HAHN-BANACH THEOREMS

NATHAN SHKOLNIK

ABSTRACT. In this paper, we will first go through some preliminary definitions and go through some important results about Banach Spaces and Dual spaces. Then we will introduce the Hahn-Banach extension theorem and some of its equivalents. Using the extension theorem we will discuss the functional implications and then move on to the Hahn-Banach Separation theorems. Using this we will discuss open problems related to Hahn-Banach.

1. INTRODUCTION

The Hahn-Banach theorem is one of the central tools in functional analysis. It arose from attempts to solve problems that require infinite systems of equations. From when it was discovered in 1927, it has undergone some more generalizations, extending from a theorem about continuous, linear, real-valued functionals to a theorem that can be applied to sublinear, complex-valued functionals and even found its way into the study of convex spaces and convex geometry.

For, the first part of this paper we will introduce some preliminary concepts and some theory about Banach spaces. In the following sections, we will discuss the axiom of choice and some properties of *weak* and *weak** topologies, with the main result for this section being Banach Alaoglu's theorem. We will then introduce the Hahn-Banach theorem for extending functionals, its relation to measure, and introduce some preliminary concepts about measure that will be used later.

The last section of this paper will focus on the geometric Hahn-Banach theorem and the Hahn-Banach separation theorems and its relationship to locally convex sets. And then we will focus on two open problems. One discusses the unique extensions of functionals in L^p spaces. And the other problem is the moment problem. We will see how Hahn-Banach allowed us to change the moment problem into a different, hopefully easier problem to solve.

PREREQUISITE KNOWLEDGE

To discuss the Hahn-Banach theorem and the corresponding Banach spaces we need to define both what a vector space is and what a Cauchy sequence is. Unlike in normal linear algebra, where the vector spaces are of finite dimension. However, for the Hahn-Banach theorems and Banach spaces the dimensions of the vector spaces aren't necessarily finite.

Definition 1.1. *Vector Space A vector space V over a field of scalars $\mathbb{K}$ with elements v called vectors satisfying the two following properties.*

- *There exists a binary operation $+$ such that For any two elements v, v_1 in a vector space V, $v + v_a = v_b$*
- *There exists an operation $\cdot$ such that $a \times v = v_n$ for all $a \in \mathbb{K}$*

For the sake of this paper the field of scalars $\mathbb{K}$ will either be the field $\mathbb{C}$ or $\mathbb{R}$

Date: June 2023.

Definition 1.2. *A Cauchy sequence is a bounded sequence $\{X_\alpha\}_{\alpha \in A}$ such that there exists an $n \in \mathbb{N}$ such that $|x_n - x_m| < \varepsilon$ for some $m, n \in A$ where $m > n$*

2. Banach Spaces

When talking about vector spaces there are certain subsets which has nice properties. The first type of space is a Banach space. A Banach space is a complete vector space which will be formally defined below.

Definition 2.1. *A Norm on a space is an operation $\|\cdot\|$ on a vector space V over a field of scalars $\mathbb{F}$ such that the following three properties are satisfied.*

- *Subadditivity $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in V$*
- *Absolute Homogeneity $\|\alpha x\| = \alpha\|x\|$ for all $x \in V$ and $\alpha \in \mathbb{K}$*
- *Positive definiteness $\|x\| = 0$ if and only if $x = 0$*

One of the classic examples and often used norms are the ℓ^pnorms. Which are defined as

$$\|\cdot\|_p = (\sum_n |a_n|^p)^{1/p}$$

where the $a_n's$ are elements of a vector and $1 \leq p < \infty$. The familiar example of an ℓ^p norm is the ℓ^2 norm which is the euclidean distance norm that is defined on euclidean space. Under this norm distances between two vectors x, y is defined as $\|x - y\|_2$ which is analogous to $\sqrt{x^2 - y^2}$. If we include the ∞ for the value of p we get the ℓ^∞ norm, which is otherwise known as the max norm. It is defined by

$$\|\cdot\|_\infty := \sup\{|a_1|, |a_2|, |a_3|, \ldots |a_n|, |a_{n+1}| \ldots\}$$

. One can show that the limit as p approaches ∞ converges to the infinity norm.
When we have a sequence that is an infinite sum, we run the risk of having norms whose values aren't finite.So, in order to get rid of the vectors whose norm is infinite we define ℓ^p spaces as follows.

Definition 2.2. *An ℓ^p space is a space containing all infinite sequences $\{x_i : \|x_i\|_p < \infty\}$*

In a normed space, we can also define a metric using norms. Given two points $x.y \in V$, a metric $d(x, y)$ id defined to be $\|x - y\|$ which, by the triangle inequality is less then or equal to $\|x\| - \|y\|$

Definition 2.3. *A Banach Space X is a normed vector space such that all Cauchy sequences $\{x_n\} \to x \in X$.*

Theorem 2.4. *If $A \subseteq X$ is a closed subset of a Banach Space X it is also a Banach Space*

Proof. Let $\{v_i\}_i$ be a Cauchy sequence in A. Since X is a Banach space and $\{v_i\}_i$ is a Cauchy sequence in X $\{v_i\}_i \to v \in X$. However, since A is closed in X, any sequence $\{v_n\}n \in A$ converges to a $v \in A$. This means that every sequence in A converges in A. □

Proposition 2.5. *A linear operator $T : X \to Y$ is a continuous function between two Banach Spaces if and only if there exists a $C \in \mathbb{F}$ such that*

$$\|Tx\| \leq \alpha\|x\| \text{ for all } x \in X \tag{1}$$

Proof. First we show that if T is a bounded linear operator then it is continuous. Take a sequence $v_n \to v$ for some $v \in X$. Then by (1) there exists a C such that

$$\|Tv_n - Tv\| \le C\|v_n - v\|$$

and since $v_n \to v$ in X that means that as $n \to \infty, \|v_n - v\| \to 0$. And since the norm operator is always non negative and since $\|Tv_n - Tv\| \le 0$ as shown by the limit above. That means that as $n \to \infty, \|Tv - Tv_n\| \to 0$ which means that the sequence Tv_n is convergent in Y which means that T is continuous.

Now we show the other way around. Let $T : X \to Y$ be continuous function between X and Y. This means that

$$T^{-1}(B_Y(0,1)) = \{x \in X : Tx \in B_Y(0,1)\}$$

. Then since $T(0) = 0$ and 0 is in $B_Y(0,1)$ that means that 0 is in $T^{-1}(B_Y(0,1))$ and since the T is continuous, there exists an $B_X(0,\varepsilon)$ which is contained $T^{-1}(B_Y(0,1))$ which implies that $T(B_X(0,\varepsilon))$is contained in $B_Y(0,1)$. Now we take any $v \in V \setminus \{0\}$ (since $0 \le 0$ means that the 0 vector already satisfies the boundedness property.) Then, scaling v by $\frac{\varepsilon}{2\|v\|}$ will give a vector contained in $B_X(0,\varepsilon)$ which implies $T(\frac{\varepsilon}{2\|v\|}v) \in B_Y(0,1)$ and therefore $\|T(\frac{\varepsilon}{2\|v\|}v)\| \le 1$ and pulling out scalars and through using the homogeneity of the norm we get that $\|Tv\| \le \frac{2}{\varepsilon}\|v\|$ which means that T is bounded by $C = \frac{2}{\varepsilon}$ □

Definition 2.6. *We define $\|f\|$, the norm of f to be $\{sup\|f(v)\| : \|v\| \le 1\}$. This is called the operator norm.*

Definition 2.7. *We call a series summable if*

$$\sum_n^\infty v_n < \infty$$

and we call the series absolutely summable if

$$\sum_n^\infty \|v_n\| < \infty$$

Theorem 2.8. *If $\sum_n v_n$ is summable, then the sequence of sums of the form $\{\sum_{n=1}^m v_n\}_{m=1}^\infty$ is a Cauchy sequence.*

Proof. Let N_a denoted $\{\sum_{n=1}^m v_n\}_{m=a}$. So then, using this definition we see that $|N_a - N_{a+1}| = v_{a+1}$. Since the series $\sum_n v_n$ is summable that means that $\lim n \to \infty v_n = 0$ which means that there is eventually an a such that $v_{a+1} \to 0$ □

Theorem 2.9. *A Vector space is a Banach Space if and only if every absolutely summable series is summable.*

Proof. We will start by proving the forwards direction. let V be a Banach Space, then, since V is complete, any absolutely summable series is Cauchy in V and therefore converges in V. Now we go the other way. Assume that every absolutely summable series in V is summable. Take a sequence $\{v_n\} \in V$. Them we just need to show that there exists a subsequence in $\{v_n\}$ that converges, since that will imply that entire series converges since there is only a finite number of elements not in the subsequence. This will therefore imply a convergent

Cauchy sequence.
So let's choose an $N_k \in \mathbb{N}$ such that for any two $n, m \geq N_k$.

$$\|v_m - v_n\| < 2^{-k}$$

We now define

$$s_k = N_1 + N_2 + \cdots + N_K$$

We now have a sequence of n_k where for all $k \in \mathbb{N}, n_k < n_{k+1}$ and now we can show that this series converges since for $n_k > N_k$

$$\|v_{n_k} - v_{n_{k+1}}\| < 2^{-k}$$

and therefore the series

$$\sum_{k\in\mathbb{N}} \|v_{n_k+1} - v_{n_k}\| = \sum_{k\in\mathbb{N}2^{-k}} = 1$$

by how we defined the n_k. This implies that the sequence of partial sums

$$\sum k = 0^m (v_{n_{k+1}} - v_{n_k}) = v_{n_m+1} - v_{n_1}$$

and now we can add back v_{n_1} and since that is one finite element that doesn't change convergence. And means that we found a summable subsequence which shows that we found a convergent sequence. And this means that we have our Cauchy property of a Banach space. □

This is something that is easier to check then checking for Cauchy convergence. And we will use this definition of Banach space to describe a space of bounded linear operator.

Theorem 2.10. *$\mathcal{B}(X,Y)$ the space of all bounded linear functions from X to Y is a Banach space if Y is a Banach space*

Proof. For this proof we will used the previous theorem and show that every absolutely summable series in $\mathcal{B}(X,Y)$ is summable.
So we choose a sequence of linear operators $\{T_n\}$ such that

$$S = \sum_n \|T_n\| < \infty$$

This gives us an absolutely summable series of $\|T_n\|s$, where the norm of T_n is just the operator norm. Now to show that this space is a Banach space we just need to show that $\sum_n T_n$ is summable.
So, lets find our series convergent candidate. Firstly, we know that for any $x \in X$, $m \in \mathbb{N}$

$$\sum_{n=0}^{m} \|T_n x\| \leq \sum_{n=0}^{m} \|T_n\|\|x\| \leq \|x\| \sum_{n=0}^{m} \|T_n\| = S\|v\|$$

Thus, we have a convergent sequence of partial sums of nonnegative real numbers for any $T_n x \in Y$.An since$\sum_n T_n x$ is an absolutely summable series in Y and since Y is a Banach space we have that $\sum_n T_n x$ is a summable series in Y. Now we define the sum of all these operators to be its own operator $T : X \to Y$ where T is defined as

$$Tx = \lim_{m\to\infty} \sum_{n=0}^{m} T_n x$$

and this is our candidate for the summation convergence. So let's show that this is in fact a bounded linear operator. So first let's show linearity.
To do this we will chose any $\lambda_1, \lambda_2 \in \mathbb{K}$ and any $x_1, x_2 \in X$ and

$$T(\lambda_1 X_1 + \lambda_2 x_2) = \lim_{m\to\infty} \sum_{n=0}^{m} T_n(\lambda_1 x_1 + \lambda_2 x_2)$$

and now since each T_n is linear

$$= \lim_{m\to\infty} \lambda_1 \sum_{n=1}^{m} T_n x_1 + \lim_{m\to\infty} \lambda_2 \sum_{n=1}^{m} T_n x_2$$

which then converges to

$$= \lambda_1 T x_1 + \lambda_2 T x_2$$

Thus, T_v is a linear operator.
And we now show that this operator is a bounded operator. Let $x \in$X the

$$Tx = \lim_{m\to\infty} \|\sum_{n=1}^{m} T_n x\| \leq \lim_{m\to\infty} \sum_{n=1}^{m} \|T_n x\|$$

and now, by the triangle inequality, this is bounded by

$$\|\sum_{n=1}^{m} \|T_n\| \|x\|\| \leq S\|x\|$$

and thus, T is a bounded linear function.
Now that we have confirmed that our candidate meets the requirements of a bounded linear operator, it remains to show that the series of T_n actually converges to our candidate.
Take any $x \in X$ with $\|x\| = 1$ then we get that

$$\|Tx - \sum_{n=1}^{m} T_n x\| = \|\lim_{m'\to\infty} \sum_{n=1}^{m'} T_n x - \sum_{n=1}^{m} T_n x\| = \|\lim_{m\to\infty} \sum_{n=m+1}^{m'} T_n x\|$$

and now we bring the norm inside the sum and since $\|x\| = 1$ we use the triangle inequality to get

$$\leq \lim_{m'\to\infty} \sum_{n=m+1}^{m'} \|T_n\|$$

and now we get that this is a series of nonnegative numbers

$$= \sum_{n=m+1}^{\infty} \|T_n\|$$

And now, taking the supremum over all v with unit length that

$$\|T - \sum_{n=0}^{m} T_n\| \leq \sum_{n=m+1}^{\infty} \|T_n\| \to 0$$

Therefore, we have a convergent series, since $\sum_{n=m+1}^{\infty} \|T_n\|$ is the tail of a convergent series. So we do indeed have convergence with respect to the operator norm. □

Definition 2.11. *A linear functional φ on a vector space V is a function $\varphi : X \to \mathbb{K}$ is a linear map from a vector space to a field of scalars. (i.e a map from V to K such that $\varphi(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 \varphi(v_1) + \lambda_2 \varphi(v_2)$ for all $\lambda \in \mathbb{K}, v_1, v_2 \in V$*

Functionals are any functions that quantify the vector space. Functionals like norms, measures, and integrals are examples of functionals and being able to describe these is one the reasons that the Hahn-Banach theorem is powerful.

Definition 2.12. *A dual space of a vector space V denoted V^* is the space consisting of all linear functionals on V.*

Theorem 2.13. *The dual space V^* of a vector space V is a vector space.*

Proof. This theorem is simple to check but takes quite some space so it's left to the reader. □

Theorem 2.14. *The continuous dual space V' is a Banach space*

Proof. This follows directly from theorem 2.13 and theorem 2.10. □

Now with this one can describe functionals on a space and begin to talk about functionals on a Vector space. To begin to proof the Hahn-Banach theorem we will use the AC and its equivalents.

3. Axiom of Choice and its equivalents

The axiom of choice is a axiom in the ZFC axiomatic framework. It allows one to choose, via a choice function, certain elements in a collection of sets. A more formal definition is given below.

Definition 3.1. *(Axiom of choice)For all $\{X_\alpha\}_{\alpha \in A}$ there exists a choice function $f : A \to \bigcup_{\alpha \in A} X_A$.*

There are also some statements that can be taken to be equivalent to the Axiom of Choice. And these are the statements that we will need later in order to prove some of the theorems in this paper.

Theorem 3.2. *(Zorns Lemma) If every nonempty chain C in a set $(H, \leq)$ with a partial ordering has an upper bound then there exists a maximal element in H.*

Theorem 3.3. *(Tychnoffs theorem) Let $\{U_\alpha\}_{\alpha \in A}$ be a collection of non empty compact sets. Then $\Pi_{\alpha \in A} U_\alpha$ is compact.*

There are also statements which are strict weaker than the axiom of choice. This means that the axiom of choice implies theses statements but the same can't be said in reverse. The ones we will use are mentioned below.

Theorem 3.4. *(Ultrafilter Lemma)Every filter $\mathcal{F}$ on a set A has an ultrafilter $\mathcal{G}$ containing $\mathcal{F}$*

There is another theorem that is weaker that the axiom of choice which we will discuss later on in the paper.
One of the reasons that the Hahn-Banach theorem is so widely used is that it acts as a sort of weaker version of the axiom of choice. It means that if one doesn't wish to use the axiom of choice, they can still use Hahn-Banach.

Topologies on The dual space

When talking about vector spaces it is sometimes useful to consider them as topological spaces with the properties of vector spaces. These vector spaces that have a topological structure on them are called topological vector spaces. There are three main types of topologies on the dual space that we will focus on for this section. The first is the strong topology. This is the topology that unless said otherwise is usually used by default.

Definition 3.5. *The strong topology, or norm topology on a Topological Vector Space V is the topology induced by open balls of the form $\mathcal{B}(v, r) := \{u \in V : \|v - u\| \leq r\}$*

We define the strong topology on V' in a similar manner to how we will define the weak* topology in the sense that the relationship with the weak and *weak** topology is analogous to the relationship between the strong topology and the strong topology on the dual space. There are two weak topologies that we are going to discuss. One of them is the "weak topology" and the other is the weak* topology.

Definition 3.6. *The weak topology which is induced by the dual space X' is the coarsest topology on X such that for all $x' \in X', x'$ is continuous.*

The weak* topology is defined in a similar manner but instead of being a topology on the initial set it is a topology on the dual space of a vector space V. However to make this definition we need to define a dual space V^{**} of V^* functions on V^* which go from V^* to $\mathbb{K}$. And we define the following function.

$$T_v(v') = v'(v)$$

Definition 3.7. *Let V be a vector space and V^* be the dual space. The weak* topology on V^* is the coarsest topology on V^* such that the maps $T_v(v')$ are continuous.*

4. Banach-Alaoglu

The Banach-Alaoglu theorem is a theorem describing the dual space of a vector space with respect to the *weak**topology. It is a theorem which is weaker than the axiom of choice but is equivalent to the Ultrafilter lemma. This theorem gives important results on the reflexivity and compactness of normed vector spaces. Using this theorems we can show weak convergence of bounded theorems in reflexive spaces.

Theorem 4.1. *Let X be a normed vector space. Then, the closed unit ball under the operator norm in X' is compact with respect to the weak* topology.*

To prove this theorem we will need one extra bit of information on how to represent a collection of functions. We can represent the collection of functions from X to $\mathbb{K}$ as $\Pi_{x \in X}\mathbb{K}$.

Proof. To proof this theorem we will begin by assuming the following proposition which leads almost directly to the theorem.

Proposition 4.2. *Let X be vector space over a field of scalars $\mathbb{K}$. Then we define U to be a subset of X. Now let r be a real number and $B_r = \{k \in K \; |k| \leq r\}$. Endow B_r with its usual topology. Then define*

$$U^{\#} := \{f \in X^* : \sup_{u \in U} |f(u)| \leq 1\}$$

If for every $x \in X$,$r_x > 0$ is a real number such that Ur_x contains x then U^* is a closed and compact subspace of $\Pi_{x\in X} B_{r_x}$, the family of functions from X to the balls of radius r_x. and since this product topology is equivalent to the weak topology that means that $U^\#$ is closed and compact with respect to the weak* topology as well.*

Now we show that the Banach Alaoglu theorem follows from the proposition above and then we will prove the proposition itself.

Let X be a topological vector space and U a subspace of X around the origin. Since U contains the origin, for every $x \in X$ there exists an r_x such that $x \in Ur_x$. Now, the above proposition implies that $U^\#$ is compact with respect to the weak topology. now we show that $U^\#$ is in X', the continuous dual space on X. To show this we just note that any function is $U^\#$ is bounded and by Proposition 4.2 is continuous.

Now we will prove the proposition stated above.

Using Tychnoffs theorem we know that $\Pi_{x\in X} B_{r_x}$ is compact since each B_{r_x} is compact. Since closed and bounded subsets of compact spaces are compact we can just show that $U^\# = \{f \in X' : f(U) \subseteq B_1\}$ is a closed subset of $\Pi_{x\in X} B_{r_x}$ and to show this we can just proof the two following statements.

- $U^\# \subseteq \Pi_{x\in X} B_{r_x}$
- $U^\#$ is a closed subset of $\Pi_{x\in X}\mathbb{K}$

Proof of first statement
Let $\pi_z := \Pi_{x\in X} K \to K$ be a projection function to the zth coordinate of the product space. Then to show that statement 1 holds we just need to show that the $\pi_x(U^\#) \subseteq B_{r_x}$ holds for all $x \in X$. So let x be fixed. Then, since $\pi_x(f) = f(x)$ for all $f \in U^\#$ is is sufficient to show that $f(x) \in B_{r_x}$.Since, by definition of r_x in the proposition, $x \in r_x U$ we can define $u_x = x * \frac{1}{r_x} \in U$. Since $\sup(f(x)) \leq 1$:

$$\frac{1}{r_x}|f(x)| = |f(\frac{1}{r_x}x)| = |f(u)| \leq \sup_{u\in U} f(u) \leq 1$$

which implies $f(x) \leq r_x$ and therefore $f(x) \in B_{r_x}$
Now we will show that proposition 2 holds.
Since X^* is a closed subset of $\Pi_{x\in X}\mathbb{K}$ the set

$$\begin{aligned} U_{B_1} &:= \{(f_x)_{x\in X} \in \Pi_{x\in X} K : f_u \in B_1 \text{ for all } u \in U\} \\ &= \Pi_{x\in X} C_x \text{ ,where } C_x := \begin{cases} B_1 & \text{if } x \in U \\ \mathbb{K} & \text{if } x \notin U \end{cases} \end{aligned}$$

is closed under the product topology on $\Pi_{x\in X}\mathbb{K}$ since it is a product of closed sets. And therefore $U_{B_1} \cap X^\# U^\#$ is closed since it is the intersection of two closed sets. □

An important theorem that follows from this is Jame's theorem.

Theorem 4.3. *(James' Theorem) A Banach space X is reflexive if and only if for all $f \in X'$ there exists an $a \in X$ with $\|A\| \leq 1$ such that $f(a) = \|f\|$*

By the way that we will define reflectivity and double duals, we can infer that this meas that a Banach space is reflexive if and only if the closed unit ball in X is weakly compact. Banach-Alaoglu as a statement is strictly stronger than Hahn-Banach which means that we can prove Hahn-Banach By using Banach-Alaoglu or it's equivalents, namely Ultrafilter lemma. [Lux62]

5. Hahn-Banach extension theorems

The Hahn-Banach theorem arose from attempts to solve infinite systems of linear equations. Infinite systems were needed to solve problems such as the Fourier cosine problem, or the moment problem. It was built on the back of other theorems. The first mention of something that can be said to be the precursor to the Hahn-Banach theorem was in a 1907 paper by Frigyes Riesz. However, it was later in the 1920s when Eduard Helly laid the basis of what would become proof of the Hahn-Banach theorem. Eduard Helly had been interested in finding a proof for the Hamburger problem(a type of moment problem) and his way to the solution led him to come across a trick that Banach would later use to prove his theorem. Helly proved a theorem which worked for only a specific class of cases. In his 1927 paper, Banach proved a Hahn-Banach theorem and gave some credit to Helly. He then in 1932 proved another version of the Hahn-Banach theorem but wasn't as generous with his crediting.
The Hahn- Banach theorem is one of the most important theorems in functional analysis. It allows one to define specific functionals on a vector space. There are many versions Hahn-Banach theorem which have very similar statements, uses, and proofs. Since the proofs are similar throughout most of the theorems we will give a proof for the continuous norm preserving version over $\mathbb{C}$ and the other proofs follow a similar line of reasoning.

Theorem 5.1. *Let V be a normed vector space, and let $M \subseteq V$ be a subspace. If $u : M \to \mathbb{C}$is a linear map such that $|u(t)| \leq C\|t\|$ for all $t \in H$ (in other words, we have a bounded linear functional), then there exists a continuous extension $U : V \to \mathbb{C}$ (which is an element of V') such that $U|_H = u$ and $\|U(t)\| \leq C\|t\|$ for all $t \in V$ (with the same C as above).*

To prove this theorem we will use the following lemma to first show that it is possible to continuously extend a linear functional to a larger subspace of a vector space and then we will show that after extending this functional we eventually get to a functional which is defined on the whole space.

Lemma 5.2. *Let V be a normed space, and let $M \subseteq V$ be a subspace. Let $u : M \to C$ be linear with $|u(t)| \leq C\|t\|$ for all $t \in M$. If $\notin M$, then there exists a function $u_a : M_a \to \mathbb{C}$ which is linear on the space $M_a = M + \mathbb{C}x = \{t + ax : t \in M, a \in \mathbb{C}\}$, with $u_a|_{M_a} = u$ and $|u_a(t_a)| \leq C|t_a|$ for all $t_a \in M_a$.*

Proof. Proof of lemma. First we'll show that any element $m_a \in M_a$ can be represented uniquely as a $t + ax$ for $t \in M, x \in M_a \setminus M$ and $a \in \mathbb{C}$. To see this assume that

$$t + ax = t_1 + a_1x \implies t - t_1 = (a_1 - a)x$$

which implies that $x \in M$ which is a contradiction unless $a = a_1$ which then means that $t = t_1$. This means that we have a well-defined way to define our continuous extension. Now we choose a $\lambda \in \mathbb{C}$ such that

$$u_1(x + ta) = u(t) + a\lambda$$

and this map is well defined on M_1 and therefore the map u_1 is linear. If the bounding constant C is 0 then the extension map is just the zero function which isn't interesting. To simply the further proof we can (without loss of generality) assume that $C = 1$. For there to be a continuous extension we need to show that there exists a λ such

$$|u(t) + a\lambda| \leq \|t + ax\|$$

holds for all $t \in M, a \in \mathbb{C}$. We can ignore the case when $a = 0$ because then the inequality is in M which we know holds by the requirements for the theorem. So now we can divide by $|a|$ to get

$$|u(\frac{t}{-a}) + \lambda| \leq \|\frac{t}{-a} + x\|$$

since a is a scalar we know that $\frac{t}{-a} \in M$ and that means it is sufficient to show that

$$|u(t) - \lambda| \leq \|t - x\|$$

for all $t \in M$ for this we will split λ into its real partα and its imaginary part $i\beta$ for $\alpha, \beta \in \mathbb{R}$. We will do the real part first and the imaginary part can be done in the same manner. We first show that

$$|g(t) - \alpha| \leq \|t - x\|$$

where $g(t)$ is the real part of $u(t)$. Now notice that Re $|g(t)| \leq |u(t)| \leq \|t\|$ since by assumption $u(t)$ is bounded by a C of 1. Since g is real valued

$$|g(t_1)| - |g(t_2))| \leq |g(t_1 - t_2)| \leq \|t_1 - t_2\|$$

. And using the triangle inequality and the previous equation we get that

$$g(t_1) - g(t_2) \leq \|t_1 - x\| + \|t_2 - x\| \implies g(t_1) - \|t_1 - x\| \leq g(t_2) + \|t_2 - x\|$$

by taking the supremum of the left hand side and the infinum of the right hand side we get that

$$\sup_{t \in M} g(t)) - \|t - x\| \leq \inf_{t \in M} g(t) + \|t - x\|$$

Now we insert an $\alpha \in \mathbb{R}$ between the inequality giving us

$$\sup_{t \in M} g(t) - \|t - x\| \leq \alpha \leq \inf_{t \in M} g(t) + \|t - x\|$$

and now we rearrange this equation to get that for all $t \in M$

$$-\|t - x\| \leq \alpha - g(t) \leq \|t - x\| \implies |g(t) - \alpha \leq \|t - x\|$$

And we can repeat this process but instead of using x we will instead use ix this defines a function u_1 on M_1 which is an extension of the original u function. □

And we have now shown that it is possible to extend a function from a subspace to a larger subspace. Now we have to show that it is possible to extend this function to a function that goes over the whole space. To do this we will use Zorn's lemma.

Proof. Proof of Hahn-Banach extension theorem. We will define a space E to be the space of all continuous extensions of u

$$E := \{(v, H) : v \text{is a continuous extension of} u \text{and} M \subseteq H \subseteq V\}$$

We the define a partial ordering $\preccurlyeq$ on the elements of E defined by

$$(v_a, H_a) \preccurlyeq (v_b, H_b) \text{ if } H_a \subseteq H_b, v_b|_{H_a} = v_a$$

one can check that this is a partial order for themselves. Now, in order to be able to use Zorns lemma we have to in some way create chains in E. So let C = $(v_i, H_i)_{i\in I}$ be a chain and i be an indexing set. We claim that every chain C has an upper bound.
Let $H = \bigcup_{i\in I} H_i$. We will check that this is a subset of E. First we take two elements $x_1, x_2 \in H$ and then $z_1, z_2 \in \mathbb{C}$. Then let's take two sets N_{i_1}, N_{i_2} such that $x_1 \in N_{i_1}, x_2 \in N_{i_2}$ be the partial ordering defined on N_i and since $N_i \in C$ we know that $x_1, x_2 \in N_{i_2}$. And now since the N_i's are closed we know that $z_1x_1 + z_2x_2 \in N_{i_2} \subseteq N$ and this makes N a subset. Now we have to show that (v, N) is a maximal element for chains in E. And we define v to be the extension of u to that acts ion N. More formally $v : N \to \mathbb{C}$ so that for any t_i in N_i , $v(t_i) = v_i(t_i)$. And for any $n \in N_i v(n) = v_i(n)$ and therefore when v is restricted to any N_i it is defined as the respective v_i which means that any $(v_i, N_i) \preccurlyeq (v, N)$ making (v, N) a upper bound for any chain. Now we have the prerequisites needed to use Zorns lemma. First let (U, N) be a maximal element in E. We claim that $N = V$. Suppose that this isn't the case. Then using the above lemma, there exists an x such that there is a continuous extensionU_1 of U to $N + \mathbb{C}x$. However, this would contradict (U, N) being a maximal element since $(U_1, N = \mathbb{C}x) \preccurlyeq (U, N)$. Therefore, $N = V$ and U is a continuous extension of u acting on all of V □

The following theorem deals with sublinear functionals dominating other functionals and this therefore acts as a more generalized version of the above proof.

Definition 5.3. *A function $p : X \to \mathbb{K}$ is called sublinear if*

- $p(x + y) \le p(x) + p(y)$ *for all* $x, y \in X$
- $p(\lambda x) = \lambda p(x)$ *for all* $\lambda \in \mathbb{K}, x \in X$

Theorem 5.4. *(Hahn-Banach Sublinear functional) Let p be a sublinear function from E to* $\mathbb{R}$. *Let* E_0 *be a vector subspace of E and let* f_0 *be a linear function from* E_0 *to* $\mathbb{R}$ *for which*

$$f_0(x) \le p(x) \text{ for all } x \in E_0$$

Then there exists a linear function f from E to $\mathbb{R}$ *that extends* f_0*and for which*

$$f(x) \le p(x) \text{ for all } x \in E.$$

When we have that $f(x) \le p(x)$ *for all* $x \in E$ *we say that* f *is dominated by* p

Theorem 5.5. *(Hahn-Banach Theorem for normed spaces Version) Let E be a normed vector space and let* E_0 *be a vector subspace of E. Let* f_0 *be a continuous linear function from* E_0 *to R. Then there exists a continuous linear function from E to* $\mathbb{R}$ *that extends* f_0*such that*

$$\|f(x)\| = \|f_0(x)\|$$

We won't write out the full proof for these two scenarios since the proofs are practically the same as the previous proof and there is not much reason to repeat the steps. However, the fact that there are many types of Hahn-Banach extension theorems means that depending on what properties we want the functional to have. This is one of the reasons that the Hahn-Banach theorem is such a powerful theorem for functional analysis. It allows us to define functionals on the entire vector space based solely on how the function is defined on a specific subspace. One simple example of how the Hahn-Banach theorem can be used to define functionals in the dual space is the following proposition.

Corollary 5.6. *Let* V *be a normed vector space then, for all* $v \in V$ *there exists a function* ϕ *such that* $\|\varphi\| = 1$ *and* $\|\varphi(v)\| = v$

Proof. Define a functional $\phi : \mathbb{C}v \to \mathbb{C}$ by $\phi(\lambda v) = \lambda\|v\|$.Then $|\phi(z)| \leq \|z\|$ for any $z \in \mathbb{C}v$ and clearly $\phi(v) = \|v\|$. Therefore by Hahn-Banach extension theorem.(We use the version that we proved). There exists a φ in V^* extending ϕ such that $\|\varphi(z)\| \leq \|z\|$ for all z. So, now we've found our desired functional φ with $\|\varphi(v)\| = \|v\|$ and $\|\varphi\| = 1$ □

This corollary shows how there is, in a sense a way to have well defined elements in the dual space of a vector space, that allows one to choose specific functionals as they fit their need. This corollary also implies the following proposition.

Definition 5.7. *We call a operator $T : V \to W$ isometric, or measure preserving is for all $v \in V$, $\|v\| = \|Tv\|$*

Proposition 5.8. *There exists a natural map from V to V^{**} which is isometric.*

We will define the elements of the dual space $T_v : V' \to \mathbb{C}$ by $T_v(v') = v'(v)$ for any $v \in V$ and any $v' \in V'$.

Proof. We will leave it as an exercise to the reader to check that $v \mapsto T_v$ is a linear operator. We will show that $v \mapsto T_v$ is bounded.
Take any $T_v(v')$ then

$$|T_v(v')| = |v'(v)| \leq \|v\|\|v'\|$$

Using this we see that $\|T_v\| \leq \|v\|$ and therefore $\|T\| \leq 1$ Now we need to show that $\|T\| = 1$. To do this we show that $\|v\| \leq \|T_v\|$ So now using corollary 31 we know there exists some $\varphi \in V'$ such that $\|\varphi = 1\|$ and $\varphi(v) = \|v\|$ and therefore we have that

$$\|v\| = f(v) = |f(v)| = T_v(f) \leq \|T_v\|\|f\|$$

and therefore we have that $\|v\| = \|T_v\|$ and therefore T is isometric. □

Corollary 32 also leads us to another interesting result.

Proposition 5.9. *For a closed subspace $A \subseteq X$ there exists an $f \in X^*$ on every $x \in X/A$ such that $\|f\| = 1, f|_A = 0$ and $f(x) = \inf_{a \in A}\|x - a\|$*

Using the above proposition we can define an indicator which will give us 1 when we are in the desired subspace and otherwise.

Proof. The proof for this proposition is very similar for the proof for the above corollary. Just define the function f to be the infimum of the distance from any point x to a point $a \in A$ □

Definition 5.10. *We call a space reflexive if the map $V \mapsto V^{**}$ is bijective.*

Reflexive spaces will come up later and are spaces where there is, in a sense the same amount of elements in V and V^*
The ℓ^p space is reflexive for all $1 < p < \infty$ (since the dual of p is q and the dual of q is p again). However, using Hahn-Banach we can show that the ℓ^1 and ℓ^∞ are not reflexive since the dual of ℓ^∞ isn't ℓ^1. [Mel21] Another form of Hahn-Banach that acts is a similar way to the form of .

Theorem 5.11. *(Dominated Hahn-Banach) Let ρ be a function between a vector space X and R. And let $f : M \to \mathbb{R}$ where $M \subseteq X$. Such that $|f(x)| \leq \rho(x)$ for all $x \in M$. Then there exists and F so that $|F(x)| \leq |\rho(x)|$ X for all x in*

We will not show the proof for this version of the theorem as well since it is very similar to the above proof.
The main difference between this version of the Hahn-Banach theorem and the initial Hahn-Banach theorem that we mentioned is that this version of the theorem doesn't require there to be a norm on the vector space and also doesn't require for the functional on the vector space to be continuous functional.

6. Measure and Banach

One of the fields where Hahn-Banach is used is in measure theory. Since measures are types of functionals, there is a place where Hahn-Banach comes into play with dealing with measure. To define measure we firstly need to define what a $\sigma - algebra$ is.

Definition 6.1. *A $\sigma - algebra \Sigma$ on X is a collection of subsets of X such that if $U \in \Sigma$ then $U^c \in \Sigma$ and such that Σ is closed under countable unions.*

Definition 6.2. *Let Σ_X be a $\sigma - algebraonasetX$ then we call $\mu : \Sigma_x \to [0, +\infty]$*

- *(Non-negativity) $\mu(A) \geq 0$ for all sets A with $\mu() = 0$.*
- *(Countable Additivity) For any countable collection of disjoint $E_n \in \Sigma_X$*

$$\sum_{k}^{\infty} \mu(E_k) = \mu(\bigcup_{k}^{\infty} E_k)$$

There is a certain type of $\sigma - algebra$ that we will be interested in later. It is called the Borel $\sigma - algebra$,denoted $\mathfrak{B}(X)$ and it is a $\sigma - algebra$ consisting of only/closed sets induced by the topology on that set.
Hahn-Banach leads to two conclusions about the extension of Lebesgue measure.

Definition 6.3. *The Lebesgue outer measure μ^* of a set A in the sigma algebra is*

$$\inf(\sum_{k=1}^{\infty}(vol(A_k)) : \textit{ for all } A_k \textit{ such that } \bigcup_k (A_k) \supseteq A$$

. The Lebesgue measure $\mu^*$is then defined to be the same as the Lebesgue outer measure if

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E^c)$$

This is the typical measure that is used for subsets of Euclidean space. And it is used to define things like Lebesgue integration and L^p spaces. So, the Hahn-Banach theorem can be used to show extension of a finitely-additive, translation invariant Lebesgue measure up to $\mathbb{R}^s$ and can be used to show the nonexistence of such a measure in $\mathbb{R}^p$ where $p \geq 3$ [FW91]. This existence of a non-Lebesgue measurable set can be also used to imply the Banach-Tarski theorem.

Theorem 6.4. *(Banach-Tarski) Given a solid ball in three-dimensional space, there exists a decomposition of the ball into a finite number of disjoint subsets, which can then be put back together in a different way to yield two identical copies of the original ball.*

The Banach-Tarski paradox is a statement that is equivalent to the Hahn-Banach theorem. More details are here [Paw91]. The Hahn-Banach theorems role in this proof is to show that we can define a measure which then leads to the proof of the theorem.
There is another type of measure that we will define in this section called Radon Measure.

Definition 6.5. *Radon measure μ is a measure on a Borel algebra induced by a Hausdorff topology $\mathfrak{B}(X)$ such that for a compact subset $C \in \mathfrak{B}(X), \mu(C)$ is finite, all sets in the Borel algebra have an regular outer measure and all open sets have the regular inner measure.*

As I mentioned earlier, we will not need to use radon Measure now but it will come up later.

7. Locally convex Sets and Hausdorff spaces

Normally, when working with a topological space it is often the case that the space is a Hausdorff space. A Hausdorff space is a topological space in which points are separated.

Definition 7.1. *A space X endowed with a topology τ is Hausdorff if for any point $x, x_1 \in X$, there exists a neighborhood V around x and neighborhood U around x_1 such that $U \cap V = \emptyset$*

These types of spaces also have a lot of properties which makes them nice to work with. One of these properties is that limits of nets and filters are unique. Also, in Hausdorff spaces, compactness of a subset implies that the set is closed.
These are just some of the properties that make Hausdorff spaces unique and nice to work with. It is also the case that all sets can be extended to be a Hausdorff set. However, in certain topologies using Hahn-Banach we can impose Hausdorff separation onto the topological vector space.
Locally convex spaces are spaces such that the neighborhood basis around the origin consists of convex, balanced sets. Convex sets are sets such that any two points in the set can be connected with a line segment that stays in the set, and balanced sets are sets such that any line from a pint to the origin is entirely within the set. The formal definition is below.

Definition 7.2. *We call V a locally convex set if the neighborhoods around the origin are sets $\{C_a\}$ such that*

- *(Convex) If $x, x_1 \in C$ then the line $px + (1-p)x \in C$ where $0 \leq p \leq 1$*
- *(Balance) For any point $x \in C$ and for any scalar c such that $|c| \leq 1$, $cx \in C$*

The reason that we only have to define this for the neighborhood around the origin is because we can shift any of the sets by writing a neighbor hood of x as $O(x)$ which is equivalent to $O(0) + x$ in a vector space.

8. Geometric Separation Hahn-Banach

The Hahn-Banach separation theorems are a set theorems which we can use to show and build separating hyperplanes between convex sets in a vector space. For this theorem we should define cones and hyperplanes are and what the closure of a set is.

Definition 8.1. *The closure of a set X denoted by $\bar{X}$ is $\bigcap_\alpha D_\alpha$the intersection of all closed sets containing X.*

Definition 8.2. *A hyperplane H is a subset of a vector space V that is of one dimension lower and is of the form*

$$H = \{x : f(x) = a\} \text{ for some } a \in \mathbb{R}, f \in E^*$$

The closure is the smallest closed set that contains a set Y. And the hyperplane acts as a sort of line that, in a sense, splits a space in half with one half being elements in a V such that $f(v) \leq a$ and the other half being$f(v) \geq a$. The original Geometric Hahn-Banach

theorem, most often credited to Mazur deals with the existence of this hyperplane.
The Geometric Hahn-Banach Theorem

Theorem 8.3. *Let X be a topological vector space over $\mathbb{R}$, N a linear subspace of X, and O a non-empty open convex subset of X such that $N \cap O = \emptyset$. Then there exists a closed hyperplane H of X such that*

$$N \subseteq H \text{ and } H \cap O = \emptyset$$

Proof. The proof for this will be done in 3 parts. Fist we will show that there indeed exists a maximal linear subspace $H \supseteq N$ disjoint from O. Then we show that H is closed in X with respect to the topology induced by a seminorm(a norm without Positive definiteness). And then we will finally show that H is indeed a hyperplane.

8.1. **H Is a maximal extension of N.** This proof is just an application of Zorn's lemma. First we consider the family of linear subspaces E of X such that

$$N \subseteq E \text{ and } E \cap O = \emptyset$$

. And now, just with how we proved the maximal extension for the Hahn-Banach Extension Theorem, we will choose a chain C in X under the ordering of $\subseteq$ and then we take the upperbound to be the union of all $S's$ in the chain. And now, by Zorns we have a maximal subspace H that contains N and who's intersect with O is empty. This now gives us a subspace over all of X satisfying our desired proprieties.

8.2. **H is closed in X.** Since H is disjoint from O, we know that $H \subseteq O^C$ which implies that $\bar{H}$ is also disjoint from O. And since O is open in X we get that

$$\bar{H} \subseteq \overline{X/O} = X/O$$

And since $\overline{H}$ contains N and is disjoint from O, $\overline{H}$ is in the family of subspaces E. But, since H is a maximal element that means that H coincides with $\overline{H}$

8.3. **H is a Hyperplane.** To prove this we will show that $\dim(X/H) = 1$, which will be sufficient to show that H is a hyperplane.
Let $\phi : X \to X/hH$ be the quotient mapping. Then, since ϕ is an open mapping(it takes open sets in X to open sets in X/H), $\phi(O)$ is an open convex set which doesn't contain the origin $\hat{o}$ of X/H.(If it did contain the origin that would mean that there is an $x \in O$ such that $\phi(x) = \hat{o}$ and since ϕ is a quotient mapping that means that $x \in H$ which contradicts the "disjointness" of H and O. So now we set

$$C = \bigcup_{a>0} a\phi(O)$$

This makes C a convex cone without the origin point $\hat{(o)}$. We show that $dim(X/H) \geq 1$. Assume that $dim(X/H) = 0$ Then $X/H = \{\hat{(o)}\}$ and therefore $H = X$ which implies that O is empty which contradicts the definition of O. So we will now show that $dim(X/H) \ngeq 2$. Assume for the sake of contradiction that $dim(X/H) \geq 2$ then we only need to show the following two claims in order two get our conclusion.

- Claim 1 The boundary δA of A must contain at least one point x $\neq \hat{o}$.
- Claim 2 The point $-x$ cannot be in A

With Claim 1 we can show that $x \notin A$, since $x \in \delta A$ and since A is open. With both this fact and the second claim, we know that x an $-x$ are both in the complement of A which is in X/H. So we define L to be the line between x an $-x$ Since A is a cone, we can see that $L \cap A = \emptyset$ Then:

- $\phi^-1(L)$ is a linear subspace
- $\phi^-1(L) \cap O = \emptyset$ since $\phi^-1(L) \cap A = \emptyset$ and A is a union of O
- $\phi^-1 \supset H$ because $\cap o = \phi(H) \subseteq L$ but we can say that containment is strict since $x \neq \hat{o}$

However, this now means that H is contained in $\phi^-1(L)$ and that $\phi^-1(L)$ is one of the subspaces S which contradicts the maximality of H. Now we just prove the claims and we will have the finished proof of H being a hyperplane.

- Proof of claim 1 Lets assume that $\delta A = \{\cap o\}$. Then that means that A has an empty boundary with respect to $X/H \setminus \{\hat{o}\}$ And since dim $dim(X/H) \geq 2$ we can see that $X/H \setminus \{\hat{o}\}$ is path-connected and therefore connected. This would now imply that A $= X/H \setminus \{\hat{o}\}$. This contradicts that A is convex since $X/H \setminus \{\hat{o}\}$ isn't convex (due to the missing origin)
- Proof of claim 2 Assume there exists an $-x \in A$. Then we can find a neighborhood U around $-x$which implies that there exists $-U \ni x$. then we can choose elements $j \in U$ and $-j \in -U$ and since A is convex that means that the line between the two points is entirely in A which implies that $\hat{o}$ is in A which contradicts the definition of A.

□

This theorem implies the extension version of the Hahn-Banach theorem.
When the subsets have different additional properties, such as compactness,closeness, and openness, there are stronger theorems that we can use in order to build separations.

Theorem 8.4.
- *If A is an open subset of X and B is a subset of X then there exists and separating function $L : X \to \mathbb{K}$ such that there exists an k such that $L(a) < k \leq L(b)$ for all $a \in A, b \in B$. If B is also open then this could be strengthened to $p(a) < k < p(b)$*
- *If we have a locally convex vector space X with an compact subset A and a closed subset B then there exists a function $s : X \to \mathbb{K}$ such that s strongly separates theses sets (i.e there exist and $v, w \in \mathbb{K}$ such that $\sup f(a) < v < w < \inf f(b)$*

We will only prove the first part. The second part follows from the first proof.

Proof. Take the set $A - B := \{a - b : a \in A, b \in B\}$ One can check that this is an open, convex set not containing the origin o. By applying the Hahn-Banach theorem to $\{o\}$ we get a closed Hyperplane containing $\{o\}$ disjoint from $A - B$. Equivalently that means that there exists a linear functional $f \not\equiv 0$ on X such that $f(A-B) \neq 0$ Then there exists a linear function of X such that

$$L(A - B) > 0 \implies L(A) - L(B) > 0 \implies L(A) > L(B)$$

And now that means that we can take an a such that $L(A) \geq a \geq L(B)$. □

The Hahn-Banach separation theorems separate also give rise to the following theorem regarding cones which is a stronger version of the aforementioned theorems since it allows us to choose the hyperplane constant to be 0.

Theorem 8.5. *Let $C \subseteq X$ be a convex cone in X and $B \subseteq X$ an open subset of X. Then there exists a linear functional $f \not\equiv 0$ $f : X \to \mathbb{R}$ such that $f(c) \geq 0$ for all $c \in C$ and such that $f(a) \leq 0$ for all $a \in B$*

Proof. The proof follows from the previous theorems. Since we have a hyperplane that separates C and B that means that there exists an α such that $f(c) \geq \alpha \geq f(b)$ for all $c \in C, b \in B$. Since C is closed under scalar multiplication we have that

$$f(\lambda c) \geq \alpha \text{ for a } \lambda \in \mathbb{R} \implies \lambda f(c) \geq \alpha \implies f(c) \geq \frac{\alpha}{\lambda}$$

and taking the minimum we get that

$$f(c) \geq \inf_{\lambda > 0} \frac{a}{\lambda} = 0$$

□

When working with topological vector spaces there are some properties which make it nice to work with these spaces. In most cases if we can fine a topological vector space that is Hausdorff, it is nicer to work with than a lot of other spaces. Same goes for spaces that are locally convex. The separation theorems allow us to have weaker initial requirements in order to further assume that a space has both of these properties.

Proposition 8.6. *Let V be a topological vector space then under the weak topology; V is Hausdorff and furthermore, under the weak* topology V^* is also Hausdorff*

Proof. One of the properties that a topological vector space has with respect to the weak topology is that the space is locally convex.
Let $v, w \in V$ with neighborhoods U, Z respectively. Then by The Hahn-Banach separation theorem, we get that there exists a function L such that $L(U) < k < L(Z)$. Which means that the sets are separated by a hyperplane of the form $x \in V : L(x) = k$.
The proof for the dual space is similar but uses the weak* topology and an element from V^{**}. □

9. Unique extensions of functions

While in general it is possible a functional on a space, it is not always guaranteed that this extension is unique. So, there could be many functionals that work to extend a functional on a subspace. in this section we will discuss the times where the extension of functionals is unique. (i.e there exists only one functional g which can extend f from a subset to the whole space.
To set the stage for these functionals we will need to define a new type of norm called a Gateux differentiable norm and we will need to restrict the type of spaces to spaces that are uniformly convex.

Definition 9.1. *We call a norm on a vector space V Gateux differentiable if for all $x, y \in V$*

$$\lim_{t \to 0} \frac{\|x + ty\| - \|x\|}{\|t\|}$$

exists.

Definition 9.2. *A uniformly convex space is a normed vector space such that for every $\varepsilon > 0$ there exists an $\delta > 0$ so that for all x, y where $\|x\| = \|y\| = 1$*

$$\|x - y\| \geq \varepsilon \implies \|\frac{x+y}{2}\| \leq 1 - \delta$$

This type of norm is a restriction that allows for us to make unique extensions. The following theorem is another version of the Hahn-Banach theorem. It is a theorem credited to Hajime Ishihara and it is based on a constructionist axiomatic framework that differs from the ZFC framework. We will not discuss this further but more reading can be found here [BB85].

Theorem 9.3. *(Ishihara's Hahn-Banach) Let U be a uniformly convex space with a Gateux differentiable norm. Let V be a subspace of U and $f : v \to \mathbb{R}$ a nonzero linear functional. Then, there exists a unique extension F of f such that $\|F\| = \|F\|$.*

We will not prove this theorem here as it is based on a different axiomatic system that would take to much time to properly introduce. This above theorem gives rise to the following theorem.

Theorem 9.4. *If the unit ball in X is weakly compact and if the norm in E is Gateaux differentiable then every continuous linear functional on every vector subspace of E can be uniquely extended to a continuous linear functional on E with the same norm.*

And this, by James' theorem, implies that

Proposition 9.5. *If E is a reflexive Banach space and the norm on E is Gateaux differentiable then Hahn–Banach extensions are unique.*

We refer again to [BB85] for more detail. Another example for a search for uniqueness of extensions is the following conjecture.

Proposition 9.6. *Let $B(\ell^2)$ be the Hilbert space of all continuous linear functions from $\ell_2 \to \ell_2$. It follows that $\ell^\infty \subset B(\ell_2)$. We can extend the extreme points of the unit ball of $\ell^{\infty *}$ to the extreme points of the unit ball around $B(\ell_2)$.*

We won't discuss this proposition further due to lack of space but more reading can be found here [BT91]

10. The moment problems

And now we're back to the problem that started it all. Helly's work which was used to establish a basis for the Hahn-Banach theorem was based on trying to solve a certain type of moment problem. Now we will use Hahn-Banach to make progress on the general moment problem. We will discuss the general form for the moment problems. The moment problems deal with the question of "If a sequence of moments exist does a probability measure on a function?". So let's define some of the terms.

Definition 10.1. *The nth moment m^n of a measure μ be a non-negative Borel measure defined on $\mathbb{R}$ is defined by*

$$m_n^\mu = \int_{\mathbb{R}} x^n \mu(dx)$$

If the measures of all these measures exists and is finite then $(m^{\mu}_n)_{n=0}^{\infty}$ is called the sequence of moments of μ. This means that given a certain measure we can create a sequence of moments. The moment problems ask the question in reverse. We will introduce the question in the univariate sense then extend it to a multivariate sense

Definition 10.2. *(Univariate moment problem) Let C be a closed subset of $\mathbb{R}$ with $m := (m^{\mu}_n)_{n=0}^{\infty}$. Does there exist a finite Radon measure μ such that the the $supp(\mu) \subseteq K$ and the moment of the measure is m where a m_n is defined as*

$$m_n = \int_C x^n \mu(dx) \text{ for all } n \in \mathbb{N}$$

If there is such a measure then we call that measure a C-representing measure and therefore a solution to the C-moment problem for m.

To make an easier way to work with these problems, at least through the lens of functional analysis, we will rephrase the moment problem in terms of a linear functional. So first we define a linear functional L_m from $\mathbb{R}[x]$ the set off all polynomials with x to $\mathbb{R}$. Then

$$p(x) := \sum_{n=0}^{N} p_n x^n \text{ and we define } L_m(p) := \sum_{n=0}^{N} p_n m_n$$

then let μ be a C representing measure for m and we get

$$L_m(p) = \sum_{n=0}^{N} p_n m_n = \sum_{n=0}^{N} p_n \int_C x^n \mu(dx) = \int_C p^x \mu(dx)$$

This meas that there is a one to one relationship between the linear functional $L_m : \mathbb{R}[x] \to \mathbb{R}$ and all sequences of real numbers.Now that we found that there is a relationship between sequences of moments and functionals we will be able to rewrote the problem in terms of linear functionals.

Definition 10.3. *(Univariate Moment problem in Functional form)*
Let C be a closed set of $\mathbb{R}$ an $L : \mathbb{R}[x] \to \mathbb{R}$. Does there exist a nonnegative Radon measure μ such that

$$L(p) = \int_{\mathbb{R}} p(x)\mu(dx) \text{ for all } p \in \mathbb{R}[x], supp\mu \subseteq C$$

Now we can extend this problem to higher dimensions. First we replace $\mathbb{R}$ with $\mathbb{R}^d$ where $d \in \mathbb{N}$ and instead of only having one x in the ring $\mathbb{R}[x]$ we define the collection of d variables $x_D = (x_1, x_2 \ldots, x_d)$ and now we can pose the problem for higher dimensions.

Definition 10.4. *Let C be a closed set of $\mathbb{R}^d$ an $L : \mathbb{R}[x_D] \to \mathbb{R}$. Does there exist a nonnegative Borel measure μ such that*

$$L(p) = \int_{\mathbb{R}^d} p(x)\mu(dx) \text{for all } p \in \mathbb{R}[x_{\mathbb{D}}]$$

Now we say that if such a measure exists that it's a C representing measure for L and therefore a solution to the moment problem for L

The following theorem gives another approach to solve this problem. This theorem is due to Reisz and Havilad an it replace the moment problem with another problem.

Definition 10.5. *(Reisz-Havilad theorem). Let $K \subseteq \mathbb{R}^d$ and a linear function $L : \mathbb{R}[x_D] \to \mathbb{R}$. Then, L has a K representing measure if and only if L(Psd(K))≥ 0 where Psd(K) := $\{p \in \mathbb{R}[x_d] : p(x) \geq 0$ for all $x \in K\}$*

Using this we can find a solution to the moment problem. However, the one downside is that this has just turned one hard problem into another hard problem.
To start to make progress on this problems we define the dual space of a cone as follows.

Definition 10.6. *The first dual $C^\vee$ of a cone is given by*

$$C^\vee = \{\ell : X \to \mathbb{R} \text{ linear } : \ell(C) \text{is continuous and } \ell(C) \geq 0\}$$

and we define the double dual as

$$C^{\vee\vee}\{x \in X : \text{ for all } \ell \in C^\vee \ell(c) \geq 0\}$$

Now using the Hahn- Banach separation theorem to prove the following proposition.

Proposition 10.7. *Let X be real vector space endowed with the finest locally convex topology φ. If C is a nonempty convex cone in X, then $\overline{C}$ with respect to φ then $C^\varphi = C^{\vee\vee}$*

Proof. Firstly, we can see that $C^\varphi \subseteq C^{\vee\vee}$. That is because if $x \in c^\varphi$ then for any $\ell \in c^{vee}$ $\ell(x) = 0$ and therefore $x \in C^{\vee\vee}$
On the other hand, suppose there exists $x_0 \in C^{\vee\vee} \setminus \overline{C^\varphi}$ Then by the separation theorem with cones, there exists a linear functional $L : X \to R$ $L(\overline{C}^\phi) \geq 0 \text{and} L(x_o) < 0$. Since $L(C) \geq 0$ and every linear functional is continuous, we have $L \in C^\vee$. This now means that since $L(x_o) < 0$ that $x_o \notin C^{\vee\vee}$ which contradicts the definition of v_o □

Now using this proposition we can create another new form of the measure problem that goes as follows.

Proposition 10.8. *Let $S := \{g_1, ..., g_s\}$ be a finite subset of $\mathbb{R}[x]$ and $L : \mathbb{R}[x] \to \mathbb{R}$ linear. Assume that M_S is Archimedean. Then there exists a K_S-representing measure μ for L if and only if $L(M_S) \geq 0$, where $K_S := \{x_D \in \mathbb{R} : g_i(x_D) \geq 0$ for all $g_i$in a finite set of polynomialsS and where $M_S := \{\sum_i = 0^s \sigma_i g_i : \sigma_i \in \sum \mathbb{R}[x_d]^2, i = 1, 2, 3 \dots s\}$*

Proof. One fact that we will keep in mind comes from Putinar Positivstellesatz (1993) and it states that if M_S is Archemdian(i.e there exists an N in $\mathbb{N}$ such that $N - \sum_{i=1}^d x_i^2 \in M$)

$$Psd(K_S) \subseteq \overline{M_S}^\varphi$$

. Now, lets suppose that $L(M_S) \geq 0$ and we have a topology φ on $\mathbb{R}[x_D]$ Then L is continuous and therefore an element of $(M_s)^\vee$. And, since M_S is Archemidian,

$$Psd(K_S) \subseteq \overline{M_S}\varphi \text{ Which by the above proposition is equal to } {M_S}^{\vee\vee}$$

And this implies that any $p \in Psd(K_s)$ is in $\{M_S\}^{\vee\vee}$. This means that for any $\ell \in \{M_S\}^\vee, \ell(Psd(K_S) \geq 0)$ which means that $L(Psd(K_S) \geq 0)$. And by Reisz-Havilad theorem we get that we have a K_S representing measure for L.
Now we show the other way around. Assume that we have a K_S representing measure μ for L. Then that means that for all $p \in M_S$ we have that

$$L(p) = \int_{\mathbb{R}^d} p(x_d)\mu(dx_d)$$

which is nonnegative since μ is a nonnegative measure and therefore we get that $L(p) \geq 0$ for all $p \in M_S \subseteq Psd(K_S)$ While this doesn't give a solution to the moment problem it simplifies the problem significantly and gives us another way to approach the problem. $\square$

11. Bibliography

References

[BB85] Errett Bishop and Douglas Bridges. *Constructive analysis*, volume 279 of *Grundlehren Math. Wiss.* Springer, Cham, 1985.

[BT91] J. Bourgain and L. Tzafriri. On a problem of Kadison and Singer. *J. Reine Angew. Math.*, 420:1–43, 1991.

[FW91] Matthew Foreman and Friedrich Wehrung. The Hahn-Banach theorem implies the existence of a non-Lebesgue measurable set. *Fundam. Math.*, 138(1):13–19, 1991.

[Lux62] W. A. J. Luxemburg. Two applications of the method of construction by ultrapowers to analysis. *Bull. Am. Math. Soc.*, 68:416–419, 1962.

[Mel21] "Richard Melrose". " introduction to functional analysis lecture notes, "Spring" "2021".

[Paw91] Janusz Pawlikowski. The Hahn-Banach theorem implies the Banach-Tarski paradox. *Fundam. Math.*, 138(1):21–22, 1991.