

Chapter II
Elements of Functional Analysis

LECTURE 8

1. Hahn-Banach Theorems

The result we are going to discuss is one of the most fundamental theorems in the whole field of Functional Analysis. Its statement is simple but quite technical.

DEFINITIONS. Let \mathbb{K} be either of the fields \mathbb{R} or \mathbb{C} . Suppose \mathcal{X} is a \mathbb{K} -vector space.

A. A map $q : \mathcal{X} \rightarrow \mathbb{R}$ is said to be a *quasi-seminorm*, if

- (i) $q(x + y) \leq q(x) + q(y)$, for all $x, y \in \mathcal{X}$;
- (ii) $q(tx) = tq(x)$, for all $x \in \mathcal{X}$ and all $t \in \mathbb{R}$ with $t \geq 0$.

B. A map $q : \mathcal{X} \rightarrow \mathbb{R}$ is said to be a *seminorm* if, in addition to the above two properties, it satisfies:

- (ii') $q(\lambda x) = |\lambda|q(x)$, for all $x \in \mathcal{X}$ and all $\lambda \in \mathbb{K}$.

Remark that if $q : \mathcal{X} \rightarrow \mathbb{R}$ is a seminorm, then $q(x) \geq 0$, for all $x \in \mathcal{X}$. (Use $2q(x) = q(x) + q(-x) \geq q(0) = 0$.)

There are several versions of the Hahn-Banach Theorem.

THEOREM 1.1 (Hahn-Banach, \mathbb{R} -version). *Let \mathcal{X} be an \mathbb{R} -vector space. Suppose $q : \mathcal{X} \rightarrow \mathbb{R}$ is a quasi-seminorm. Suppose also we are given a linear subspace $\mathcal{Y} \subset \mathcal{X}$ and a linear map $\phi : \mathcal{Y} \rightarrow \mathbb{R}$, such that*

$$\phi(y) \leq q(y), \text{ for all } y \in \mathcal{Y}.$$

Then there exists a linear map $\psi : \mathcal{X} \rightarrow \mathbb{R}$ such that

- (i) $\psi|_{\mathcal{Y}} = \phi$;
- (ii) $\psi(x) \leq q(x)$ for all $x \in \mathcal{X}$.

PROOF. We first prove the Theorem in the following:

Particular Case: Assume $\dim \mathcal{X}/\mathcal{Y} = 1$.

This means there exists some vector $x_0 \in \mathcal{X}$ such that

$$\mathcal{X} = \{y + sx_0 : y \in \mathcal{Y}, s \in \mathbb{R}\}.$$

What we need is to prescribe the value $\psi(x_0)$. In other words, we need a number $\alpha \in \mathbb{R}$ such that, if we define $\psi : \mathcal{X} \rightarrow \mathbb{R}$ by $\psi(y + sx_0) = \phi(y) + s\alpha$, $\forall y \in \mathcal{Y}, s \in \mathbb{R}$, then this map satisfies condition (ii). For $s > 0$, condition (ii) reads:

$$\phi(y) + s\alpha \leq q(y + sx_0), \forall y \in \mathcal{Y}, s > 0,$$

and, upon dividing by s (set $z = s^{-1}y$), is equivalent to:

$$(1) \quad \alpha \leq q(z + x_0) - \phi(z), \forall z \in \mathcal{Y}.$$

For $s < 0$, condition (ii) reads (use $t = -s$):

$$\phi(y) - t\alpha \leq q(y - tx_0), \forall y \in \mathcal{Y}, t > 0,$$

and, upon dividing by t (set $w = t^{-1}y$), is equivalent to:

$$(2) \quad \alpha \geq \phi(w) - q(w - x_0), \quad \forall w \in \mathcal{Y}.$$

Consider the sets

$$\begin{aligned} Z &= \{q(z + x_0) - \phi(z); z \in \mathcal{Y}\} \subset \mathbb{R} \\ W &= \{\phi(w) - q(w - x_0) : w \in \mathcal{Y}\} \subset \mathbb{R}. \end{aligned}$$

The conditions (1) and (2) are equivalent to the inequalities

$$(3) \quad \sup W \leq \alpha \leq \inf Z.$$

This means that, in order to find a real number α with the desired property, it suffices to prove that $\sup W \leq \inf Z$, which in turn is equivalent to

$$(4) \quad \phi(w) - q(w - x_0) \leq q(z + x_0) - \phi(z), \quad \forall z, w \in \mathcal{Y}.$$

But the condition (4) is equivalent to

$$\phi(z + w) \leq q(z + x_0) + q(w - x_0),$$

which is obviously satisfied because

$$\phi(z + w) \leq q(z + w) = q((z + x_0) + (w - x_0)) \leq q(z + x_0) + q(w - x_0).$$

Having proved the Theorem in this particular case, let us proceed now with the general case. Let us consider the set Ξ of all pairs (\mathcal{Z}, ν) with

- \mathcal{Z} is a subspace of \mathcal{X} such that $\mathcal{Z} \supset \mathcal{Y}$;
- $\nu : \mathcal{Z} \rightarrow \mathbb{R}$ is a linear functional such that
 - (i) $\nu|_{\mathcal{Y}} = \phi$;
 - (ii) $\nu(z) \leq q(z)$, for all $z \in \mathcal{Z}$.

Put an order relation \succ on Ξ as follows:

$$(\mathcal{Z}_1, \nu_1) \succ (\mathcal{Z}_2, \nu_2) \Leftrightarrow \begin{cases} \mathcal{Z}_1 \supset \mathcal{Z}_2 \\ \nu_1|_{\mathcal{Z}_2} = \nu_2 \end{cases}$$

Using Zorn's Lemma, Ξ possesses a maximal element (\mathcal{Z}, ψ) . The proof of the Theorem is finished once we prove that $\mathcal{Z} = \mathcal{X}$. Assume $\mathcal{Z} \subsetneq \mathcal{X}$ and choose a vector $x_0 \in \mathcal{X} \setminus \mathcal{Z}$. Form the subspace $\mathcal{V} = \{z + tx_0 : z \in \mathcal{Z}, t \in \mathbb{R}\}$ and apply the particular case of the Theorem for the inclusion $\mathcal{Z} \subset \mathcal{V}$, for $\psi : \mathcal{Z} \rightarrow \mathbb{R}$ and for the quasi-seminorm $q|_{\mathcal{V}} : \mathcal{V} \rightarrow \mathbb{R}$. It follows that there exists some linear functional $\eta : \mathcal{M} \rightarrow \mathbb{R}$ such that

- (i) $\eta|_{\mathcal{Z}} = \psi$ (in particular we will also have $\eta|_{\mathcal{Y}} = \phi$);
- (ii) $\eta(v) \leq q(v)$, for all $v \in \mathcal{V}$.

But then the element $(\mathcal{V}, \eta) \in \Xi$ will contradict the maximality of (\mathcal{Z}, ψ) . \square

THEOREM 1.2 (Hahn-Banach, \mathbb{C} -version). *Let \mathcal{X} be an \mathbb{C} -vector space. Suppose $q : \mathcal{X} \rightarrow \mathbb{R}$ is a quasi-seminorm. Suppose also we are given a linear subspace $\mathcal{Y} \subset \mathcal{X}$ and a linear map $\phi : \mathcal{Y} \rightarrow \mathbb{C}$, such that*

$$\operatorname{Re} \phi(y) \leq q(y), \quad \text{for all } y \in \mathcal{Y}.$$

Then there exists a linear map $\psi : \mathcal{X} \rightarrow \mathbb{R}$ such that

- (i) $\psi|_{\mathcal{Y}} = \phi$;
- (ii) $\operatorname{Re} \psi(x) \leq q(x)$ for all $x \in \mathcal{X}$.

PROOF. Regard for the moment both \mathcal{X} and \mathcal{Y} as \mathbb{R} -vector spaces. Define the \mathbb{R} -linear map $\phi_1 : \mathcal{Y} \rightarrow \mathbb{R}$ by $\phi_1(y) = \operatorname{Re} \phi(y)$, for all $y \in \mathcal{Y}$, so that we have

$$\phi_1(y) \leq q(y), \quad \forall y \in \mathcal{Y}.$$

Use Theorem 1 to find an \mathbb{R} -linear map $\psi_1 : \mathcal{X} \rightarrow \mathbb{R}$ such that

- (i) $\psi_1|_{\mathcal{Y}} = \phi_1$;
- (ii) $\psi_1(x) \leq q(x)$, for all $x \in \mathcal{X}$.

Define the map $\psi : \mathcal{X} \rightarrow \mathbb{C}$ by

$$\psi(x) = \psi_1(x) - i\psi_1(ix), \quad \text{for all } x \in \mathcal{X}.$$

Claim 1: ψ is \mathbb{C} -linear.

It is obvious that ψ is \mathbb{R} -linear, so the only thing to prove is that $\psi(ix) = i\psi(x)$, for all $x \in \mathcal{X}$. But this is quite obvious:

$$\begin{aligned} \psi(ix) &= \psi_1(ix) - i\psi_1(i^2x) = \psi_1(ix) - i\psi_1(-x) = \\ &= -i^2\psi_1(x) + i\psi_1(x) = i(\psi_1(x) - i\psi_1(ix)) = i\psi(x), \quad \forall x \in \mathcal{X}. \end{aligned}$$

Because of the way ψ is defined, and because ψ_1 is real-valued, condition (ii) in the Theorem follows immediately

$$\operatorname{Re} \psi(x) = \psi_1(x) \leq q(x), \quad \forall x \in \mathcal{X},$$

so in order to finish the proof, we need to prove condition (i) in the Theorem, (i.e. $\psi|_{\mathcal{Y}} = \phi$). This follows from the fact that $\phi_1 = \psi_1|_{\mathcal{Y}}$, and from:

Claim 2: For every $y \in \mathcal{Y}$, we have $\phi(y) = \phi_1(y) - i\phi_1(iy)$.

But this is quite obvious, because

$$\operatorname{Im} \phi(y) = -\operatorname{Re}(i\phi(y)) = -\operatorname{Re} \phi(iy) = -\phi_1(iy), \quad \forall y \in \mathcal{Y}.$$

□

THEOREM 1.3 (Hahn-Banach, for seminorms). *Let \mathcal{X} be a \mathbb{K} -vector space (\mathbb{K} is either \mathbb{R} or \mathbb{C}). Suppose q is a seminorm on \mathcal{X} . Suppose also we are given a linear subspace $\mathcal{Y} \subset \mathcal{X}$ and a linear map $\phi : \mathcal{Y} \rightarrow \mathbb{K}$, such that*

$$|\phi(y)| \leq q(y), \quad \text{for all } y \in \mathcal{Y}.$$

Then there exists a linear map $\psi : \mathcal{X} \rightarrow \mathbb{K}$ such that

- (i) $\psi|_{\mathcal{Y}} = \phi$;
- (ii) $|\psi(x)| \leq q(x)$ for all $x \in \mathcal{X}$.

PROOF. We are going to apply Theorems 1 and 2, using the fact that q is also a quasi-seminorm.

THE CASE $\mathbb{K} = \mathbb{R}$. Remark that

$$\phi(y) \leq |\phi(y)| \leq q(y), \quad \forall y \in \mathcal{Y}.$$

So we can apply Theorem 1 and find $\psi : \mathcal{X} \rightarrow \mathbb{R}$ with

- (i) $\psi|_{\mathcal{Y}} = \phi$;
- (ii) $\psi(x) \leq q(x)$, for all $x \in \mathcal{X}$.

Using condition (ii) we also get

$$-\psi(x) = \psi(-x) \leq q(-x) = q(x), \text{ for all } x \in \mathcal{X}.$$

In other words we get

$$\pm\psi(x) \leq q(x), \text{ for all } x \in \mathcal{X},$$

which of course gives the desired property (ii) in the Theorem.

THE CASE $\mathbb{K} = \mathbb{C}$. Remark that

$$\operatorname{Re} \phi(y) \leq |\phi(y)| \leq q(y), \quad \forall y \in \mathcal{Y}.$$

So we can apply Theorem 2 and find $\psi : \mathcal{X} \rightarrow \mathbb{R}$ with

- (i) $\psi|_{\mathcal{Y}} = \phi$;
- (ii) $\operatorname{Re} \psi(x) \leq q(x)$, for all $x \in \mathcal{X}$.

Using condition (ii) we also get

$$(5) \quad \operatorname{Re} (\lambda\psi(x)) = \operatorname{Re} \psi(\lambda x) \leq q(\lambda x) = q(x), \text{ for all } x \in \mathcal{X} \text{ and all } \lambda \in \mathbb{T}.$$

(Here $\mathbb{T} = \{\lambda \in \mathbb{C} : |\lambda| = 1\}$.) Fix for the moment $x \in \mathcal{X}$. There exists some $\lambda \in \mathbb{T}$ such that $|\psi(x)| = \lambda\psi(x)$. For this particular λ we will have $\operatorname{Re} (\lambda\psi(x)) = |\psi(x)|$, so the inequality (5) will give

$$|\psi(x)| \leq q(x).$$

□

In the remainder of this section we will discuss the geometric form of the Hahn-Banach theorems. We begin by describing a method of constructing quasi-seminorms.

PROPOSITION 1.1. *Let \mathcal{X} be a real vector space. Suppose $\mathcal{C} \subset \mathcal{X}$ is a convex subset, which contains 0, and has the property*

$$(6) \quad \bigcup_{t>0} t\mathcal{C} = \mathcal{X}.$$

For every $x \in \mathcal{X}$ we define

$$Q_{\mathcal{C}}(x) = \inf\{t > 0 : x \in t\mathcal{C}\}.$$

(By (6) the set in the right hand side is non-empty.) Then the map $Q_{\mathcal{C}} : \mathcal{X} \rightarrow \mathbb{R}$ is a quasi-seminorm.

PROOF. For every $x \in \mathcal{X}$, let us define the set

$$T_{\mathcal{C}}(x) = \{t > 0 : x \in t\mathcal{C}\}.$$

It is pretty clear that, since $0 \in \mathcal{C}$, we have

$$T_{\mathcal{C}}(0) = (0, \infty),$$

so we get

$$Q_{\mathcal{C}}(0) = \inf T_{\mathcal{C}}(0) = 0.$$

Claim 1: For every $x \in \mathcal{X}$ and every $\lambda > 0$, one has the equality

$$T_{\mathcal{C}}(\lambda x) = \lambda T_{\mathcal{C}}(x).$$

Indeed, if $t \in T_{\mathcal{C}}(\lambda x)$, we have $\lambda x \in t\mathcal{C}$, which means that $\lambda^{-1}tx \in \mathcal{C}$, i.e. $\lambda^{-1}t \in T_{\mathcal{C}}(x)$. Consequently we have

$$t = \lambda(\lambda^{-1}t) \in \lambda T_{\mathcal{C}}(x),$$

which proves the inclusion

$$T_{\mathcal{C}}(\lambda x) \subset \lambda T_{\mathcal{C}}(x).$$

To prove the other inclusion, we start with some $s \in \lambda T_{\mathcal{C}}(x)$, which means that there exists some $t \in T_{\mathcal{C}}(x)$ with $\lambda t = s$. The fact that $t = \lambda^{-1}s$ belongs to $T_{\mathcal{C}}(x)$ means that $x \in \lambda^{-1}s\mathcal{C}$, so get $\lambda x \in s\mathcal{C}$, so s indeed belongs to $T_{\mathcal{C}}(\lambda x)$.

Claim 2: For every $x, y \in \mathcal{X}$, one has the inclusion¹

$$T_{\mathcal{C}}(x + y) \supset T_{\mathcal{C}}(x) + T_{\mathcal{C}}(y).$$

Start with some $t \in T_{\mathcal{C}}(x)$ and some $s \in T_{\mathcal{C}}(y)$. Define the elements $u = t^{-1}x$ and $v = s^{-1}y$. Since $u, v \in \mathcal{C}$, and \mathcal{C} is convex, it follows that \mathcal{C} contains the element

$$\frac{t}{t+s}u + \frac{s}{t+s}v = \frac{1}{t+s}(x+y),$$

which means that $x+y \in (t+s)\mathcal{C}$, so $t+s$ indeed belongs to $T_{\mathcal{C}}(x+y)$.

We can now conclude the proof. If $x \in \mathcal{X}$ and $\lambda > 0$, then the equality

$$Q_{\mathcal{C}}(\lambda x) = \lambda Q_{\mathcal{C}}(x)$$

is an immediate consequence of Claim 1. If $x, y \in \mathcal{X}$, then the inequality

$$Q_{\mathcal{C}}(x+y) \leq \lambda Q_{\mathcal{C}}(x) + Q_{\mathcal{C}}(y)$$

is an immediate consequence of Claim 2. □

DEFINITION. Under the hypothesis of the above proposition, the quasi-semi-norm $Q_{\mathcal{C}}$ is called the *Minkowski functional* associated with the set \mathcal{C} .

REMARK 1.1. Let \mathcal{X} be a real vector space. Suppose $\mathcal{C} \subset \mathcal{X}$ is a convex subset, which contains 0, and has the property (6). Then one has the inclusions

$$\{x \in \mathcal{X} : Q_{\mathcal{C}}(x) < 1\} \subset \mathcal{C} \subset \{x \in \mathcal{X} : Q_{\mathcal{C}}(x) \leq 1\}.$$

The second inclusion is pretty obvious, since if we start with some $x \in \mathcal{C}$, using the notations from the proof of Proposition 2.1, we have $1 \in T_{\mathcal{C}}(x)$, so

$$Q_{\mathcal{C}}(x) = \inf T_{\mathcal{C}}(x) \leq 1.$$

To prove the first inclusion, start with some $x \in \mathcal{X}$ with $Q_{\mathcal{C}}(x) < 1$. In particular this means that there exists some $t \in (0, 1)$ such that $x \in t\mathcal{C}$. Define the vector $y = t^{-1}x \in \mathcal{C}$ and notice now that, since \mathcal{C} is convex, it will contain the convex combination $ty + (1-t)0 = x$.

Exercise 1. Let \mathcal{X} be a real vector space, and let $q : \mathcal{X} \rightarrow \mathbb{R}$ be a quasi-seminorm. Define the sets

$$\begin{aligned} \mathcal{C}_0 &= \{x \in \mathcal{X} : q(x) < 1\}, \\ \mathcal{C}_1 &= \{x \in \mathcal{X} : q(x) \leq 1\}. \end{aligned}$$

- (i) Prove that \mathcal{C}_0 and \mathcal{C}_1 are both convex, they contain 0, and they both have property (6).

¹For subsets $T, S \subset \mathbb{R}$ we define $T + S = \{t + s : t \in T, s \in S\}$.

(ii) Let \mathcal{C} is any convex set with

$$\mathcal{C}_0 \subset \mathcal{C} \subset \mathcal{C}_1.$$

Analyze the relationship between $Q_{\mathcal{C}}$ and q .

DEFINITION. A *topological vector space* is a vector space \mathcal{X} over \mathbb{K} (which is either \mathbb{R} or \mathbb{C}), which is also a topological space, such that the maps

$$\mathcal{X} \times \mathcal{X} \ni (x, y) \longmapsto x + y \in \mathcal{X}$$

$$\mathbb{K} \times \mathcal{X} \ni (\lambda, x) \longmapsto \lambda x \in \mathcal{X}$$

are continuous.

REMARK 1.2. Let \mathcal{X} be a real topological vector space. Suppose $\mathcal{C} \subset \mathcal{X}$ is a convex *open* subset, which contains 0. Then \mathcal{C} has the property (6). Moreover (compare with Remark 2.1), one has the equality

$$(7) \quad \{x \in \mathcal{X} : Q_{\mathcal{C}}(x) < 1\} = \mathcal{C}.$$

To prove this remark, we define for each $x \in \mathcal{X}$, the function

$$F_x : \mathbb{R} \ni t \longmapsto tx \in \mathcal{X}.$$

Since \mathcal{X} is a topological vector space, the map F_x , $x \in \mathcal{X}$ are continuous. To prove the property (6) we start with an arbitrary $x \in \mathcal{X}$, and we use the continuity of the map F_x at 0. Since \mathcal{C} is a neighborhood of 0, there exists some $\rho > 0$ such that

$$F_x(t) \in \mathcal{C}, \quad \forall t \in [-\rho, \rho].$$

In particular we get $\rho x \in \mathcal{C}$, which means that $x \in \rho^{-1}\mathcal{C}$.

To prove the equality (7) we only need to prove the inclusion “ \supset ” (since the inclusion “ \subset ” holds in general, by Remark 2.1). Start with some element $x \in \mathcal{C}$. Using the continuity of the map F_x at 1, plus the fact that $F_x(1) = x \in \mathcal{C}$, there exists some $\varepsilon > 0$, such that

$$F_x(t) \in \mathcal{C}, \quad \forall t \in [1 - \varepsilon, 1 + \varepsilon].$$

In particular, we have $F(1 + \varepsilon) \in \mathcal{C}$, which means precisely that

$$x \in (1 + \varepsilon)^{-1}\mathcal{C}.$$

This gives the inequality

$$Q_{\mathcal{C}}(x) \leq (1 + \varepsilon)^{-1},$$

so we indeed get $Q_{\mathcal{C}}(x) < 1$.

The first geometric version of the Hahn-Banach Theorem is:

LEMMA 1.1. *Let \mathcal{X} be a real topological vector space, and let $\mathcal{C} \subset \mathcal{X}$ be a convex open set which contains 0. If $x_0 \in \mathcal{X}$ is some point which does not belong to \mathcal{C} , then there exists a linear continuous map $\phi : \mathcal{X} \rightarrow \mathbb{R}$, such that*

- $\phi(x_0) = 1$;
- $\phi(v) < 1, \forall v \in \mathcal{C}$.

PROOF. Consider the linear subspace

$$\mathcal{Y} = \mathbb{R}x_0 = \{tx_0 : t \in \mathbb{R}\},$$

and define $\psi : \mathcal{Y} \rightarrow \mathbb{R}$ by

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

It is obvious that ψ is linear, and $\psi(x_0) = 1$.

Claim: One has the inequality

$$\psi(y) \leq Q_{\mathcal{C}}(y), \quad \forall y \in \mathcal{Y}.$$

Let y be represented as $y = tx_0$ for some $t \in \mathbb{R}$. If $t \leq 0$, the inequality is clear, because $\psi(y) = t \leq 0$ and the right hand side $Q_{\mathcal{C}}(y)$ is always non-negative. Assume $t > 0$. Since $Q_{\mathcal{C}}$ is a quasi-seminorm, we have

$$(8) \quad Q_{\mathcal{C}}(y) = Q_{\mathcal{C}}(tx_0) = tQ_{\mathcal{C}}(x_0),$$

and the fact that $x_0 \notin \mathcal{C}$ will give (by Remark 2.2) the inequality $Q_{\mathcal{C}}(x_0) \geq 1$. Since $t > 0$, the computation (8) can be continued with

$$Q_{\mathcal{C}}(y) = tQ_{\mathcal{C}}(x_0) \geq t = \psi(y),$$

so the Claim follows also in this case.

Use now the Hahn-Banach Theorem, to find a linear map $\phi : \mathcal{X} \rightarrow \mathbb{R}$ such that

- (i) $\phi|_{\mathcal{Y}} = \psi$;
- (ii) $\phi(x) \leq Q_{\mathcal{C}}(x), \quad \forall x \in \mathcal{X}$.

It is obvious that (i) gives $\phi(x_0) = \psi(x_0) = 1$. If $v \in \mathcal{C}$, then by Remark 2.2 we have $Q_{\mathcal{C}}(v) < 1$, so by (ii) we also get $\phi(v) < 1$. This means that the only thing that remains to be proven is the continuity of ϕ . Since ϕ is linear, we only need to prove that ϕ is continuous at 0. Start with some $\varepsilon > 0$. We must find some open set $\mathcal{U}_{\varepsilon} \subset \mathcal{X}$, with $\mathcal{U}_{\varepsilon} \ni 0$, such that

$$|\phi(u)| < \varepsilon, \quad \forall u \in \mathcal{U}_{\varepsilon}.$$

We take $\mathcal{U}_{\varepsilon} = (\varepsilon\mathcal{C}) \cap (-\varepsilon\mathcal{C})$. Notice that, for every $u \in \mathcal{U}_{\varepsilon}$, we have $\pm u \in \varepsilon\mathcal{C}$, which gives $\varepsilon^{-1}(\pm u) \in \mathcal{C}$. By Remark 2.2 this gives $Q_{\mathcal{C}}(\varepsilon^{-1}(\pm u)) < 1$, which gives

$$Q_{\mathcal{C}}(\pm u) < \varepsilon.$$

Then using property (ii) we immediately get

$$\phi(\pm u) < \varepsilon,$$

and we are done. □

It turns out that the above result is a particular case of a more general result:

THEOREM 1.4 (Hahn-Banach Separation Theorem - real case). *Let \mathcal{X} be a real topological vector space, let $\mathcal{A}, \mathcal{B} \subset \mathcal{X}$ be non-empty convex sets with \mathcal{A} open, and $\mathcal{A} \cap \mathcal{B} = \emptyset$. Then there exists a linear continuous map $\phi : \mathcal{X} \rightarrow \mathbb{R}$, and a real number α , such that*

$$\phi(a) < \alpha \leq \phi(b), \quad \forall a \in \mathcal{A}, b \in \mathcal{B}.$$

PROOF. Fix some points $a_0 \in \mathcal{A}, b_0 \in \mathcal{B}$, and define the set

$$\mathcal{C} = \mathcal{A} - \mathcal{B} + b_0 - a_0 = \{a - b + b_0 - a_0 : a \in \mathcal{A}, b \in \mathcal{B}\}.$$

It is straightforward that \mathcal{C} is convex and contains 0. The equality

$$\mathcal{C} = \bigcup_{b \in \mathcal{B}} (\mathcal{A} + b_0 - a_0)$$

shows that \mathcal{C} is also open. Define the vector $x_0 = b_0 - a_0$. Since $\mathcal{A} \cap \mathcal{B} = \emptyset$, it is clear that $x_0 \notin \mathcal{C}$.

Use Lemma 2.1 to produce a linear continuous map $\phi : \mathcal{X} \rightarrow \mathbb{R}$ such that

- (i) $\phi(x_0) = 1$;
- (ii) $\phi(v) < 1, \forall v \in \mathcal{C}$.

By the definition of x_0 and \mathcal{C} , we have $\phi(b_0) = \phi(a_0) + 1$, and

$$\phi(a) < \phi(b) + \phi(a_0) - \phi(b_0) + 1, \quad \forall a \in \mathcal{A}, b \in \mathcal{B},$$

which gives

$$(9) \quad \phi(a) < \phi(b), \quad \forall a \in \mathcal{A}, b \in \mathcal{B}.$$

Put

$$\alpha = \inf_{b \in \mathcal{B}} \phi(b).$$

The inequalities (9) give

$$(10) \quad \phi(a) \leq \alpha \leq \phi(b), \quad \forall a \in \mathcal{A}, b \in \mathcal{B}.$$

The proof will be complete once we prove the following

Claim: One has the inequality

$$\phi(a) < \alpha, \quad \forall a \in \mathcal{A}.$$

Suppose the contrary, i.e. there exists some $a_1 \in \mathcal{A}$ with $\phi(a_1) = \alpha$. Using the continuity of the map

$$\mathbb{R} \ni t \longmapsto a_1 + tx_0 \in \mathcal{X}$$

there exists some $\varepsilon > 0$ such that

$$a_1 + tx_0 \in \mathcal{A}, \quad \forall t \in [-\varepsilon, \varepsilon].$$

In particular, by (10) one has

$$\phi(a_1 + \varepsilon x_0) \leq \alpha,$$

which means that

$$\alpha + \varepsilon \leq \alpha,$$

which is clearly impossible. \square

THEOREM 1.5 (Hahn-Banach Separation Theorem - complex case). *Let \mathcal{X} be a complex topological vector space, let $\mathcal{A}, \mathcal{B} \subset \mathcal{X}$ be non-empty convex sets with \mathcal{A} open, and $\mathcal{A} \cap \mathcal{B} = \emptyset$. Then there exists a linear continuous map $\phi : \mathcal{X} \rightarrow \mathbb{C}$, and a real number α , such that*

$$\operatorname{Re} \phi(a) < \alpha \leq \operatorname{Im} \phi(b), \quad \forall a \in \mathcal{A}, b \in \mathcal{B}.$$

PROOF. Regard \mathcal{X} as a real topological vector space, and apply the real version to produce an \mathbb{R} -linear continuous map $\phi_1 : \mathcal{X} \rightarrow \mathbb{R}$, and a real number α , such that

$$\phi_1(a) < \alpha \leq \phi_1(b), \quad \forall a \in \mathcal{A}, b \in \mathcal{B}.$$

Then the function $\phi : \mathcal{X} \rightarrow \mathbb{C}$ defined by

$$\phi(x) = \phi_1(x) - i\phi_1(ix), \quad x \in \mathcal{X}$$

will clearly satisfy the desired properties. \square

There is another version of the Hahn-Banach Separation Theorem, which holds for a special type of topological vector spaces. Before we discuss these, we shall need a technical result.

LEMMA 1.2. *Let \mathcal{X} be a topological vector space, let $\mathcal{C} \subset \mathcal{X}$ be a compact set, and let $\mathcal{D} \subset \mathcal{X}$ be a closed set. Then the set*

$$\mathcal{C} + \mathcal{D} = \{x + y : x \in \mathcal{C}, y \in \mathcal{D}\}$$

is closed.

PROOF. Start with some point $p \in \overline{\mathcal{C} + \mathcal{D}}$, and let us prove that $p \in \mathcal{C} + \mathcal{D}$. For every neighborhood \mathcal{U} of 0, the set $p + \mathcal{U}$ is a neighborhood of p , so by assumption, we have

$$(11) \quad (p + \mathcal{U}) \cap (\mathcal{C} + \mathcal{D}) \neq \emptyset.$$

Define, for each neighborhood \mathcal{U} of 0, the set

$$\mathcal{A}_{\mathcal{U}} = (p + \mathcal{U} - \mathcal{D}) \cap \mathcal{C}.$$

Using (11), it is clear that $\mathcal{A}_{\mathcal{U}}$ is non-empty. It is also clear that, if $\mathcal{U}_1 \subset \mathcal{U}_2$, then $\mathcal{A}_{\mathcal{U}_1} \subset \mathcal{A}_{\mathcal{U}_2}$. Using the compactness of \mathcal{C} , it follows that

$$\bigcap_{\substack{\mathcal{U} \text{ neighborhood} \\ \text{of } 0}} \overline{\mathcal{A}_{\mathcal{U}}} \neq \emptyset.$$

Choose then a point q in the above intersection. It follows that

$$(q + \mathcal{V}) \cap \mathcal{A}_{\mathcal{U}} \neq \emptyset,$$

for any two neighborhoods \mathcal{U} and \mathcal{V} of 0. In other words, for any two such neighborhoods of 0, we have

$$(12) \quad (q + \mathcal{V} - \mathcal{U}) \cap (p - \mathcal{D}) \neq \emptyset.$$

Fix now an arbitrary neighborhood \mathcal{W} of 0. Using the continuity of the map

$$\mathcal{X} \times \mathcal{X} \ni (x_1, x_2) \longmapsto x_1 - x_2 \in \mathcal{X},$$

there exist neighborhoods \mathcal{U} and \mathcal{V} of 0, such that $\mathcal{U} - \mathcal{V} \subset \mathcal{W}$. Then $q + \mathcal{V} - \mathcal{U} \subset q - \mathcal{W}$, so (12) gives

$$(q - \mathcal{W}) \cap (p - \mathcal{D}) \neq \emptyset,$$

which yields

$$(p - q + \mathcal{W}) \cap \mathcal{D} \neq \emptyset.$$

Since this is true for all neighborhoods \mathcal{W} of 0, we get $p - q \in \overline{\mathcal{D}}$, and since \mathcal{D} is closed, we finally get $p - q \in \mathcal{D}$. Since, by construction we have $q \in \mathcal{C}$, it follows that the point $p = q + (p - q)$ indeed belongs to $\mathcal{C} + \mathcal{D}$. \square

DEFINITION. A topological vector space \mathcal{X} is said to be *locally convex*, if every point has a fundamental system of convex open neighborhoods. This means that for every $x \in \mathcal{X}$ and every neighborhood N of x , there exists a convex open set D , with $x \in D \subset N$.

THEOREM 1.6 (Hahn-Banach Separation Theorem for Locally Convex Spaces). *Let \mathbb{K} be one of the fields \mathbb{R} or \mathbb{C} , and let \mathcal{X} be a locally convex \mathbb{K} -vector space. Suppose $\mathcal{C}, \mathcal{D} \subset \mathcal{X}$ are convex sets, with \mathcal{C} compact, \mathcal{D} closed, and $\mathcal{C} \cap \mathcal{D} = \emptyset$. Then there exists a linear continuous map $\phi : \mathcal{X} \rightarrow \mathbb{K}$, and two numbers $\alpha, \beta \in \mathbb{R}$, such that*

$$\operatorname{Re} \phi(x) \leq \alpha < \beta \leq \operatorname{Re} \phi(y), \quad \forall x \in \mathcal{C}, y \in \mathcal{D}.$$

PROOF. Consider the convex set $\mathcal{B} = \mathcal{D} - \mathcal{C}$. By Lemma ??, \mathcal{B} is closed. Since $\mathcal{C} \cap \mathcal{D} = \emptyset$, we have $0 \notin \mathcal{B}$. Since \mathcal{B} is closed, its complement $\mathcal{X} \setminus \mathcal{B}$ will then be a neighborhood of 0. Since \mathcal{X} is locally convex, there exists a convex open set \mathcal{A} , with $0 \in \mathcal{A} \subset \mathcal{X} \setminus \mathcal{B}$. In particular we have $\mathcal{A} \cap \mathcal{B} = \emptyset$. Applying the suitable version of the Hahn-Banach Theorem (real or complex case), we find a linear continuous map $\phi : \mathcal{X} \rightarrow \mathbb{K}$, and a real number ρ , such that

$$\operatorname{Re} \phi(a) < \rho \leq \operatorname{Re} \phi(b), \quad \forall a \in \mathcal{A}, b \in \mathcal{B}.$$

Notice that, since $\mathcal{A} \ni 0$, we get $\rho > 0$. Then the inequality

$$\rho \leq \operatorname{Re} \phi(b), \quad b \in \mathcal{B}$$

gives

$$\operatorname{Re} \phi(y) - \operatorname{Re} \phi(x) \geq \rho > 0, \quad \forall x \in \mathcal{C}, y \in \mathcal{D}.$$

Then if we define

$$\beta = \inf_{y \in \mathcal{D}} \operatorname{Re} \phi(y) \quad \text{and} \quad \alpha = \sup_{x \in \mathcal{C}} \operatorname{Re} \phi(x),$$

we get $\beta \geq \alpha + \rho$, and we are done. □