The Hahn–Banach Theorem

The Hahn–Banach theorem has many forms and many important consequences. Here are a few relatively simple and fun ones.

Theorem 1. *Let V be a real vector space. Let $N\colon V\to\mathbb{R}$ be sublinear, in the sense that*

$$N(ax+by)\le aN(x)+bN(y),\ \text{for any } x,y\in V,\ a,b\ge 0.$$

Let W be a subspace of V and $\ell\colon W\to\mathbb{R}$ a linear function such that

$$\ell(x)\le N(x),\ \text{for any } x\in W.$$

Then ℓ can be extended to $L\colon V\to\mathbb{R}$ in such a way that

$$L(x)\le N(x),\ \text{for any } x\in V.$$

Proof. The main step is showing that if $z\notin W$, then ℓ can be extended to $\operatorname{Span}(W\cup\{z\})$. For this we must show that it is possible to define $L(z)$ in such a way that for all $x\in W$ and $a\in\mathbb{R}$ we have

$$\ell(x)+aL(z)\le N(x+az); \tag{1}$$

once that is done we can define $L(x+az)=\ell(x)+aL(z)$, because the representation of an element in $\operatorname{Span}(W\cup\{z\})$ as $x+az$ is unique.

If $a>0$, then satisfying (1) is the same as satisfying

$$L(z)\le a^{-1}(N(x+az)-\ell(x)).$$

If $a<0$, thenthe requirement is

$$L(z)\ge a^{-1}(N(x+az)-\ell(x)).$$

To check that these requirements can be simulatenously satisfied, we must check that

$$b^{-1}(\ell(y)-N(y-bz))\le a^{-1}(N(x+az)-\ell(x)),$$

for all $x,y\in W$ and $a,b>0$. In other words, we must check that

$$a\ell(y)+b\ell(x)\le aN(y-bz)+bN(x+az).$$

For this, write

$$a\ell(y)+b\ell(x)=\ell(ay+bx)\le N(ay+bx)\le aN(y-bz)+bN(x+az),$$

where for the last step we used (2) with x replaced by $y-bx$ and y replaced by $x+az$.

Now order all extensions by inclusion, where $(W',\ell')\subset(W'',\ell'')$ if $W'\subset W''$ and if the graph of ℓ' is a subset of the graph of ℓ''. Each totally ordered family of extensions is bounded by the union of that family, so by Zorn's Lemma there is a maximal element $(W_{\max},\ell_{\max})$, and since $(W_{\max},\ell_{\max})$ is not a proper subset of any extension we must have $W_{\max}=V$. □

A stronger assumption on N allows us to bound $|L(x)|$, and also to cover the complex case.

Kiril Datchev, February 12, 2025. Questions, comments, and corrections are gratefully received at kdatchev@purdue.edu.

Theorem 2. *Let V be a vector space over $\mathbb{F}$, with $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. Let $N\colon V \to [0,\infty)$ be sublinear, in the sense that*

$$N(ax+by) \le |a|N(x) + |b|N(y), \text{ for any } x,y \in V,\ a,b \in \mathbb{F}. \tag{2}$$

Let W be a subspace of V and $\ell\colon W \to \mathbb{F}$ a linear function such that

$$\operatorname{Re}\ell(x) \le N(x), \text{ for any } x \in W.$$

Then ℓ can be extended to $L\colon V \to \mathbb{R}$ in such a way that

$$|L(x)| \le N(x), \text{ for any } x \in V.$$

Proof. If $\mathbb{F} = \mathbb{R}$, then $\operatorname{Re}\ell(x) = \ell(x)$ and Theorem 1 yields an extension with $L(x) \le N(x)$ for all x. The remaining conclusion is $L(x) \ge -N(x)$, and this follows from $L(-x) \le N(-x) \le N(x)$.

If $\mathbb{F} = \mathbb{C}$, then define a real-linear functional on W by $\ell_{\mathbb{R}}(x) = \operatorname{Re}\ell(x)$. As above, there is a real-linear extension $L_{\mathbb{R}}\colon V \to \mathbb{R}$ obeying $L_{\mathbb{R}}(x) \le N(x)$. Let

$$L(x) = L_{\mathbb{R}}(x) - iL_{\mathbb{R}}(ix).$$

Then $L|_W = \ell$ because $z = \operatorname{Re} z - i\operatorname{Re}(iz)$. Next, L is complex linear because $L(ix) = iL(x)$. Finally, if $L(x) \neq 0$ then let $\lambda = \frac{|L(x)|}{L(x)}$ and write

$$|L(x)| = L(\lambda x) = L_{\mathbb{R}}(\lambda x) \le N(\lambda x) \le N(x),$$

where for the second equality we used $\operatorname{Im} L(\lambda x) = 0 \Longrightarrow L_{\mathbb{R}}(i\lambda x) = 0$. ☐

Theorem 3 (Using linear functionals to measure distance)**.** *Let V be a real of complex normed vector space. Let $Z \subset V$ be a subspace, and let $v \in V$. There is $\ell \in V^*$ such that $\ell(v) = d(v,Z)$ and $|\ell(x)| \le d(x,Z)$ for all $x \in V$.*

Proof. Let W be the span of v, and define $N\colon V \to \mathbb{R}$ and $\ell \in W^*$ by

$$N(x) = d(x,Z), \qquad \ell(cv) = cN(v).$$

By Theorem 2, and using $N(x) \le \|x\|$, it is enough to check (2). If $a \neq 0$, then

$$N(ax) = \inf_{z\in Z}\|ax - z\| = \inf_{z\in Z}\|ax - az\| = |a|\inf_{z\in Z}\|x - z\| = |a|N(x).$$

Meanwhile,

$$N(x+y) = \inf_{z\in Z}\|x+y-z\| = \inf_{z_1,z_2\in Z}\|x+y-z_1-z_2\| \le \inf_{z_1\in Z}\|x-z_1\| + \inf_{z_2\in Z}\|y-z_2\| = N(x)+N(y).$$

☐

Exercise 1. Let V be a complex normed vector space, and let $u\colon V \to \mathbb{R}$ be a continuous real-linear functional. Prove that if u attains a maximum on the unit sphere at some x_0, then $u(ix_0) = 0$.

Theorem 4 (A closed subspace is an intersection of hyperplanes)**.** *Let V be a normed vector space, let $Z \subset V$ be a subspace and let $v \in V$. The following are equivalent:*

(1) $v \in \overline{Z}$.
(2) *If $\ell \in V^*$ and $\ell|_Z = 0$, then $\ell(v) = 0$.*

In other words, $\overline{Z} = \bigcap_{\ell \in Z^\perp} \operatorname{Ker}(\ell)$*, where* $Z^\perp = \{\ell \in V^* \colon \ell|_Z = 0\}$.

Proof. (1) $\Longrightarrow$ (2): $\ell|_{\overline{Z}} = 0$ by continuity.

(2) $\Longrightarrow$ (1): If $v \notin \overline{Z}$, then $d(v, Z) > 0$, so Theorem 3 yields ℓ with the desired properties. □

Theorem 5 (Runge approximation)**.** *Let $K \subset \mathbb{C}$ be a compact set with connected complement. If f is analytic in a neighborhood of K, then f can be uniformly approximated by polynomials in z.*

Recall that the Stone–Weierstrass Theorem cannot be applied here, because it requires the approximating family to be closed under complex conjugation. The standard counterexample, which also shows the relevance of requiring connectedness of $\mathbb{C} \setminus K$, is that $1/z$ cannot be uniformly approximated on the unit circle by polynomials $P(z)$ because $\int_0^{2\pi} P(e^{i\theta})e^{i\theta} d\theta = 0$ for every polynomial P.

Exercise 2. Let $K \subset \mathbb{C}$ be the unit circle, and let N be an integer. Prove that there is $\ell \in C(K)^*$ of norm 1 such that $\ell(z^N) = 1$ and $\ell(z^n) = 0$ for every integer $n \neq N$. Do this in two ways, once by invoking Theorem 3 and once by defining the functional via an explicit formula and checking the relevant properties directly.

Proof of Theorem 5. Let $V = C(K)$, let $Z \subset C(K)$ be the space of polynomials in z, and let ℓ be an element of V^* that vanishes on Z. By Theorem 4, it is enough to check that $\ell(f) = 0$. By the Riesz representation theorem (see Theorem 6.19 of [Rud]) there is a complex Borel measure μ_ℓ on K such that $\ell(g) = \int_K g\, d\mu_\ell$ for every $g \in C(K)$.

Since a complex measure is a linear combination of ordinary measures, the problem thus reduces to showing that for any Borel measure μ on K, we have

$$\int_K z^n\, d\mu(z) = 0, \text{ for } n = 0, 1, 2, \ldots \qquad \Longrightarrow \qquad \int_K f(z)\, d\mu(z) = 0.$$

By the Cauchy integral formula,

$$\int_K f(z)\, d\mu(z) = \int_K \int_\Gamma \frac{f(w)dw}{w - z} d\mu(z),$$

where Γ is any contour enclosing K, contained in the set where f is analytic.[1] By switching the order of integration, we see that it is enough to show that the function

$$g(w) = \int_K \frac{d\mu(z)}{w - z}$$

vanishes on $\mathbb{C} \setminus K$. But if w is sufficiently large, specifically $|w| > \max\{|z| \colon z \in K\}$, then

$$g(w) = \int_K \sum_{n=0}^{\infty} \frac{1}{w} \Big(\frac{z}{w}\Big)^n d\mu(z) = \sum_{n=0}^{\infty} w^{-n-1} \int_K z^n\, d\mu(z) = 0.$$

On the other hand, g is analytic (check complex differentiability directly from the definition), and $\mathbb{C} \setminus K$ is connected, so g must vanish identically on $\mathbb{C} \setminus K$ (because the zeroes of a nonconstant analytic function are isolated; see Theorem 10.18 of [Rud]). □

[1] To construct such a Γ, cover K by a finite collection of open rectangles whose closures are contained in the set where f is analytic, and let Γ be the boundary of the union of these rectangles. See Theorem 13.5 of [Rud] for a more detailed and general construction.

The above proof is adapted from Theorem 13.6 of [Rud], which also shows how to handle the case where $\mathbb{C} \setminus K$ is not connected, by replacing the polynomials with suitable rational functions. Example 2.3.10 in [Hör] contains a similar proof: in place of the contour Γ, it uses a smooth cutoff function and a Cauchy integral formula for compactly supported functions. Near that example in [Hör] are several other fun applications of the Hahn–Banach Theorem.

Mergelyan's theorem improves Runge's theorem by weakening its hypotheses. Specifically, instead of requiring that f be analytic near K, it requires only that f be continuous on K and analytic on the interior of K. It is the final theorem of [Rud], and its final step is an application of Runge's theorem.

References

[Hör] Hörmander *Linear Functional Analysis*, 1988.
[Rud] Walter Rudin, *Real and Complex Analysis*, Third Edition, 1987.