Hahn-Banach Theorem and Applications

Let E be a vector space over $\mathbb{R}$. We recall that a *functional* is a function defined on E, or on some subspace of E, *with values in* $\mathbb{R}$.

Theorem 1.1 (Helly, Hahn–Banach analytic form). *Let $p : E \to \mathbb{R}$ be a function satisfying*

$$p(\lambda x) = \lambda p(x) \quad \forall x \in E \quad \text{and} \quad \forall \lambda > 0, \tag{1}$$

$$p(x + y) \leq p(x) + p(y) \quad \forall x, y \in E. \tag{2}$$

Let $G \subset E$ be a linear subspace and let $g : G \to \mathbb{R}$ be a linear functional such that

$$g(x) \leq p(x) \quad \forall x \in G. \tag{3}$$

Under these assumptions, there exists a linear functional f defined on all of E that extends g, i.e., $g(x) = f(x)\ \forall x \in G$, and such that

$$f(x) \leq p(x) \quad \forall x \in E. \tag{4}$$

Proof

Consider the set

$$P = \left\{ h : D(h) \subset E \to \mathbb{R} \;\middle|\; \begin{array}{l} D(h) \text{ is a linear subspace of } E, \\ h \text{ is linear, } G \subset D(h), \\ h \text{ extends } g, \text{ and } h(x) \leq p(x) \quad \forall x \in D(h) \end{array} \right\}.$$

On P we define the order relation

$$(h_1 \leq h_2) \Leftrightarrow (D(h_1) \subset D(h_2) \text{ and } h_2 \text{ extends } h_1).$$

On P we define the order relation

$$(h_1 \leq h_2) \Leftrightarrow (D(h_1) \subset D(h_2) \text{ and } h_2 \text{ extends } h_1) \,.$$

It is clear that P is nonempty, since $g \in P$. We claim that P is *inductive*. Indeed, let $Q \subset P$ be a totally ordered subset; we write Q as $Q = (h_i)_{i \in I}$ and we set

$$D(h) = \bigcup_{i \in I} D(h_i), \quad h(x) = h_i(x) \quad \text{if } x \in D(h_i) \text{ for some } i.$$

It is easy to see that the definition of h makes sense, that $h \in P$, and that h is an upper bound for Q. We may therefore apply Zorn's lemma, and so we have a maximal element f in P. We claim that $D(f) = E$, which completes the proof of Theorem

Suppose, by contradiction, that $D(f) \neq E$. Let $x_0 \notin D(f)$; set $D(h) = D(f) + \mathbb{R}x_0$, and for every $x \in D(f)$, set $h(x + tx_0) = f(x) + t\alpha$ $(t \in \mathbb{R})$, where the constant $\alpha \in \mathbb{R}$ will be chosen in such a way that $h \in P$. We must ensure that

$$f(x) + t\alpha \leq p(x + tx_0) \quad \forall x \in D(f) \quad \text{and} \quad \forall t \in \mathbb{R}.$$

In view of (1) it suffices to check that

$$\begin{cases} f(x)+\alpha \le p(x+x_0) & \forall x \in D(f), \\ f(x)-\alpha \le p(x-x_0) & \forall x \in D(f). \end{cases}$$

In other words, we must find some α satisfying

$$\sup_{y\in D(f)} \{f(y)-p(y-x_0)\} \le \alpha \le \inf_{x\in D(f)} \{p(x+x_0)-f(x)\}.$$

Such an α exists, since

$$f(y)-p(y-x_0) \le p(x+x_0)-f(x) \quad \forall x \in D(f), \quad \forall y \in D(f);$$

indeed, it follows from (2) that

$$f(x)+f(y) \le p(x+y) \le p(x+x_0)+p(y-x_0).$$

We conclude that $f \le h$; but this is impossible, since f is maximal and $h \ne f$.

Notation. We denote by $E^\star$ the *dual space* of E, that is, the space of all *continuous linear functionals on* E; the (dual) *norm on* $E^\star$ is defined by

$$\|f\|_{E^\star} = \sup_{\substack{\|x\|\le 1\\ x\in E}} |f(x)| = \sup_{\substack{\|x\|\le 1\\ x\in E}} f(x). \tag{5}$$

When there is no confusion we shall also write $\|f\|$ instead of $\|f\|_{E^\star}$.

Given $f \in E^\star$ and $x \in E$ we shall often write $\langle f, x\rangle$ instead of $f(x)$; we say that $\langle\ ,\ \rangle$ is the *scalar product for the duality* $E^\star, E$.

• **Corollary 1.2.** *Let $G \subset E$ be a linear subspace. If $g : G \to \mathbb{R}$ is a continuous linear functional, then there exists $f \in E^\star$ that extends g and such that*

$$\|f\|_{E^\star} = \sup_{\substack{x \in G \\ \|x\| \le 1}} |g(x)| = \|g\|_{G^\star}.$$

Proof. Use Theorem 1.1 with $p(x) = \|g\|_{G^\star}\|x\|$.

• **Corollary 1.3.** *For every $x_0 \in E$ there exists $f_0 \in E^\star$ such that*

$$\|f_0\| = \|x_0\| \text{ and } \langle f_0, x_0\rangle = \|x_0\|^2.$$

Proof. Use Corollary 1.2 with $G = \mathbb{R}x_0$ and $g(tx_0) = t\|x_0\|^2$, so that $\|g\|_{G^\star} = \|x_0\|$.

• **Corollary 1.4.** *For every $x \in E$ we have*

$$\|x\| = \sup_{\substack{f \in E^\star \\ \|f\| \le 1}} |\langle f, x\rangle| = \max_{\substack{f \in E^\star \\ \|f\| \le 1}} |\langle f, x\rangle|. \tag{6}$$

Proof. We may always assume that $x \neq 0$. It is clear that

$$\sup_{\substack{f \in E^\star \\ \|f\| \le 1}} |\langle f, x\rangle| \le \|x\|.$$

On the other hand, we know from Corollary 1.3 that there is some $f_0 \in E^\star$ such that $\|f_0\| = \|x\|$ and $\langle f_0, x\rangle = \|x\|^2$. Set $f_1 = f_0/\|x\|$, so that $\|f_1\| = 1$ and $\langle f_1, x\rangle = \|x\|$.

Lemma For any nonzero element x of a normed space X there is a functional f such that $\|f\| = 1$ and $f(x) = \|x\|$.

Corollary (X^* separates points) *If $x, y \in X$ with $x \neq y$, then there exists $f \in X^*$ such that $f(x) \neq f(y)$. Consequently, if $x, y \in X$ and $f(x) = f(y)$ for every $f \in X^*$, then $x = y$.*

Proof If $x \neq y$, then by the previous lemma, there exists an f with $\|f\|_{X^*} = 1$ such that $f(x) - f(y) = f(x - y) = \|x - y\| \neq 0$. $\square$

This result shows that understanding the action of elements of X^* on X is enough to 'understand' the whole of X.

In order to consider the complex version of the Hahn-Banach theorem, we first observe that any complex vector space V can be viewed as a real vector space by only allowing scalar multiplication by real numbers. This does not effect the elements of the space itself, but has a significant effect on what it means for a map to be `linear', i.e. the values of α, β allowed in the expression

$$\phi(\alpha x + \beta y) = \alpha\phi(x) + \beta\phi(y), \quad x, y \in V \qquad (*)$$

become restricted to real numbers.

We therefore make the distinction now between `real-linear' maps (allowing only $\alpha, \beta \in \mathbb{R}$ in (*)) and 'complex-linear' maps (for which we can take $\alpha, \beta \in \mathbb{C}$ in (*)).

While the 'real version' of a normed space V has the same elements, its dual space contains more maps, since they are only required to real-linear. We therefore use the notation $V_{\mathbb{R}}^*$ for the collection of all bounded real-linear functionals on V in order to distinguish it from V^* (which contains only complex-linear functionals when V is complex).

Lemma A *Let V be a complex vector space. Given any complex-linear functional $f: V \to \mathbb{C}$ there exists a unique real-linear $\psi: V \to \mathbb{R}$ such that*

$$f(v) = \psi(v) - \mathrm{i}\psi(\mathrm{i}v) \qquad \textit{for all} \qquad v \in V. \qquad (**)$$

Furthermore,

(i) *if $p: V \to \mathbb{R}$ is a seminorm and $|f(v)| \le p(v)$, then $|\psi(v)| \le p(v)$;*
(ii) *if V is a normed space and $f \in V^*$, then $\psi \in V_{\mathbb{R}}^*$ with*

$$\|\psi\|_{V_{\mathbb{R}}^*} = \|f\|_{V^*}.$$

Conversely, if $\psi: V \to \mathbb{R}$ is real-linear and satisfies $|\psi(v)| \le p(v)$, then $f: V \to \mathbb{C}$ defined by (**) is complex-linear and satisfies $|f(v)| \le p(v)$ for all $v \in V$. Moreover, if $\psi \in V_{\mathbb{R}}^*$, then $f \in V^*$ with $\|f\|_{V^*} = \|\psi\|_{V_{\mathbb{R}}^*}$.

Proof If $v \in V$, then we can write

$$f(v) = \psi(v) + \mathrm{i}\phi(v),$$

where $\psi, \phi: V \to \mathbb{R}$ are real-linear. Since

$$\psi(\mathrm{i}v) + \mathrm{i}\phi(\mathrm{i}v) = f(\mathrm{i}v) = \mathrm{i}f(v) = \mathrm{i}\psi(v) - \phi(v),$$

it follows that $\psi(\mathrm{i}v) = -\phi(v)$, which yields (**).

If $|f(v)| \le p(v)$, then this is inherited by ψ, since

$$|f(v)|^2 = |\psi(v)|^2 + |\psi(\mathrm{i}v)|^2 \qquad \Rightarrow \qquad |\psi(v)| \le |f(v)| \le p(v);$$

for bounded linear functionals on a normed space the same argument gives $\|\psi\|_{V_{\mathbb{R}}^*} \le \|f\|_{V^*}$.

To show that $\|\psi\|_{V_{\mathbb{R}}^*} \ge \|f\|_{V^*}$, observe that $|f(x)| = \mathrm{e}^{\mathrm{i}\theta} f(x)$ for some $\theta \in \mathbb{R}$. So

$$|f(x)| = \mathrm{e}^{\mathrm{i}\theta} f(x) = f(\mathrm{e}^{\mathrm{i}\theta} x) = \psi(\mathrm{e}^{\mathrm{i}\theta} x) - \mathrm{i}\psi(\mathrm{i}\mathrm{e}^{\mathrm{i}\theta} x).$$

Since $|f(x)|$ is real, we must have, for all $x \in V$,

$$|f(x)| = \psi(\mathrm{e}^{\mathrm{i}\theta} x) \le |\psi(\mathrm{e}^{\mathrm{i}\theta} x)| \le \|\psi\|_{V_{\mathbb{R}}^*} \|\mathrm{e}^{\mathrm{i}\theta} x\|_V = \|\psi\|_{V_{\mathbb{R}}^*} \|x\|_V,$$

and so $\|f\|_{V^*} \le \|\psi\|_{V_{\mathbb{R}}^*}$.

To prove the converse results (from ψ to f), we first show that, given a real-linear map $\psi : V \to \mathbb{R}$, the map $f : V \to \mathbb{C}$ defined in (**) is complex-linear. It is clear that

$$f(u + v) = f(u) + f(v),$$

since ψ has this property. We need only show that $f(\lambda u) = \lambda f(u)$ for $\lambda \in \mathbb{C}$, and for this it suffices to check that $f(\mathrm{i}u) = \mathrm{i} f(u)$, since f is clearly real-linear because ψ is. We have $f(\mathrm{i}u) = \psi(\mathrm{i}u) - \mathrm{i}\psi(-u) = \mathrm{i}[\psi(u) - \mathrm{i}\psi(\mathrm{i}u)] = \mathrm{i} f(u)$.

Suppose that $|f(x)| = \mathrm{e}^{\mathrm{i}\theta} f(x)$; then $|f(x)| = f(\mathrm{e}^{\mathrm{i}\theta} x) = \psi(\mathrm{e}^{\mathrm{i}\theta} x) - \mathrm{i}\psi(\mathrm{i}\mathrm{e}^{\mathrm{i}\theta} x)$,

and since $|f(x)|$ is real we have

$$|f(x)| = \psi(\mathrm{e}^{\mathrm{i}\theta} x) \le |\psi(\mathrm{e}^{\mathrm{i}\theta} x)| \le p(\mathrm{e}^{\mathrm{i}\theta} x) = |\mathrm{e}^{\mathrm{i}\theta}| p(x) = p(x). \qquad \square$$

Definition If V is a vector space, then a function $p\colon V \to \mathbb{R}$ is

- *sublinear* if for $x, y \in V$

$$p(x+y) \leq p(x) + p(y) \qquad \text{and} \qquad p(\lambda x) = \lambda p(x),\ \lambda \in \mathbb{R},\ \lambda \geq 0;$$

- a *seminorm* if for $x, y \in V$

$$p(x+y) \leq p(x) + p(y) \qquad \text{and} \qquad p(\lambda x) = |\lambda| p(x),\ \lambda \in \mathbb{K}.$$

Theorem (Complex Hahn–Banach Theorem) *Let X be a complex vector space, U a subspace of X, and p a seminorm on X. Suppose that $\phi\colon U \to \mathbb{C}$ is linear and satisfies*

$$|\phi(x)| \leq p(x) \qquad \textit{for all} \quad x \in U.$$

Then there exists a linear map $f\colon X \to \mathbb{C}$ such that $f(x) = \phi(x)$ for all $x \in U$ and

$$|f(x)| \leq p(x) \qquad \textit{for all} \quad x \in X.$$

In particular, if X is a normed space, then any $\phi \in U^*$ can be extended to some $f \in X^*$ with $\|f\|_{X^*} = \|\phi\|_{U^*}$.

Proof By Lemma A there exists a real-linear $\psi : U \to \mathbb{R}$ such that

$$\phi(v) = \psi(v) - i\psi(iv), \qquad (***)$$

and

$$|\psi(w)| \leq p(w)$$

for all $w \in U$.

We can now use the real Hahn–Banach Theorem to extend ψ from U to X to give a real-linear map $\Psi : X \to \mathbb{R}$ that satisfies

$$|\Psi(x)| \leq p(x) \qquad \text{for every} \quad x \in X.$$

Finally, if we define

$$f(u) = \Psi(u) - i\Psi(iu),$$

then $f : X \to \mathbb{C}$ is complex-linear, extends ϕ, and satisfies

$$|f(w)| \leq p(w),$$

using the second half of Lemma A.

If $\phi \in U^*$, then we follow a very similar argument, first writing ϕ as in (***), where now $\psi \in U^*_{\mathbb{R}}$ with $\|\psi\|_{U^*_{\mathbb{R}}} = \|\phi\|_{U^*}$. We then extend ψ to an element $\Psi \in X^*_{\mathbb{R}}$ with $\|\Psi\|_{X^*_{\mathbb{R}}} = \|\psi\|_{U^*_{\mathbb{R}}}$ and use the second part of Lemma A to guarantee that the complex-linear functional f on X defined by setting $f(u) := \Psi(u) - i\Psi(iu)$ satisfies

$$\|f\|_{X^*} = \|\Psi\|_{X^*_R} = \|\psi\|_{U^*} = \|\phi\|_{U^*}. \qquad \square$$

Let E be a normed linear space.

Definition. An affine *hyperplane* is a subset H of E of the form

$$H = \{x \in E\,;\, f(x) = \alpha\},$$

where f is a linear functional that does not vanish identically and $\alpha \in \mathbb{R}$ is a given constant. We write $H = [f = \alpha]$ and say that $f = \alpha$ is the equation of H.

A subset $A \subset E$ is convex if

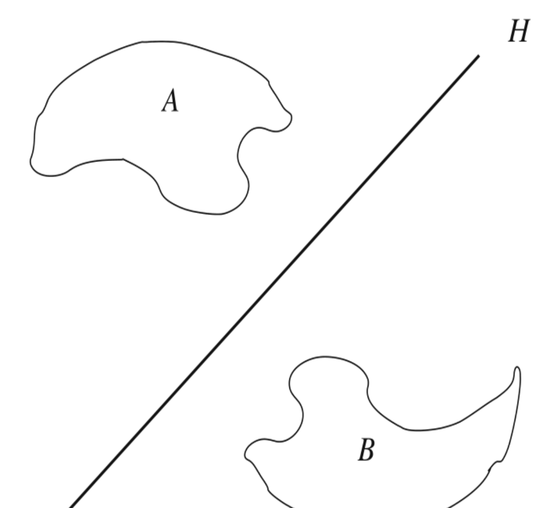
$$\boxed{tx + (1-t)y \in A \quad \forall x, y \in A, \quad \forall t \in [0,1].}$$

Definition. Let A and B be two subsets of E. We say that the hyperplane $H = [f = \alpha]$ *separates* A *and* B if

$$\boxed{f(x) \le \alpha \quad \forall x \in A \quad \text{and} \quad f(x) \ge \alpha \quad \forall x \in B.}$$

We say that H *strictly separates* A and B if there exists some $\varepsilon > 0$ such that

$$\boxed{f(x) \le \alpha - \varepsilon \quad \forall x \in A \text{ and } f(x) \ge \alpha + \varepsilon \quad \forall x \in B.}$$



Geometrically, the separation means that A lies in one of the half-spaces determined by H, and B lies in the other.

Proposition 1.5. *The hyperplane $H = [f = \alpha]$ is closed if and only if f is continuous.*

Proof. It is clear that if f is continuous then H is closed. Conversely, let us assume that H is closed. The complement H^c of H is open and nonempty (since f does not vanish identically). Let $x_0 \in H^c$, so that $f(x_0) \neq \alpha$, for example, $f(x_0) < \alpha$.

Fix $r > 0$ such that $B(x_0, r) \subset H^c$, where

$$B(x_0, r) = \{x \in E\, ;\, \|x - x_0\| < r\}.$$

We have that

$$f(x) < \alpha \quad \forall x \in B(x_0, r). \tag{7}$$

Indeed, suppose by contradiction that $f(x_1) > \alpha$ for some $x_1 \in B(x_0, r)$. The segment

$$\{x_t = (1 - t)x_0 + tx_1\, ;\, t \in [0, 1]\}$$

is contained in $B(x_0, r)$ and thus $f(x_t) \neq \alpha$, $\forall t \in [0, 1]$; on the other hand, $f(x_t) = \alpha$ for some $t \in [0, 1]$, namely $t = \frac{\alpha - f(x_0)}{f(x_1) - f(x_0)}$, a contradiction, and thus (7) is proved.

It follows from (7) that

$$f(x_0 + rz) < \alpha \quad \forall z \in B(0, 1).$$

Consequently, f is continuous and $\|f\| \leq \frac{1}{r}(\alpha - f(x_0))$.

Lemma 1.2. *Let $C \subset E$ be an open convex set with $0 \in C$. For every $x \in E$ set*

$$p(x) = \inf\{\alpha > 0; \alpha^{-1}x \in C\} \tag{8}$$

(p is called the gauge of C or the Minkowski functional of C).

Then p satisfies (1), (2), *and the following properties:*

$$\text{there is a constant } M \text{ such that } 0 \le p(x) \le M\|x\| \quad \forall x \in E, \tag{9}$$

$$C = \{x \in E\,;\, p(x) < 1\}. \tag{10}$$

Proof of (9). Let $r > 0$ be such that $B(0, r) \subset C$; we clearly have

$$p(x) \le \frac{1}{r}\|x\| \quad \forall x \in E.$$

Proof of (10). First, suppose that $x \in C$; since C is open, it follows that $(1+\varepsilon)x \in C$ for $\varepsilon > 0$ small enough and therefore $p(x) \le \frac{1}{1+\varepsilon} < 1$. Conversely, if $p(x) < 1$ there exists $\alpha \in (0, 1)$ such that $\alpha^{-1}x \in C$, and thus $x = \alpha(\alpha^{-1}x) + (1-\alpha)0 \in C$.

Proof of (2). Let $x, y \in E$ and let $\varepsilon > 0$. Using (1) and (10) we obtain that $\frac{x}{p(x)+\varepsilon} \in C$ and $\frac{y}{p(y)+\varepsilon} \in C$. Thus $\frac{tx}{p(x)+\varepsilon} + \frac{(1-t)y}{p(y)+\varepsilon} \in C$ for all $t \in [0, 1]$. Choosing the value $t = \frac{p(x)+\varepsilon}{p(x)+p(y)+2\varepsilon}$, we find that $\frac{x+y}{p(x)+p(y)+2\varepsilon} \in C$. Using (1) and (10) once more, we are led to $p(x + y) < p(x) + p(y) + 2\varepsilon$, $\forall \varepsilon > 0$.

Lemma 1.3. *Let $C \subset E$ be a nonempty open convex set and let $x_0 \in E$ with $x_0 \notin C$. Then there exists $f \in E^\star$ such that $f(x) < f(x_0) \quad \forall x \in C$. In particular, the hyperplane $[f = f(x_0)]$ separates $\{x_0\}$ and C.*

Proof of Lemma 1.3. After a translation we may always assume that $0 \in C$. We may thus introduce the gauge p of C (see Lemma 1.2). Consider the linear subspace $G = \mathbb{R}x_0$ and the linear functional $g : G \to \mathbb{R}$ defined by

$$g(tx_0) = t, \quad t \in \mathbb{R}.$$

It is clear that

$$g(x) \leq p(x) \quad \forall x \in G$$

(consider the two cases $t > 0$ and $t \leq 0$). It follows from Theorem 1.1 that there exists a linear functional f on E that extends g and satisfies

$$f(x) \leq p(x) \quad \forall x \in E.$$

In particular, we have $f(x_0) = 1$ and that f is continuous by (9). We deduce from (10) that $f(x) < 1$ for every $x \in C$.

Theorem 1.6 (Hahn–Banach, first geometric form). *Let $A \subset E$ and $B \subset E$ be two nonempty convex subsets such that $A \cap B = \emptyset$. Assume that one of them is open. Then there exists a closed hyperplane that separates A and B.*

Proof of Theorem 1.6. Set $C = A - B$, so that C is convex (check!), C is open (since $C = \bigcup_{y \in B}(A - y)$), and $0 \notin C$ (because $A \cap B = \emptyset$). By Lemma 1.3 there is some $f \in E^\star$ such that

$$f(z) < 0 \quad \forall z \in C,$$

that is,

$$f(x) < f(y) \quad \forall x \in A, \quad \forall y \in B.$$

Fix a constant α satisfying

$$\sup_{x \in A} f(x) \leq \alpha \leq \inf_{y \in B} f(y).$$

Clearly, the hyperplane $[f = \alpha]$ separates A and B.

Theorem 1.7 (Hahn–Banach, second geometric form). *Let $A \subset E$ and $B \subset E$ be two nonempty convex subsets such that $A \cap B = \emptyset$. Assume that A is closed and B is compact. Then there exists a closed hyperplane that strictly separates A and B.*

Proof. Set $C = A - B$, so that C is convex, closed (check!), and $0 \notin C$. Hence, there is some $r > 0$ such that $B(0, r) \cap C = \emptyset$. By Theorem 1.6 there is a closed hyperplane that separates $B(0, r)$ and C. Therefore, there is some $f \in E^\star$, $f \not\equiv 0$, such that

$$f(x - y) \leq f(rz) \quad \forall x \in A, \quad \forall y \in B, \quad \forall z \in B(0, 1).$$

It follows that $f(x - y) \leq -r\|f\| \quad \forall x \in A, \forall y \in B$. Letting $\varepsilon = \frac{1}{2}r\|f\| > 0$, we obtain

$$f(x) + \varepsilon \leq f(y) - \varepsilon \quad \forall x \in A, \quad \forall y \in B.$$

Choosing α such that

$$\sup_{x \in A} f(x) + \varepsilon \leq \alpha \leq \inf_{y \in B} f(y) - \varepsilon,$$

we see that the hyperplane $[f = \alpha]$ strictly separates A and B.

Proposition 1 (Distance functional) *Let X be a normed space and Y be a proper closed subspace of X. Take $x \in X \setminus Y$ and set*

$$d = \operatorname{dist}(x, Y) := \inf\{\|x - y\| : \ y \in Y\} > 0.$$

Then there is an $f \in X^*$ such that $\|f\|_{X^*} = 1$, $f(y) = 0$ for every $y \in Y$, and $f(x) = d$.

Proof First note that $d > 0$ since Y is closed.
Let $U = \operatorname{Span}\{Y \cup \{x\}\}$ and define $\phi : U \to \mathbb{K}$ by setting

$$\phi(y + \lambda x) := \lambda d, \qquad y \in Y, \ \lambda \in \mathbb{K}.$$

To see that ϕ is bounded on U, observe that

$$|\phi(y + \lambda x)| = |\lambda| d \leq |\lambda| \, \|x - (-y/\lambda)\| = \|\lambda x + y\|$$

since $(-y/\lambda) \in Y$ and d is the distance between x and Y; it follows that $\|\phi\|_{U^*} \leq 1$.

To see that $\|\phi\|_{U^*} \geq 1$, take $y_n \in Y$ such that

$$\|x - y_n\| \leq d\left(1 + \frac{1}{n}\right)$$

Then

$$\phi(-y_n + x) = d \geq \frac{n}{n+1} \|x - y_n\|$$

and so $\|\phi\|_{U^*} \geq n/(n+1)$ for every n, i.e. $\|\phi\|_{U^*} \geq 1$.

We have therefore shown that $\|\phi\|_{U^*} = 1$. We now extend ϕ to an element $f \in X^*$ using the Hahn–Banach Theorem; the resulting f satisfies $\|f\|_{X^*} = 1$, $f(x) = d$, and $f(y) = 0$ for every $y \in Y$, as required. $\square$

Lemma Let X be a normed space. If X* is separable, then X is separable.

Proof Since X^* is separable,

$$S_{X^*} = \{f \in X^* : \|f\|_{X^*} = 1\}$$

(the unit sphere in X^*) is separable. Let $\{f_n\}$ be a countable dense subset of S_{X^*}. Since $\|f_n\|_{X^*} = 1$, for each n, there exists an $x_n \in X$ with $\|x_n\| = 1$ such that $|f_n(x_n)| \geq 1/2$, by the definition of the norm in X^*.

We now show that

$$M := \operatorname{clin}(\{x_n\})$$

is all of X, which will imply that X is separable.

Suppose for a contradiction that $M \neq X$. Then M is a proper closed subspace of X, and so Proposition 1 (existence of a distance functional) provides an $f \in X^*$ with $\|f\|_{X^*} = 1$ (i.e. $f \in S_{X^*}$) and $f(x) = 0$ for every $x \in M$. But then $f(x_n) = 0$ for every n and so

$$\frac{1}{2} \leq |f_n(x_n)| = |f_n(x_n) - f(x_n)| \leq \|f_n - f\|_{X^*}\|x_n\| = \|f_n - f\|_{X^*}$$

for every n, which contradicts the fact that $\{f_n\}$ is dense in S_{X^*}. $\square$

Let X, Y be normed spaces. For every operator $T \in \mathcal{B}(X,Y)$ and every functional $y^* \in Y^*$, the function T^*y^* on X given by the formula

$$T^*y^*(x) := y^*(Tx),$$

is linear and continuous on X. The obtained linear mapping

$$T^*\colon\ Y^* \to X^*$$

is called the *adjoint operator*. It is continuous and satisfies the equality

$$\|T^*\| = \|T\|.$$

Indeed,

$$\|T^*y^*\| = \sup_{\|x\|\leqslant 1} |T^*y^*(x)| = \sup_{\|x\|\leqslant 1} |y^*(Tx)| \leqslant \|T\|\|y^*\|,$$

since $\|Tx\| \leqslant \|T\|$. On the other hand, for every $\varepsilon > 0$ there exists $x \in X$ with $\|x\| = 1$ and $\|Tx\| > \|T\| - \varepsilon$. By the Hahn–Banach theorem there exists a functional $y^* \in Y^*$ with $\|y^*\| = 1$ and

$$|T^*y^*(x)| = |y^*(Tx)| = \|Tx\| > \|T\| - \varepsilon,$$

which gives $\|T^*\| \geqslant \|T\|$.

Theorem Let E and F be Banach spaces. If $T \in K(E,F)$, then $T^* \in K(F^*,E^*)$.

Proof. We have to show that $T^\star(B_{F^\star})$ has compact closure in $E^\star$. Let (v_n) be a sequence in $B_{F^\star}$. We claim that $(T^\star(v_n))$ has a convergent subsequence. Set $K = \overline{T(B_E)}$; this is a compact metric space. Consider the set $\mathcal{H} \subset C(K)$ defined by

$$\mathcal{H} = \{\varphi_n : x \in K \longmapsto \langle v_n, x\rangle ; n = 1, 2, \ldots\}.$$

The assumptions of Ascoli–Arzelà's theorem are satisfied. Thus, there is a subsequence, denoted by φ_{n_k}, that converges uniformly on K to some continuous function $\varphi \in C(K)$. In particular, we have

$$\sup_{u \in B_E} \left| \langle v_{n_k}, Tu\rangle - \varphi(Tu)\right| \underset{k\to\infty}{\longrightarrow} 0 .$$

Thus

$$\sup_{u \in B_E} \left| \langle v_{n_k}, Tu\rangle - \langle v_{n_\ell}, Tu\rangle\right| \underset{k,\ell\to\infty}{\longrightarrow} 0,$$

i.e., $\|T^\star v_{n_k} - T^\star v_{n_\ell}\|_{E^\star} \underset{k,\ell\to\infty}{\longrightarrow} 0$. Consequently $T^\star v_{n_k}$ converges in $E^\star$.