*Chapter 5*

# The Hahn-Banach Theorem and its applications

## 5.1 The Hahn-Banach Theorem

One of the most important results in the theory of t.v.s. is the Hahn-Banach theorem (HBT). It is named for Hans Hahn and Stefan Banach who proved this theorem independently in the late 1920s, dealing with the problem of extending continuous linear functionals defined on a subspace of a seminormed vector space to the whole space. We will see that actually this extension problem can be reduced to the problem of separating by a closed hyperplane a convex open set and an affine submanifold (the image by a translation of a linear subspace) which do not intersect. Indeed, there are several versions of HBT in literature, but we are going to present for the moment just two of them as representatives of the analytic and the geometric side of this result.

Before stating these two versions of HBT, let us recall the notion of hyperplane in a vector space (we always consider vector spaces over the field $\mathbb{K}$ which is either $\mathbb{R}$ or $\mathbb{C}$). A *hyperplane* $H$ in a vector space $X$ over $\mathbb{K}$ is a maximal proper linear subspace of $X$ or, equivalently, a linear subspace of codimension one, i.e. $dim(X/H) = 1$. Another equivalent formulation is that a hyperplane is a set of the form $\varphi^{-1}(\{0\})$ for some linear functional $\varphi : X \to \mathbb{K}$ not identically zero. The translation by a non-null vector of a hyperplane will be called *affine hyperplane*.

**Theorem 5.1.1** (Analytic form of Hahn-Banach thm (for seminormed spaces)). *Let $p$ be a seminorm on a vector space $X$ over $\mathbb{K}$, $M$ a linear subspace of $X$, and $f$ a linear functional on $M$ such that*

$$|f(x)| \leq p(x), \forall\, x \in M. \tag{5.1}$$

*There exists a linear functional $\tilde{f}$ on $X$ such that $\tilde{f}(x) = f(x), \forall\, x \in M$ and*

$$|\tilde{f}(x)| \leq p(x), \forall\, x \in X. \tag{5.2}$$

**Theorem 5.1.2** (Geometric form of Hahn-Banach theorem)**.**
*Let $X$ be a topological vector space over $\mathbb{K}$, $N$ a linear subspace of $X$, and $\Omega$ a non-empty open convex subset of $X$ such that $N \cap \Omega = \emptyset$. Then there exists a closed hyperplane $H$ of $X$ such that*

$$N \subseteq H \quad \textit{and} \quad H \cap \Omega = \emptyset. \tag{5.3}$$

It should be remarked that the vector space $X$ does not apparently carry any topology in Theorem 5.1.1, but actually the datum of a seminorm on $X$ is equivalent to the datum of the topology induced by this seminorm. It is then clear that the conditions (5.1) and (5.2) imply the $p-$continuity of the functions $f$ and $\tilde{f}$, respectively.

Let us also stress the fact that in Theorem 5.1.2 neither local convexity nor the Hausdorff separation property are assumed on the t.v.s. $X$. Moreover, it is easy to see that the geometric form of HBT could have been stated also in an affine setting, namely starting with any affine submanifold $N$ of $X$ which does not intersect the open convex subset $\Omega$ and getting a closed affine hyperplane fulfilling (5.3).

We will first show how to derive Theorem 5.1.1 from Theorem 5.1.2 and then give a proof of Theorem 5.1.2.

Before starting the proofs, let us fix one more definition. A *convex cone* $C$ in a vector space $X$ over $\mathbb{R}$ is a subset of $X$ which is closed under addition and multiplication by positive scalars.

*Proof. Theorem 5.1.2*

**1) Existence of a linear subspace $H$ of $X$ maximal for (5.3).**
This first part of the proof is quite simple and consists in a straightforward application of Zorn's lemma. In fact, consider the family $\mathcal{F}$ of all the linear subspaces $S$ of $X$ such that

$$N \subseteq S \quad \text{and} \quad S \cap \Omega = \emptyset. \tag{5.4}$$

$\mathcal{F}$ is clearly non-empty since $N$ belongs to it by assumption. If we take now a totally ordered subfamily $\mathcal{C}$ of $\mathcal{F}$ (totally ordered for the inclusion relation $\subseteq$), then the union of all the linear subspaces belonging to $\mathcal{C}$ is a linear subspace of $X$ having the properties in (5.4) and containing each element of $\mathcal{C}$. Hence, we can apply Zorn's lemma and conclude that there exists at least a maximal element $H$ in $\mathcal{F}$.

**2) $H$ is closed in $X$.**
The fact that $H$ and $\Omega$ do not intersect gives that $H$ is contained in the

complement of $\Omega$ in $X$. This implies that also its closure $\overline{H}$ does not intersect $\Omega$. Indeed, since $\Omega$ is open, we get

$$\overline{H} \subseteq \overline{X \setminus \Omega} = X \setminus \Omega.$$

Then $\overline{H}$ is a linear subspace of $X$ (as closure of a linear subspace in a t.v.s.), which is disjoint from $\Omega$ and which contains $H$ and so $N$, i.e. $\overline{H} \in \mathcal{F}$. Hence, as $H$ is maximal in $\mathcal{F}$, it must coincide with its closure. Note that the fact that $H$ is closed guarantees that the quotient space $X/H$ is a Hausdorff t.v.s. (see Proposition 2.3.5).

**3) $H$ is an hyperplane**
We want to show that $H$ is a hyperplane, i.e. that $dim(X/H) = 1$. To this aim we distinguish the two cases when $\mathbb{K} = \mathbb{R}$ and when $\mathbb{K} = \mathbb{C}$.

**3.1) Case $\mathbb{K} = \mathbb{R}$**
Let $\phi : X \to X/H$ be the canonical map. Since $\phi$ is an open linear mapping (see Proposition 2.3.2), $\phi(\Omega)$ is an open convex subset of $X/H$. Also we have that $\phi(\Omega)$ does not contain the origin $\hat{o}$ of $X/H$. Indeed, if $\hat{o} \in \phi(\Omega)$ holds, then there would exist $x \in \Omega$ s.t. $\phi(x) = \hat{o}$ and so $x \in H$, which would contradict the assumption $H \cap \Omega = \emptyset$. Let us set:

$$A = \bigcup_{\lambda > 0} \lambda \phi(\Omega).$$

Then the subset $A$ of $X/H$ is an open convex cone which does not contain the origin $\hat{o}$.

Let us observe that $X/H$ has at least dimension 1. Indeed, if $dim(X/H) = 0$ then $X/H = \{\hat{o}\}$ and so $X = H$ which contradicts the fact that $\Omega$ does not intersect $H$ (recall that we assumed $\Omega$ is non-empty). Suppose that $dim(X/H) \geq 2$, then to get our conclusion it will suffice to show the following claims:

Claim 1: The boundary $\partial A$ of $A$ must contain at least one point $x \neq \hat{o}$.

Claim 2: The point $-x$ cannot belong to $A$.

In fact, once Claim 1 is established, we have that $x \notin A$, because $x \in \partial A$ and $A$ is open. This together with Claim 2 gives that both $x$ and $-x$ belong to the complement of $A$ in $X/H$ and, therefore, so does the straight line $L$ defined by these two points. (If there was a point $y \in L \cap A$ then any positive multiple of $y$ would belong to $L \cap A$, as $A$ is a cone. Hence, for some $\lambda > 0$ we would have $x = \lambda y \in L \cap A$, which contradicts the fact that $x \notin A$.) Then:

- $\phi^{-1}(L)$ is a linear subspace of $X$
- $\phi^{-1}(L) \cap \Omega = \emptyset$, since $L \cap A = \emptyset$
- $\phi^{-1}(L) \supsetneq H$ because $\hat{o} = \phi(H) \subseteq L$ but $L \neq \{\hat{o}\}$ since $x \neq \hat{o}$ is in $L$.

This contradicts the maximality of $H$ and so $dim(X/H) = 1$.
To complete the proof of 3.1) let us show the two claims.

*Proof. of Claim 1*
Suppose that $\partial A = \{\hat{o}\}$. This means that $A$ has empty boundary in the set $(X/H) \setminus \{\hat{o}\}$. Since $dim(X/H) \geq 2$, the space $(X/H) \setminus \{\hat{o}\}$ is path-connected and so connected. Hence, $A = (X/H) \setminus \{\hat{o}\}$ which contradicts the convexity of $A$ since $(X/H) \setminus \{\hat{o}\}$ is clearly not convex. □

*Proof. of Claim 2*
Suppose $-x \in A$. As $A$ is open, there is a neighborhood $V$ of $-x$ entirely contained in $A$. This implies that $-V$ is a neighborhood of $x$. Since $x$ is a boundary point of $A$, there exists $y \in (-V) \cap A$. But then $-y \in V \subset A$ and so, by the convexity of $A$, the whole line segment between $y$ and $-y$ is contained in $A$, in particular $\hat{o}$, which contradicts the definition of $A$.

□

**3.2) Case $\mathbb{K} = \mathbb{C}$**
Although here we are considering the scalars to be the complex numbers, we may view $X$ as a vector space over the real numbers and it is obvious that its topology, as originally given, is still compatible with its linear structure. By step 3.1) above, we know that there exists a real hyperplane $H_0$ of $X$ which contains $N$ and does not intersect $\Omega$. By a real hyperplane, we mean that $H_0$ is a linear subspace of $X$ viewed as a vector space over the field of real numbers such that $dim_{\mathbb{R}}(X/H_0) = 1$.

Now it is easy to see that $iN = N$ (here $i = \sqrt{-1}$). Hence, setting $H := H_0 \cap iH_0$, we have that $N \subseteq H$ and $H \cap \Omega = \emptyset$. Then to complete the proof it remains to show that this $H$ is a complex hyperplane. It is obviously a complex linear subspace of $X$ and its real codimension is $\geq 1$ and $\leq 2$ (since the intersection of two distinct hyperplanes is always a linear subspace with codimension two). Hence, its complex codimension is equal to one.

□

*Proof. Theorem 5.1.2 ⇒ Theorem 5.1.1*
Let $p$ be a seminorm on the vector space $X$, $M$ a linear subspace of $X$, and $f$ a linear functional defined on $M$ fulfilling (5.1). As already remarked before, this means that $f$ is continuous on $M$ w.r.t. the topology induced by $p$ on $X$ (which makes $X$ a l.c. t.v.s.).

Consider the subset $N := \{x \in M : f(x) = 1\}$. Taking any vector $x_0 \in N$, it is easy to see that $N - x_0 = Ker(f)$ (i.e. the kernel of $f$ in

$M$), which is a hyperplane of $M$ and so a linear subspace of $X$. Therefore, setting $M_0 := N - x_0$, we have the following decomposition of $M$:

$$M = M_0 \oplus \mathbb{K}x_0,$$

where $\mathbb{K}x_0$ is the one-dimensional linear subspace through $x_0$. In other words

$$\forall\, x \in M,\ \exists!\, \lambda \in \mathbb{K}, y \in M_0 :\ x = y + \lambda x_0.$$

Then

$$\forall\, x \in M,\ f(x) = f(y) + \lambda f(x_0) = \lambda f(x_0) = \lambda,$$

which means that the values of $f$ on $M$ are completely determined by the ones on $N$. Consider now the open unit semiball of $p$:

$$U := \mathring{U}_p = \{x \in X : p(x) < 1\},$$

which we know being an open convex subset of $X$ endowed with the topology induced by $p$. Then $N \cap U = \emptyset$ because if there was $x \in N \cap U$ then $p(x) < 1$ and $f(x) = 1$, which contradict (5.1).

By Theorem 5.1.2 (affine version), there exists a closed affine hyperplane $H$ of $X$ with the property that $N \subseteq H$ and $H \cap U = \emptyset$. Then $H - x_0$ is a hyperplane and so the kernel of a continuous linear functional $\tilde{f}$ on $X$ non-identically zero.

Arguing as before (consider here the decomposition $X = (H - x_0) \oplus \mathbb{K}x_0$), we can deduce that the values of $\tilde{f}$ on $X$ are completely determined by the ones on $N$ and so on $H$ (because for any $h \in H$ we have $h - x_0 \in Ker(\tilde{f})$ and so $\tilde{f}(h) - \tilde{f}(x_0) = \tilde{f}(h - x_0) = 0$). Since $\tilde{f} \not\equiv 0$, we have that $\tilde{f}(x_0) \neq 0$ and w.l.o.g. we can assume $\tilde{f}(x_0) = 1$ i.e. $\tilde{f} \equiv 1$ on $H$. Therefore, for any $x \in M$ there exist unique $\lambda \in \mathbb{K}$ and $y \in N - x_0 \subseteq H - x_0$ s.t. $x = y + \lambda x_0$, we get that:

$$\tilde{f}(x) = \lambda \tilde{f}(x_0) = \lambda = \lambda f(x_0) = f(x),$$

i.e. $f$ is the restriction of $\tilde{f}$ to $M$. Furthermore, the fact that $H \cap U = \emptyset$ means that $\tilde{f}(x) = 1$ implies $p(x) \geq 1$. Then for any $y \in X$ s.t. $\tilde{f}(y) \neq 0$ we have that: $\tilde{f}\left(\frac{y}{\tilde{f}(y)}\right) = 1$ and so that $p\left(\frac{y}{\tilde{f}(y)}\right) \geq 1$ which implies that $|\tilde{f}(y)| \leq p(y)$. The latter obviously holds for $\tilde{f}(y) = 0$. Hence, (5.2) is established. □

Theorem 5.1.2 is indeed slightly more general than Theorem 5.1.1. In fact, Theorem 5.1.2 can be deduced from the following more general analytic form of the Hahn-Banach Theorem.