The Hahn-Banach Theorem **BJG October 2011**

Conspicuous by its absence from this course (Cambridge Mathematical Tripos Part II, Linear Analysis) is the Hahn-Banach theorem. A simple version of it is as follows.

Theorem 1 (Hahn-Banach)**.** *Let $V, \tilde{V}$ be normed spaces with $V \subseteq \tilde{V}$. Let $\phi : V \to \mathbb{R}$ be a bounded linear functional. Then there is a bounded linear functional $\tilde{\phi} : \tilde{V} \to \mathbb{R}$ which extends ϕ in the sense that $\tilde{\phi}|_V = \phi$, and for which $\|\tilde{\phi}\| = \|\phi\|$.*

Proof. Roughly speaking, the idea is to extend ϕ to $\tilde{\phi}$ "one dimension at a time". Suppose, then, that $\tilde{V} = V + \langle w \rangle$, where $w \notin V$. By rescaling we may assume, without loss of generality, that $\|\phi\| = 1$. We are forced to define

$$\tilde{\phi}(v + w) = \phi(v) + t\lambda$$

for all $v \in V$ and $t \in \mathbb{R}$, where λ must not depend on v or on t. A map $\tilde{\phi}$ defined in this way will always be a linear functional, but our task is to show that by judicious choice of λ we may ensure that $\|\tilde{\phi}\| \leqslant 1$. For this we require that

$$|\phi(v) + t\lambda| \leqslant \|v + tw\|$$

for all $v \in V$ and $t \in \mathbb{R}$. By replacing v by v/t and using linearity of ϕ, it suffices to establish this in the case $t = 1$; that is to say, we must show that there is $\lambda \in \mathbb{R}$ such that

$$|\phi(v) + \lambda| \leqslant \|v + w\|$$

for all $v \in V$. Equivalently,

$$-\|v + w\| - \phi(v) \leqslant \lambda \leqslant \|v + w\| - \phi(v).$$

There will be such a λ if, and only if,

$$-\|v + w\| - \phi(v) \leqslant \|v' + w\| - \phi(v')$$

for all $v, v' \in V$. Indeed, we could then take any $\lambda \in [m, M]$ with

$$m := \sup_{v \in V} -\|v + w\| - \phi(v), \qquad M := \inf_{v' \in V} \|v' + w\| - \phi(v').$$

Rearranging, the inequality we are required to prove is

$$\phi(v') - \phi(v) \leqslant \|v' + w\| + \|v + w\|$$

for all $v, v' \in V$. However the left-hand side is $\phi(v' - v)$ which, since $\|\phi\| = 1$, has magnitude at most $\|v' - v\|$. The result is now a consequence of the triangle inequality.

We have proved the Hahn-Banach theorem when $\tilde{V}$ is obtained from V by the addition of one vector. This is already enough to prove the whole theorem when $\tilde{V}$ is finite-dimensional (by incrementing the dimension of V one step at a time). Essentially the same argument works in the infinite-dimensional case, too, although Zorn's lemma is needed to make this rigorous.

Consider the set of all extensions of ϕ, that is to say pairs (V', ϕ') where $V \subseteq V' \subseteq \tilde{V}$, $\phi'|_V = \phi$, and $\|\phi'\| \leqslant \|\phi\|$. There is an obvious partial order on this set: namely, say that $(V_1, \phi_1) \preceq (V_2, \phi_2)$ if and only if $V_1 \subseteq V_2$ and $\phi_2|_{V_1} = \phi_1$. Every chain in this partial order has an upper bound. Indeed if $(V_i, \phi_i)_{i \in I}$ is a chain, then an upper bound for it is (V', ϕ'), where $V' = \bigcup_{i \in I} V_i$ and ϕ' equals ϕ_i on V_i, for all i. By Zorn's lemma, there is a maximal element (V_0, ϕ_0). However by the special case of the theorem proved above, we could extend ϕ_0 to $V_0 + \langle w \rangle$ for any $w \notin V_0$. The only possible conclusion is that there *is* no $w \notin V_0$, or in other words $V_0 = \tilde{V}$ and ϕ_0 is defined on all of $\tilde{V}$. $\square$

Remark. Zorn's lemma is equivalent to the axiom of choice, and so we have used the axiom of choice in proving Hahn-Banach. It is known that Hahn-Banach is strictly weaker than the axiom of choice, but cannot be proven in ZF.

Let us derive some consequences of the theorem. The main point is that, without it, we are essentially powerless to construct a good supply of bounded linear functionals on a typical normed space X. With it, however, we immediately see that X^* is quite rich; indeed for any $x \in X$ there is some $\phi \in X^*$ such that $\phi(x) \neq 0$. More specifically, there is some $\phi \in X^*$ with $\|\phi\| = 1$ such that $\phi(x) = \|x\|$. To see these facts, simply take V to be the subspace spanned by $\langle x \rangle$ and $\tilde{V} := X$, and extend the linear functional $\phi_0 : V \to \mathbb{R}$ defined by $\phi_0(tx) = t\|x\|$.

One may think of this geometrically in terms of convex bodies admitting supporting hyperplanes. Consider the unit ball $B := \{x \in X : \|x\| \leqslant 1\}$ (which is the most general form of a convex set) and let $x_0 \in B$ have norm 1. As just remarked, there is a linear functional $\phi : X \to \mathbb{R}$ with $\phi(x_0) = 1$ and $\|\phi\| \leqslant 1$. Consider the hyperplane $H := \{x \in X : \phi(x) = 1\}$. Then H meets B at x_0 (and possibly at other points). However if $x \in B$ then $\phi(x) \leqslant \|\phi\|\|x\| \leqslant 1$, and so all of B lies in the half-space $\{x \in X : \phi(x) \leqslant 1\}$. H is called a supporting hyperplane for B.

The following interesting fact is little more than a rephrasing of the above.

Theorem 2. *Let X be a normed space. Then the natural map from X to X^{**} is an isometry.*

Proof. The natural map in question associates to $x \in X$ the functional $\hat{x}$ on X^* defined by $\hat{x}(\phi) := \phi(x)$. It is easy to see that $\|\hat{x}\| \leqslant \|x\|$. To get an inequality in the other direction, choose ϕ as described above. Then $|\hat{x}(\phi)| = |\phi(x)| = \|x\| = \|x\|\|\phi\|$, and so indeed $\|\hat{x}\| \geqslant \|x\|$. $\square$

A slightly more complicated observation in the same vein is the following.

Theorem 3. *Let X, Y be normed spaces, and suppose that $T : X \to Y$ is a bounded linear map. Let $T^* : Y^* \to X^*$ be its dual. Then $\|T^*\| = \|T\|$.*

Proof. We remark that it is very easy to see that $\|T^*\| \leqslant \|T\|$; indeed we already remarked on this in the main part of the course. Let $\varepsilon > 0$ be arbitrary. By definition of $\|T\|$, there is some $x \in X$, $x \neq 0$, with $\|Tx\| \geqslant (\|T\| - \varepsilon)\|x\|$. Using the remark above, choose a linear functional $\phi \in X^*$ with $\|\phi\| = 1$ and $\phi(Tx) = \|Tx\|$. Then

$$T^*\phi(x) = \phi(Tx) = \|Tx\| \geqslant (\|T\| - \varepsilon)\|x\|,$$

which certainly means that

$$\|T^*\phi\| \geqslant \|T\| - \varepsilon = (\|T\| - \varepsilon)\|\phi\|.$$

It follows that

$$\|T^*\| \geqslant \|T\| - \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, the result follows. □

Here is another fact we were unable to establish in the course proper.

Theorem 4. *The dual of ℓ^∞ is strictly bigger than ℓ^1.*

Proof. Certainly the dual $(\ell^\infty)^*$ contains ℓ^1, since if $(b_i)_{i \in \mathbb{N}} \in \ell^1$ then the map $(a_i)_{i \in \mathbb{N}} \mapsto \sum_i a_i b_i$ is a bounded linear functional. Note that any functional of this type is determined by its values on ℓ_0^∞, the closed subspace of ℓ^∞ consisting of sequences which tend to zero. This is obviously a *proper* subspace of ℓ^∞, and so the quotient space $\ell^\infty/\ell_0^\infty$ is a nontrivial normed space. By Hahn-Banach we may find a nontrivial bounded linear functional ψ on it. This pulls back under the quotient map $\pi : \ell^\infty \to \ell^\infty/\ell_0^\infty$ to give a nontrivial functional $\phi \in (\ell^\infty)^*$ defined by $\phi(x) := \psi(\pi(x))$. Since ϕ is trivial on ℓ_0^∞, it does not come from ℓ^1. □

The applications we have given so far are perhaps not very "surprising". The next one is rather more so.

Theorem 5 (Finitely additive measure on $\mathbb{Z}$)**.** *Write $\mathcal{P}(\mathbb{Z})$ for the set of all subsets of $\mathbb{Z}$. Then there is a "measure" $\mu : \mathcal{P}(\mathbb{Z}) \to [0, 1]$ which is normalised so that $\mu(\mathbb{Z}) = 1$, is shift-invariant in the sense that $\mu(A + 1) = \mu(A)$, and is finitely-additive in the sense that $\mu(A_1 \cup \cdots \cup A_k) = \mu(A_1) + \mu(A_2) + \cdots + \mu(A_k)$ whenever $A_1, \ldots, A_k$ are disjoint.*

Proof. For the purposes of this proof write ℓ^∞ for the Banach space of bounded sequences $(x_n)_{n \in \mathbb{Z}}$ indexed by $\mathbb{Z}$. We will in fact construct a linear functional $\phi \in (\ell^\infty)^*$ which is *shift-invariant* in the sense that $\phi((x_n)_{n \in \mathbb{Z}}) = \phi((x_{n+1})_{n \in \mathbb{Z}})$, *positive* in the sense that $\phi((x_n)_{n \in \mathbb{Z}}) \geqslant 0$ whenever $x_n \geqslant 0$ for all n, and normalised so that $\phi(\mathbf{1}) = 1$, where $\mathbf{1}$ is the constant sequence of 1s.

Clearly such a ϕ gives rise to a finitely additive measure on $\mathbb{Z}$ by defining $\mu(A) := \phi(x_A)$, where $(x_A)_n = 1$ if $n \in A$ and 0 otherwise.

Let V be the subspace of ℓ^∞ spanned by the constant sequence $\mathbf{1}$ and the space V_0 of sequences of the form $(x_{n+1} - x_n)_{n\in\mathbb{Z}}$. It is trivial to check that $\mathbf{1}$ is not in V_0, so we may unambiguously define

$$\phi_0((x_{n+1} - x_n + c)_{n\in\mathbb{Z}}) := c$$

on V. For any $\varepsilon > 0$ there must be some n such that $x_{n+1} - x_n \geqslant -\varepsilon$, and so if $c > 0$ we certainly have $\|(x_{n+1} - x_n - c)_{n\in\mathbb{Z}}\| = \sup_n |x_{n+1} - x_n + c| \geqslant |c| - \varepsilon$. Since ε was arbitrary, we actually have $\|(x_{n+1} - x_n + c)_{n\in\mathbb{Z}}\| \geqslant |c|$. The same conclusion holds if $c \leqslant 0$, and therefore $\|\phi_0\| \leqslant 1$. By the Hahn-Banach theorem there is an extension of ϕ_0 to a linear functional ϕ on all of ℓ^∞ such that $\|\phi\| \leqslant 1$. This is obviously normalised so that $\phi(\mathbf{1}) = 1$, and ϕ is pretty clearly shift-invariant since $(x_n)_{n\in\mathbb{Z}}$ and the shifted sequence $(x_{n+1})_{n\in\mathbb{Z}}$ differ by an element of V_0, on which ϕ_0 is defined and equal to zero.

It remains to confirm that ϕ is positive, and for this we may suppose without loss of generality that $\|x\| = 1$, so that $0 \leqslant x_n \leqslant 1$ for all n. Then $\|\mathbf{1} - x\| \leqslant 1$, and so

$$1 - \phi(x) = \phi(\mathbf{1} - x) \leqslant \|\mathbf{1} - x\| \leqslant 1,$$

which obviously means that $\phi(x) \geqslant 0$ as required. $\square$