ANNEX 1

APPENDIX 3

PROOF OF THE HAHN-BANACH THEOREM

In order to prove the Hahn-Banach theorem, we need the following preliminary results.

A set E is called *partially ordered* if there exists a binary relation, $\leq$, defined for certain pairs $(x,y) \in E \times E$ such that

(*i*) $x \leq x$,

(*ii*) if $x \leq y$ and $y \leq z$, then $x \leq z$.

Let F be a subset of E. An element $x \in E$ is an *upper bound* for F (with respect to $\leq$) if $y \leq x$ for every $y \in F$.

F is *totally ordered* if for every y and z in F, either $y \leq z$ or $z \leq y$.

An element $m \in E$ is said to be *maximal* if $m \leq x$ implies $m = x$.

For example, if E is the set of all subsets of a set S, let $A \leq B$ mean $A \subset B$. Then E is partially ordered with respect to $\leq$. If F is a subset of E, then $\cup_{A \in F} A$ is an upper bound for F. If F consists of sets $A_1 \subseteq A_2 \subseteq \ldots$, then F is totally ordered. The set S is the only maximal element in E.

ZORN'S LEMMA. *Let E be a non-empty partially ordered set. If every totally ordered subset of E has an upper bound in E, then E contains a maximal element.*

The following result is a more general version of the Hahn-Banach Theorem IX.5.1.

THEOREM. *Suppose* X *is a vector space over the real or complex numbers. Let* p *be a real-valued function defined on* X *such that for all* x, y *in* X *and numbers* α,

(a) $p(x+y) \leq p(x) + p(y)$

(b) $p(\alpha x) = |\alpha| p(x)$.

Suppose f *is a linear functional defined on a subspace* M *of* X *and*

$$|f(z)| \leq p(z) \quad \text{for all} \quad z \in M.$$

Then f *can be extended to a linear functional* F *defined on all of* X *such that*

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

PROOF. Let $\mathcal{E}$ be the set of all linear functionals g such that the domain $D(g)$ of g is contained in X, $g = f$ on M, and $|g(x)| \leq p(x)$, $x \in D(g)$. Obviously, f is in $\mathcal{E}$. We partially order $\mathcal{E}$ by letting $g \leq h$ mean that h is an extension of g. The idea of the proof is to use Zorn's lemma to show that $\mathcal{E}$ has a maximal element and that this element is defined on all of X.

Let $\mathcal{J}$ be a totally ordered subset of $\mathcal{E}$. Define the linear functional H by

$$D(H) = \bigcup_{g \in \mathcal{J}} D(g)$$

$$H(x) = g(x) \quad \text{if} \quad x \in D(g).$$

Since $\mathcal{J}$ is totally ordered, it follows that H is in $\mathcal{E}$ and that H is an upper bound of $\mathcal{J}$. Hence $\mathcal{E}$ has a maximal element F by Zorn's lemma. The theorem is proved once we show that $D(F) = X$.

Suppose there exists an $x \in X$ which is not in $D(F)$. Let $X_1 = sp\{x\} \oplus D(F)$. We shall show that there exists a $G \in \mathcal{E}$ which is an extension of F to X_1, thereby contradicting the maximality of F. It is clear that we should define G on X_1 by

(1) $\quad G(\alpha x+z) = \alpha G(x) + F(z), \quad z \in D(F),$

with $G(x)$ chosen so that

(2) $\quad |G(\alpha x+z)| \leq p(\alpha x+z), \quad z \in D(F).$

In order to determine $G(x)$, let us first assume that X is a vector space over the reals. If (2) is to hold, then in particular,

(3) $\quad G(x+z) \leq p(x+z), \quad z \in D(F)$

and

(4) $\quad G(-x-y) \leq p(-x-y) = p(x+y), \quad y \in D(F).$

Thus, from (1), (3) and (4) we want

(5) $\quad G(x) \leq p(x+z) - F(z), \quad z \in D(F).$

and

(6) $\quad G(x) \geq -p(x+y) - F(y), \quad y \in D(F).$

It is easy to check that the right side of (5) is greater than or equal to the right side of (6) for all z and y in $D(F)$. Thus, if we define

$$G(x) = \inf\{p(x+z) - F(z) : z \in D(F)\},$$

it follows that (3) and (4) hold. But then

(7) $\quad G(\alpha x+z) \leq p(\alpha x+z), \quad z \in D(F).$

To see this, suppose $\alpha > 0$. Then (3) implies

$$G(\alpha x+z) = \alpha G(x + z/\alpha) \leq \alpha p(x + z/\alpha) = p(\alpha x+z).$$

If $\alpha < 0$, then (4) implies

$$G(\alpha x+z) = -\alpha G(-x - z/\alpha) \leq -\alpha p(-x - z/\alpha) = p(\alpha x+z).$$

If $\alpha = 0$, then

$$G(z) = F(z) \leq p(z).$$

Finally, (7) implies (2) since

$$-G(\alpha x+z) = G(-\alpha x-z) \leq p(-\alpha x-z) = p(\alpha x+z).$$

To summarize, we have constructed a $G \in \mathcal{E}$ such that $F \leq G$ but $F \neq G$. This contradicts the maximality of F. Hence $D(F) = X$ and the theorem is proved for X a vector space over the reals.