsilon = A + B(0, \epsilon)$ and $B_\epsilon = B + B(0, \epsilon)$. Then, both A_ϵ and B_ϵ are nonempty, open and convex.

We claim there exists $\epsilon > 0$ such that $A_\epsilon \cap B_\epsilon = \emptyset$. To see why, suppose not. Then for every $\epsilon > 0$ we have $A_\epsilon \cap B_\epsilon \neq \emptyset$. Choose a sequence $\{\epsilon_k\} \to 0$ such that we can find sequences $\{x_k\} \subset A$ and $\{y_k\} \subset B$ satisfying

$$|x_k - y_k| < 2\epsilon_k.$$

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

Since B is compact, $\{y_k\}$ has a convergent subsequence $\{y_{kj}\}$ such that $\{y_{kj}\} \to y$ in B. So there exists $N_1 \in \mathbb{N}$ such that for all $k \geq N_1$

$$|y_{kj} - y| < 2\epsilon_{kj}.$$

Define a subsequence $\{x_{kj}\}$ of $\{x_k\}$ such that

$$|x_{kj} - y_{kj}| < 2\epsilon_{kj}.$$

Then,

$$|x_{kj} - y| \leq |x_{kj} - y_{kj}| + |y_{kj} - y| < 4\epsilon_{kj}.$$

Since A is closed, it must follow that $y \in A$ and $y \in B$. This is a contradiction since A and B are disjoint, so $A_\epsilon \cap B_\epsilon = \emptyset$ for some $\epsilon > 0$.

Now if we apply the first geometric Hahn-Banach theorem (Theorem 5.13) on A_ϵ and B_ϵ we have

$$f(x + \beta\epsilon) \leq \alpha \leq f(y - \beta\epsilon) \qquad \text{for all} \qquad x \in A,\ y \in B,$$

for some $0 < \beta < 1$. Hence, A and B are strictly separated by the hyperplane $H = \{x \in E : f(x) = \alpha\}$. □

Here we state one important corollary of the the second geometric Hahn-Banach theorem.

Corollary 5.16. *Let F be a subspace of a normed vector space E such that $\overline{F} \neq E$. Then there exists $f \in E^* \setminus \{0\}$ such that $f(x) = 0$ for all $x \in F$.*

Proof. Since $\overline{F} \neq E$, there exists $x_0 \in E \setminus \overline{F}$. Let $A = \overline{F}$ and $B = \{x_0\}$. Then by the second geometric Hahn-Banach theorem (Theorem 5.15) there exists a hyperplane determined by the pair (f, α), $f \in E^*$, that strictly separates A and B,

$$f(x) < \alpha < f(y) \qquad \text{for all} \qquad x \in A. \tag{5.17}$$

Now since A is a subspace of E it is closed under scalar multiplication and since (5.17) holds for all $x \in A$ we can write:

$$f(\beta x) = \beta f(x) < \alpha < f(y) \qquad \text{for all} \qquad x \in A,\ \beta \in \mathbb{R}.$$

In order for this to hold, $f(x)$ is necessarily 0 for all $x \in F$. □

This corollary is one of the most important results related to the Hahn-Banach theorem, because it provides an alternative way of showing density of a subspace. This is often much easier to work with than trying to verify the topological definition of a dense subspace.

6. The Baire Category Theorem and the Uniform Boundedness Principle

Another fundamental result in functional analysis is the uniform boundedness principle. In order to prove the uniform boundedness principle, we will have to use the following topological result from real analysis, the Baire Category theorem.

Theorem 6.1. (**Baire Category Theorem**). *If U_i is a countable family of dense open subsets of a Banach space X, then*

$$U = \bigcap_{n=1}^{\infty} U_n$$

is dense in X.

Proof. Take $x \in X$, and $r > 0$; we want to show that $B(x,r) \bigcap U$ is not empty. Now, since each U_n is dense and open, for some $y \in U_n$ and $s > 0$, we have

$$B(x,r) \cap U_n \supset B(y,2s) \supset \overline{B}(y,s).$$

First take $x_1 \in X$ and $r_1 < 1/2$ such that $\overline{B}(x_1,r_1) \subset U_1 \cap B(x,r)$, then take $x_2 \in X$ and $r_2 < 2^{-2}$ such that $\overline{B}(x_2,r_2) \subset U_2 \cap B(x_1,r_1)$, and in general take $x_n \in X$ and $r_n < 2^{-n}$ such that $\overline{B}(x_n,r_n) \subset U_n \cap B(x_{n-1},r_{n-1})$. By this we obtain a nested sequence of closed sets

$$\overline{B}(x_1,r_1) \supset \overline{B}(x_2,r_2) \supset \cdots. \tag{6.2}$$

Since the space X is complete we have,

$$\bigcap_{n=1}^{\infty} \overline{B}(x_n,r_n) = x_0.$$

So, the points $\{x_j\}$ form a Cauchy sequence by (6.2), and they converge to x_0, which is the intersection of all the nested sets. Now, $x_0 \in \overline{B}(x_1,r_1) \subset B(x,r)$, and also $x_0 \in \overline{B}(x_n,r_n) \subset U_n$ for all n. So $x_0 \in U \cap B(x,r)$ as asserted. □

By taking complements, we have this immediate useful corollary:

Corollary 6.3. *Let X be a Banach space and F_j a countable sequence of nowhere dense subsets. Then*

$$\bigcup_{j=1}^{\infty} F_j \neq X.$$

We now prove the uniform boundedness principle, which will come into play in proving some results about weak convergence.

Theorem 6.4. (**Uniform Boundedness Principle**). *Let X be a Banach space and Y normed space. Let $S \subset \mathcal{L}(X,Y)$, and let*

$$\sup_{T\in S} \|Tx\|_Y < \infty \qquad \textit{for all} \qquad x \in X.$$

Then

$$\sup_{T\in S} \|T\|_{\mathcal{L}(X,Y)} < \infty.$$

Proof. Consider the sets

$$G_j = \{x \in X : \|Tx\|_Y \leq j \text{ for all } T \in S\}.$$

Then $\bigcup_j G_j = X$, and since G_j is closed, Corollary 6.3 shows that at least one of the G_j must have a nonempty interior, call it G_n. Then there exists $y \in X$, and $r > 0$ such that $B(y,r) \subset G_n$.
Therefore, if $\|x\|_X \leq r$, so that $(y+x) \in G_n$,

$$\|Tx\|_Y = \|T(y+x) + T(-y)\| \leq n + \|Ty\|_Y \leq R$$

for some $R > 0$, since $\sup_{T\in S} \|Ty\|_Y$ is bounded. Thus, for any x with $\|x\|_X = r$ we have

$$\|Tx\|_Y \leq \frac{R}{r}\|x\|_X, \qquad \text{for all} \qquad T \in S.$$

This shows that,

$$\|T\|_{\mathcal{L}(X,Y)} \leq \frac{R}{r} \qquad \text{for all} \qquad T \in S,$$

□

7. Weak and Weak-* Convergence

First, we introduce the notion of a reflexive space in order to prove some basic results about weak convergence.

Let X be a general Banach space. Then for an element $x \in X$ we can define a linear functional G_x on X^* by

$$G_x(f) = f(x) \qquad \text{for all} \qquad f \in X^*. \tag{7.1}$$

with $G_x \in X^{**} \equiv (X^*)^*$. If we set $Ax = G_x$ we get a linear map $A : X \to X^{**}$. Note that

$$|G_x(f)| \leq \|f\|_{X^*}\|x\|,$$

then,

$$\|G_x\|_{X^{**}} \leq \|x\|.$$

Now by Corollary 5.7, given an $x \in X$ there exists an $f \in X^*$ with $\|f\|_{X^*} = 1$ and $f(x) = |x|$; therefore,

$$\|G_x\|_{X^{**}} = \|x\|, \tag{7.2}$$

and hence A is an isometry from X onto a subspace of X^{**}. When the isometry is onto (when $X \simeq X^{**}$), we say that X is ***reflexive***.

Definition 7.3. Let X be a Banach space. A sequence $x_n \in X$ ***converges weakly*** to $x \in X$, denoted by

$$x_n \rightharpoonup x \quad \text{in} \quad X,$$

if $f(x_n) \to f(x)$ for every $f \in X^*$.

In order to motivate the preceding terminology, we point out that the standard notion of convergence in the norm is referred to as ***strong convergence***. Now we show that strong convergence implies weak convergence, but the converse is not always true.

Lemma 7.4. *If $x_n \to x$ (strong convergence) then $x_n \rightharpoonup x$ (weak convergence).*

Proof. Every element of the dual space, X^* is a bounded linear functional and therefore is continuous. So, $f(x_n) \to f(x)$ for every $f \in X^*$, which is exactly $x_n \rightharpoonup x$. □

Example 7.5. Here we provide a simple example of weakly convergent subsequences that do not converge strongly:

Let $\{e_j\}$ be an orthonormal basis in a separable Hilbert space. The unit ball in an infinite-dimensional Hilbert space is not compact, so the $\{e_j\}$ have no convergent subsequence. Hence the sequence $\{e_j\}$ does not converge strongly. However, we will show that $e_j \rightharpoonup 0$.

By the Riesz representation theorem, every element $l \in H^*$ has a representation as $(x_l, \cdot)$ for some $x_l \in H$. It suffices to consider sequences (x_l, e_j) for each $x_l \in H$. Now since $\{e_j\}$ is an orthonormal basis, we have

$$\|x_l\|^2 = \sum_{j=l}^{\infty} |(x_l, e_j)|^2,$$

and as $j \to \infty$ we have $|(x_l, e_j)| \to 0$. And this is precisely $l(e_j) \to 0$ for every $l \in H^*$, and therefore $e_j \rightharpoonup 0$.

The following result relies on the Hahn-Banach Theorem and uniform boundedness principle.

Proposition 7.6. *Weak limits are unique, and weakly convergent sequences are bounded.*

Proof. If $x_n \rightharpoonup x$ and $x_n \rightharpoonup y$ then it follows that $f(x) = f(y)$ for every $f \in X^*$. By Corollary 5.8, this implies that $x = y$.

For boundedness, notice that for each $f \in X^*$, $\{f(x_n)\}$ is a convergent sequence of real numbers and is bounded,

$$|f(x_n)| \leq C_f \qquad \text{for all n.}$$

Now define an element $G_n \in X^{**}$ that corresponds to x_n,

$$G_n(f) = f(x_n) \qquad \text{for all} \qquad f \in X^*.$$

Then

$$|G_n(f)| \leq C_f \qquad \text{for all n,}$$

and $\{G_n(f)\}$ is a bounded sequence for every $f \in X^*$. By Prop. 4.7 X^* is complete and then by the uniform boundedness principle the sequence $\{\|G_n\|_{X^{**}}\}$ is bounded. Finally, by equation (7.2) $\|G_n\|_{X^{**}} = \|x_n\|$, so $\{x_n\}$ is bounded as wanted. □

Taking weak limits can decrease the norm, however it can never increase it. The following lemma shows this:

Lemma 7.7. *If* $x_n \rightharpoonup x$ *in* X *then*

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

Proof. If $x_n \rightharpoonup x$, choose $f \in X^*$ such that $\|f\|_{X^*} = 1$, $f(x) = \|x\|_X$ (Corollary 5.7). Then

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

□

Now we introduce the notion of weak-* convergence which is a type of weak convergence for a sequence $\{f_n\}$ in X^* defined in terms of the action of the f_n on elements of X.

Definition 7.8. A sequence $f_n \in X^*$ converges weakly-* to f, written

$$f_n \overset{*}{\rightharpoonup} f,$$

if $f_n(x) \to f(x)$ for every $x \in X$.

Proposition 7.9. -

(i) Weak- limits are unique, and weakly-* convergent sequences are bounded,*

(ii) weak convergence implies weak- convergence, and*
(iii) weak- convergence implies weak convergence if X is reflexive.*

Proof. *(i)* Uniqueness follows directly from the definition: if $f, g \in X^*$ with $f(x) = g(x)$ for all $x \in X$, then $f = g$. We can apply the uniform boundedness principle directly to the sequence $\{f_n\}$ to show it is bounded (this is identical to the proof of Prop 7.6).

(ii) By (7.1) we have $X \subset X^{**}$, using

$$G_x(f) = f(x) \qquad \text{for all} \qquad f \in X^*. \tag{7.10}$$

By weak convergence we have

$$G(f_n) \to G(f) \qquad \text{for all} \qquad G \in X^{**}. \tag{7.11}$$

In particular we have

$$G_x(f_n) \to G_x(f) \qquad \text{for some} \qquad x \in X.$$

Which gives by definition of G_x,

$$f_n(x) \to f(x) \qquad \text{for all} \qquad x \in X.$$

which is weak-* convergence.

(iii) If X is reflexive then every element of $G \in X^{**}$ can be written as G_x for some $x \in X$ (see 7.10). Now for each G we have

$$G_x(f_n) = f_n(x) \to f(x) = G_x(f)$$

by weak-* convergence. Since (7.11) holds for all $G \in X^{**}$, weak convergence follows. □

Acknowledgments. It is a pleasure to thank my mentors, Jessica Lin and Yan Zhang, for their patience and enthusiasm in teaching me functional analysis. I would like to especially thank Jessica Lin for carefully and thoroughly helping me revise this paper.

References

1. Haïm Brezis, *Analyse fonctionnelle*, Collection Mathématiques Appliquées pour la Maîtrise. [Collection of Applied Mathematics for the Master's Degree], Masson, Paris, 1983, Théorie et applications. [Theory and applications]. MR MR697382 (85a:46001)
2. Gerald B. Folland, *Real analysis*, second ed., Pure and Applied Mathematics (New York), John Wiley & Sons Inc., New York, 1999, Modern techniques and their applications, A Wiley-Interscience Publication. MR MR1681462 (2000c:00001)
3. James C. Robinson, *Infinite-dimensional dynamical systems*, Cambridge Texts in Applied Mathematics, Cambridge University Press, Cambridge, 2001, An introduction to dissipative parabolic PDEs and the theory of global attractors. MR MR1881888 (2003f:37001a)