2. Hahn-Banach Theorem

Definition: A mapping $p : E \to \mathbb{R}$ defined on a vector space E is said to be a *support* function if

(1) $p(tx) = tp(x)$ for all $x \in E$ and $t > 0$.
(2) $p(x + y) \leq p(x) + p(y)$ for all $x, y \in E$.

Remark: It follows that $p(0) = 0$ because $p(0) = p(0/2) = \frac{1}{2}p(0)$.

Example: If $E = \mathbb{K}^n$ and $p(x^1, x^2, \ldots, x^n) = \max\{x^j : 1 \leq j \leq n\}$ then p is a support function but p is not nonnegative.

The Hahn Banach Theorem concerns extending linear mappings $f : E \to \mathbb{K}$ where E is a vector space over $\mathbb{K}$ (that is $\mathbb{R}$ or $\mathbb{C}$). The set of all such linear scalar-valued mappings is the *dual* or adjoint of E and is denoted E^*. In the case $\mathbb{K} = \mathbb{R}$ we have the following result which is a stepping stone to the Hahn Banach Theorem.

THEOREM 2.1. *Suppose the E is a vector space over $\mathbb{R}$ and F is a proper subspace of E and $z \in E - F$. Suppose that $p : E \to \mathbb{R}$ is a support function defined on all of E and $f \in F^*$ and*

$$f(x) \leq p(x) \text{ for all } x \in F.$$

Let G denote the smallest linear subspace of E containing both F and z. Then there is $g \in G^*$ so that

i) $g(x) = f(x)$, *for all* $x \in F$.
ii) $g(x) \leq p(x)$, *for all* $x \in G$

Explicitly $G = \{x + \lambda z \in E : x \in F \text{ and } \lambda \in \mathbb{R}\}$ and $g(x + \lambda z) = f(x) + \lambda g(z)$ (so that i holds) and so the essence of the proof is to show that $g(z)$ can be chosen so that ii) is true.

PROOF. We observe that for any $x, y \in F$,

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

Therefore $f(y) - p(y - z) \leq p(x + z) - f(x)$ for all $x, y \in F$. We may therefore choose $r \in \mathbb{R}$ so that

$$\sup_{y \in F} f(y) - p(y - z) \leq r \leq \inf_{x \in F} p(x + z) - f(x)$$

Define

$$g(x + \lambda z) = f(x) + \lambda r$$

and it is clear that g is defined on $G = \{x + \lambda z : x \in F, \lambda \in \mathbb{R}\}$ and extends f so that i) is satisfied.

It remains to show that $g(x+\lambda z) \leq P(x+\lambda z)$. Certainly this is so if $\lambda = 0$ because $g(x) = f(x) \leq p(x)$ by the hypothesis. Suppose then that $\lambda > 0$. Then

$$\begin{aligned} g(x+\lambda z) &= f(x) + \lambda r \\ &= \lambda(f(x/\lambda) + r) \\ &\leq \lambda(f(x/\lambda) + p((x/\lambda) + z) - f(x/\lambda)) = p(x+\lambda z) \end{aligned}$$

where we have noted $r \leq p((x/\lambda)+z) - f(x/r)$ by its definition. Further we have, when $\lambda > 0$

$$\begin{aligned} g(x-\lambda z) &= f(x) - \lambda r \\ &= \lambda(f(x/\lambda) - r) \\ &\leq \lambda(f(x/\lambda) + p((x/\lambda) - z) - f(x/\lambda)) = p(x-\lambda z) \end{aligned}$$

This proves the second property and completes the proof. □

THEOREM 2.2. ***Hahn-Banach: Real Case.*** *Suppose the E is a vector space over $\mathbb{R}$ and F is a proper subspace of E. Suppose that $p : E \to \mathbb{R}$ is a support function defined on all of E and $f \in F^*$ and*

$$f(x) \leq p(x) \text{ for all } x \in F.$$

Then there is $g \in E^*$ so that

i) *$g(x) = f(x)$, for all $x \in F$.*
ii) *$g(x) \leq p(x)$, for all $x \in E$.*

The proof of the Theorem relies on Zorn's Lemma. We recall therefore the statement of Zorn's Lemma. Recall therefore that a *partial order* on a set S is a relation $\leq$ on S with the following properties

Ord 1 $x \leq x$ for all $x \in S$
Ord 2 If $x \leq y$ and $y \leq z$ then $x \leq z$
Ord 3 If $x \leq y$ and $ylex$ then $x = y$

(following S. Lang's labeling). A subset S' of S is *totally ordered* if whenever $x, y \in S'$ then either $x \leq y$ or $y \leq x$. An *upper bound* for a subset $T \subseteq S$ is an element $b \in S$ so that $x \leq b$ for all $x \in T$. We say that S is *inductively ordered* if every totally ordered subset has an upper bound.

Examples The $\mathbb{Z}$ of integers is totally ordered by the usual ordering. Also $\mathbb{R}$ is totally ordered by its usual ordering. The power set of a set is partially ordered by inclusion "$\subseteq$". The set of subgroups of a group form a partially ordered set with the relation of inclusion.

Zorn's Lemma *If S set is a nonempty inductively ordered set then it has a maximal element*

Of course a *maximal element* of a partially ordered set S is an element $x_0 \in S$ so that $x \leq x_0$ for all $x \in S$.

Zorn's Lemma is known to be equivalent to the Axiom of Choice. (Reference?) Here we treat it as an axiom.

PROOF. **(Hahn-Banach Theorem)** Let S be the set of all pairs (G, u) where G is a subspace of E and indeed $F \subseteq G \subseteq E$ and $u \in G^*$ and $u(x) \leq p(x)$ for all $x \in G$ and moreover $u|_F = f$. We define an ordering on S by $(G_1, u_1) \leq (G_2, u_2)$ if G_1 is a subspace of G_2 and $u_2|_{G_1} = u_1$. Suppose that $S' \subseteq S$ is a totaly ordered subset of S so that whenever $(G_1, u_1), (G_2, u_2) \in S'$ implies either $(G_1, u_1) \leq (G_2, u_2)$ or $(G_2, u_2) \leq (G_1, u_1)$. Define

$$G_0 = \cup_{\{(G,u) \in S'\}} G$$

Then G_0 is a subspace of E because if $x, y \in G_0$ then $x \in G_1$ and $y \in G_2$ for some $(G_1, u_1), (G_2, u_2) \in S'$. Since $(G_1, u_1) \leq (G_2, u_2)$ or $(G_2, u_2) \leq (G_1, u_1)$ both x, y belong to either G_1 or G_2 and therefore so does $x + y$ or αx where $\alpha \in \mathbb{K}$. This says $x + y \in G_0$ and $\alpha x \in G_0$ and so G_0 is a subspace of E. Define $u_0 \in G_0^*$ by $u_0(x) = u_1(x)$ if $x \in G_1$ where $(G_1, u_1) \in S'$. Then $u_0(x)$ is well defined because if $x \in G_2$ also then $(G_1, u_1) \leq (G_2, u_2)$ or $(G_2, u_2) \leq (G_1, u_1)$ and, in either case, $u_2(x) = u_1(x)$. Furthermore u_0 is linear because if $x, y \in G_0$ then $x, y \in G_1$ for some (G_1, u_1) in S' and $u_0(\alpha x + y) = u_1(\alpha x + y) = \alpha u_1(x) + u_1(y) = \alpha u_0(x) + u_0(y)$ for any scalar α. Also $u_0|_F = f$ because $u_1|_F = f$ for all $(G_1, u_1) \in S'$. Finally $u_0(x) \leq p(x)$ for any $x \in G_0$ because $x \in G_1$ for some $(G_1, u_1) \in S'$ and $u_0(x) = u_1(x) \leq p(x)$. Therefore $(G_0, u_0) \in S$ and it is an upper bound for S'. Therefore S is inductively ordered and Zorn's Lemma applies and we find that S has a maximal element, (F_0, f_0) say.

We now claim that $F_0 = E$. For suppose that F_0 were a proper subspace of E so that there is $z \in E - F_0$. Then the preceding theorem applies and we can conclude there is an extension of f_0, g say, to the smallest linear space G that contains F_0 and z so that $g|_{F_0} = f_0$ so that $g(x) \leq p(x)$ for all $x \in G$. However this says that $(G, g) \in S$ and this contradicts the maximality of (F_0, f_0). Therefore $F_0 = E$ and $f_0 \in E^*$ and $f_0 = g$ has the properties stated in the Theorem. $\square$

The Hahn-Banach theorem also holds on complex vector spaces. We shall derive it as a consequence of the theorem for real vector spaces established above. The one difference is that the support function p must be replaced with a semi-norm.

We shall therefore suppose that E is a vector space over $\mathbb{C}$ and that p is a seminorm on E and F is a proper subspace of E and $f \in F^*$

so that $|f(x)| \leq p(x)$ for all $x \in F$. We may regard E as a real vector space by simply restricting the function of scalar multiplication to real scalars. (Scalar multiplication, $(\lambda, v) \mapsto \lambda v$ defined on $\mathbb{C} \times E$ is restricted to $\mathbb{R} \times E$.) To distinguish the real vector space we shall write $E_{\mathbb{R}}$ for E regarded as a real vector space.

Remark. Every complex vector space can be made into a real vector space by simply restricting of scalar multiplication. Observe that the mapping $u(x) = ix$ where $i = \sqrt{-1}$ is an isomorphism with the property that $u^2(x) = -x$. Suppose now that $E_{\mathbb{R}}$ is a real vector space with the property that there is an isomorphism $u : E_{\mathbb{R}} \to E_{\mathbb{R}}$ with the property that $u^2(x) = -x$. For example if $E_{\mathbb{R}} = F \times F$ for some real vector space F then $u(x, y) = (-y, x)$ is such a mapping. We define a mapping of $\mathbb{C} \times E_{\mathbb{R}}$ to $E_{\mathbb{R}}$ by

$$(\alpha + \beta i)x = \alpha x + \beta u(x)$$

This mapping defines complex scalar multiplication on $E_{\mathbb{R}}$ making $E_{\mathbb{R}}$ into a complex vector space. This can be checked but most of the properties are obvious with the possible exception of $(\lambda\mu)x = \lambda(\mu x)$ for all $\lambda = \alpha_1 + i\beta_1$, $\mu = \alpha_2 + i\beta_2$ in $\mathbb{C}$ and $x \in E_{\mathbb{R}}$:

$$\begin{aligned} \lambda(\mu x) &= (\alpha_1 + i\beta_1)(\alpha_2 x + \beta_2 u(x)) \\ &= (\alpha_1\alpha_2 - \beta_1\beta_2)x + (\alpha_1\beta_2 + \beta_1\alpha_2)u(x) = (\lambda\mu)x \end{aligned}$$

Now consider $f \in F^*$. As a complex valued function $f = f_{\Re} + i f_{\Im}$ where $f_{\Re}$ and $f_{\Im}$ are linear on the real space $E_{\mathbb{R}}$. Moreover $f(ix) = if(x)$ implies $f_{\Re}(ix) + i f_{\Im}(ix) = -f_{\Im}(x) + i f_{\Re}(x)$ and this shows that $f_{\Re}(x) = f_{\Im}(ix)$. Conversely, if $f_{\Re}$ is linear on the real space $E_{\mathbb{R}}$ then $f(x) = f_{\Re}(x) - i f_{\Re}(ix)$ is complex linear because $f((\alpha + i\beta)x) = (\alpha + i\beta)f(x)$ by a straightforward computation. Also $f_{\Re}(x) \leq |f_{\Re}(x)| \leq |f(x)| \leq p(x)$

Applying the Hahn Banach Theorem to $f_{\Re}$ there is an extension $g_1 \in E_{\mathbb{R}}^*$ so that $g_1(x) \leq p(x)$. (As a semi-norm on E, p is also a support function on $E_{\mathbb{R}}$.) We define

$$g(x) = g_1(x) - i g_1(ix)$$

so that g is linear as a mapping on the complex vector space E to $\mathbb{C}$ and it extends f because g_1 extends $f_{\Re}$ by the preceding discussion. Furthermore, for some θ,

$$|g(x)| = e^{i\theta} g(x) = g(e^{i\theta}x) = g_1(e^{i\theta}x) \leq p(e^{i\theta}x) = p(x).$$

because $g(e^{i\theta}x)$ must be positive. We have established the following.

Theorem 2.3. ***Hahn-Banach: Complex Case.*** *Suppose the E is a vector space over $\mathbb{C}$ and F is a proper subspace of E. Suppose* *that* $p : E \to \mathbb{R}$ *is a semi-norm defined on all of* E *and* $f \in F^*$ *and*

$$|f(x)| \leq p(x) \text{ for all } x \in F.$$

Then there is $g \in E^*$ *so that*

i) $g(x) = f(x)$, *for all* $x \in F$.
ii) $|g(x)| \leq p(x)$, *for all* $x \in E$