Chapter 3

Hahn-Banach Theorems

3.1 Analytic Versions

The analytic form of the Hahn-Banach theorem concerns the extension of linear functionals defined on a subspace of a normed linear space to the entire space, preserving the norm of the functional. We will prove a slightly more general result in this direction.

Theorem 3.1.1 (**Hahn-Banach Theorem**) *Let V be a vector space over $\mathbb{R}$. Let $p : V \to \mathbb{R}$ be a mapping such that*

$$\left.\begin{array}{rcl} p(\alpha x) & = & \alpha p(x) \\ p(x+y) & \leq & p(x) + p(y) \end{array}\right\} \tag{3.1.1}$$

for all x and $y \in V$ and for all $\alpha > 0$ in $\mathbb{R}$. Let W be a subspace of V and let $g : W \to \mathbb{R}$ be a linear map such that

$$g(x) \leq p(x)$$

for all $x \in W$. Then, there exists a linear extension $f : V \to \mathbb{R}$ of g (i.e. $f(x) = g(x)$ for all $x \in W$) which is such that

$$f(x) \leq p(x)$$

for all $x \in V$.

Proof: Step 1. Let $\mathcal{P}$ denote the collection of all pairs (Y, h), where Y is a subspace of V containing W and $h : Y \to \mathbb{R}$ a linear map which is an extension of g and which is also such that

$$h(x) \leq p(x)$$

for all $x \in Y$. Clearly $\mathcal{P}$ is non-empty, since $(W, g) \in \mathcal{P}$. Consider the partial order defined on $\mathcal{P}$ by

$$(Y, h) \preceq (\widetilde{Y}, \widetilde{h})$$

if $Y \subset \widetilde{Y}$ and $\widetilde{h}$ is a linear extension of h. Let $\mathcal{Q} = \{(Y_i, h_i) \mid i \in I\}$ be a chain in $\mathcal{P}$. Define

$$Y = \cup_{i \in I} Y_i$$

and let $h : Y \to \mathbb{R}$ be defined by $h(x) = h_i(x)$ if $x \in Y_i$. Since $\mathcal{Q}$ is a chain, it is immediate to see that h is well defined and also that it is a linear extension of each of the h_i. Also $h(x) \leq p(x)$ for all $x \in Y$. Thus, $(Y, h) \in \mathcal{P}$ and $(Y_i, h_i) \preceq (Y, h)$ for each $i \in I$ and thus every chain has an upper bound. Hence, by Zorn's lemma, $\mathcal{P}$ has a maximal element (Z, f).

Step 2. We will show that $Z = V$, which will complete the proof. Assuming the contrary, let $x_0 \notin Z$. Consider the linear subspace of V given by

$$Y = \{x + tx_0 \mid x \in Z, \ t \in \mathbb{R}\}.$$

We will define a linear extension $h : Y \to \mathbb{R}$ of f such that $(Y, h) \in \mathcal{P}$, thus contradicting the maximality of f. Define

$$h(x + tx_0) = f(x) + \alpha t$$

where α will be suitably determined. In order that $(Y, h) \in \mathcal{P}$, we need that

$$f(x) + \alpha t \leq p(x + tx_0)$$

for all $x \in Z$ and for all $t \in \mathbb{R}$. If $t > 0$, this reduces to (in view of (3.1.1))

$$f\left(\frac{1}{t}x\right) + \alpha \leq p\left(\frac{1}{t}x + x_0\right)$$

or, equivalently,

$$f(x) + \alpha \leq p(x + x_0) \tag{3.1.2}$$

for all $x \in Z$. Similarly, by considering $t < 0$, we deduce that

$$f(x) - \alpha \leq p(x - x_0) \tag{3.1.3}$$

for all $x \in Z$. In other words, it is necessary that α be chosen such that

$$\sup_{x \in Z}[f(x) - p(x - x_0)] \leq \alpha \leq \inf_{x \in Z}[p(x + x_0) - f(x)]. \tag{3.1.4}$$

But, for all x and $y \in Z$, we have

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

or, equivalently,

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

Hence, it is possible to choose α such that (3.1.4) is true, giving the desired contradiction, which completes the proof. ■

Theorem 3.1.2 (**Hahn-Banach Theorem**) *Let V be a normed linear space over $\mathbb{R}$. Let W be a subspace of V and let $g : W \to \mathbb{R}$ be a continuous linear functional on W. Then there exists a continuous linear extension $f : V \to \mathbb{R}$ of g such that*

$$\|f\|_{V^*} \;=\; \|g\|_{W^*}.$$

Proof: Set $p(x) = \|g\|_{W^*}\|x\|$. Then p verifies (3.1.1). Also $g(x) \le p(x)$ for all $x \in W$. Thus, there exists a linear extension of g, *viz.* $f : V \to \mathbb{R}$ such that, for all $x \in V$, $f(x) \le \|g\|_{W^*}\|x\|$. This implies that f is continuous and that $\|f\|_{V^*} \le \|g\|_{W^*}$. But, since $f = g$ on W, it follows that we do have equality of the norms of f and g. This completes the proof. ■

We will now prove the same result for complex vector spaces.

Proposition 3.1.1 *Let V be a normed linear space over $\mathbb{C}$. Let $f : V \to \mathbb{C}$ be a continuous linear functional. Let $f = g + ih$ where g and h are real valued linear functionals. Then*

$$f(x) \;=\; g(x) - ig(ix)$$

for all $x \in V$ and, further, $\|f\| = \|g\|$.

Proof: Let $x \in V$. Then $f(ix) = if(x)$. Expressing this in terms of the real and imaginary parts of f, we get

$$g(ix) + ih(ix) \;=\; ig(x) - h(x)$$

which shows that $h(x) = -g(ix)$. Now, let $f(x) = e^{i\theta}|f(x)|$ where $\theta \in [0, 2\pi)$. Then,

$$|f(x)| \;=\; e^{-i\theta}f(x) \;=\; f(e^{-i\theta}x) \;=\; g(e^{-i\theta}x)$$

since the left extreme of the above relation is real. Thus $|f(x)| \leq \|g\| \, \|x\|$ which implies that $\|f\| \leq \|g\|$. On the other hand,

$$|f(x)|^2 = |g(x)|^2 + |h(x)|^2$$

which yields $|g(x)| \leq |f(x)| \leq \|f\| \, \|x\|$ whence we get $\|g\| \leq \|f\|$. This completes the proof. ■

Theorem 3.1.3 (**Hahn-Banach Theorem**) *Let V be a normed linear space over $\mathbb{C}$. Let W be a subspace of V and let $g : W \to \mathbb{C}$ be a continuous linear functional on W. Then there exists a continuous linear extension $f : V \to \mathbb{C}$ of g such that*

$$\|f\|_{V^*} = \|g\|_{W^*}.$$

Proof: Let $g = h(x) - ih(ix)$ where h is the real part of g. We consider V as a real normed linear space by restricting ourselves to scalar multiplication by reals only. Then, there exists $\widetilde{h} : V \to \mathbb{R}$ which is a linear extension of h and such that $\|\widetilde{h}\|_{V^*} = \|h\|_{W^*}$. Now set

$$f(x) = \widetilde{h}(x) - i\widetilde{h}(ix)$$

for all $x \in V$. Then, clearly, $f(x+y) = f(x) + f(y)$ and, for real scalars α, $f(\alpha x) = \alpha f(x)$. Now,

$$f(ix) = \widetilde{h}(ix) - i\widetilde{h}(-x) = i(\widetilde{h}(x) - i\widetilde{h}(ix)) = if(x)$$

and thus f is complex linear as well. Further, by the preceding proposition,

$$\|f\|_{V^*} = \|\widetilde{h}\|_{V^*} = \|h\|_{W^*} = \|g\|_{W^*}.$$

This completes the proof. ■

Corollary 3.1.1 *Let V be a normed linear space and $x_0 \in V$ a non-zero vector. Then, there exists $f \in V^*$ such that $\|f\| = 1$ and $f(x_0) = \|x_0\|$.*

Proof: Let W be the one-dimensional space spanned by x_0. Define $g(\alpha x_0) = \alpha\|x_0\|$. Then $\|g\|_{W^*} = 1$. Hence, there exists $f \in V^*$ such that $\|f\|_{V^*} = 1$ and which extends g. Hence $f(x_0) = g(x_0) = \|x_0\|$. ■

Remark 3.1.1 If V is a normed linear space and if x and y are distinct points in V, then, clearly, there exists $f \in V^*$ such that $f(x) \neq f(y)$ (consider $x_0 = x - y \neq \mathbf{0}$). We say that V^* separates points of V. ■

Corollary 3.1.2 *Let V be a normed linear space. Let $x \in V$. then*

$$\|x\| = \sup_{f\in V^*,\ \|f\|\leq 1} |f(x)| = \max_{f\in V^*,\ \|f\|\leq 1} |f(x)|. \tag{3.1.5}$$

Proof: Clearly, $|f(x)| \leq \|f\|\ \|x\| \leq \|x\|$ when $\|f\| \leq 1$. On the other hand, by the preceding corollary, there exists $f \in V^*$ such that $\|f\| = 1$ and $f(x) = \|x\|$ when x is non-zero. Thus the result is established for non-zero vectors and is trivially true for the null vector. ■

Compare the relation

$$\|f\| = \sup_{x\in V,\ \|x\|\leq 1} |f(x)|, \tag{3.1.6}$$

which is a *definition*, with the relation (3.1.5), which is a *result* of the theory. In the former, the supremum need not be attained, while in the latter the supremum is always attained and hence is a maximum.

This is the starting point for the investigation of a very nice property of Banach spaces called **reflexivity**.

Let $x \in V$ and define

$$J_x(f) = f(x)$$

for $f \in V^*$. Then, by virtue of (3.1.5), it follows that $J_x \in (V^*)^* = V^{**}$ and that, in fact,

$$\|J_x\|_{V^{**}} = \|x\|_V.$$

Thus $J : V \to V^{**}$ given by $x \mapsto J_x$ is a norm preserving linear transformation. Such a map is called an **isometry**. The map J is clearly injective and maps V isometrically onto a subspace of V^{**}.

Definition 3.1.1 *A Banach space V is said to be* **reflexive** *if the canonical imbedding $J : V \to V^{**}$, given above, is surjective.* ■

Since V^{**}, being a dual space, is always complete, the notion of reflexivity makes sense only for Banach spaces. By applying Corollary 3.1.2 to V^*, it is readily seen that the supremum in (3.1.6) is attained for reflexive Banach spaces. A deep result due to R. C. James is that the converse is also true: *if V is a Banach space such that the supremum is attained in (3.1.6) for all $f \in V^*$, then V is reflexive.* We will study reflexive spaces in greater detail in Chapter 5.

Example 3.1.1 Let $1 < p < \infty$. Let p^* be the conjugate exponent of p (cf. Definition 2.2.1). Let $y \in \ell_{p^*}$. We already saw that (cf. Example 2.3.4) the linear functional f_y defined on ℓ_p by

$$f_y(x) \;=\; \sum_{i=1}^{\infty} x_i y_i$$

for $x = (x_i) \in \ell_p$, is continuous and that, in fact,

$$\|f_y\| \;\leq\; \|y\|_{p^*}.$$

Now, let $f \in \ell_p^*$. Define

$$f_i \;=\; f(\mathsf{e}_i)$$

where e_i is the sequence whose i-th entry is unity and all other entries are zero. Set $\mathsf{f} = (f_i)$.

Let n be any positive integer. Define

$$x_i \;=\; \begin{cases} 0, & \text{if } 1 \leq i \leq n \text{ and } f_i = 0, \\ |f_i|^{p^*}/f_i, & \text{if } 1 \leq i \leq n \text{ and } f_i \neq 0, \\ 0, & \text{if } i > n. \end{cases}$$

Then, since it is a finite sequence, $x = (x_i) \in \ell_p$ and $x = \sum_{i=1}^{n} x_i \mathsf{e}_i$. Thus,

$$f(x) \;=\; \sum_{i=1}^{n} x_i f_i \;=\; \sum_{i=1}^{n} |f_i|^{p^*}.$$

Consequently

$$\sum_{i=1}^{n} |f_i|^{p^*} \;\leq\; \|f\|\|x\|_p \;=\; \|f\| \sum_{i=1}^{n} \left(|f_i|^{p^*}\right)^{\frac{1}{p}}$$

using the definition of the x_i and that of p^*. This yields

$$\left(\sum_{i=1}^{n} |f_i|^{p^*}\right)^{\frac{1}{p^*}} \;\leq\; \|f\|.$$

Since n was arbitrary, we deduce that $\mathsf{f} \in \ell_{p^*}$ and that $\|\mathsf{f}\|_{p^*} \leq \|f\|$.

For any $x = (x_i) \in \ell_p$, we have $\sum_{i=1}^{n} x_i \mathsf{e}_i \to x$ in ℓ_p and so, by the continuity of f, it follows that

$$f(x) \;=\; \sum_{i=1}^{\infty} x_i f_i$$

or, in other words, $f = f_{\mathsf{f}}$. Hence,

$$\|\mathsf{f}\|_{p^*} \;\le\; \|f\| \;\le\; \|\mathsf{f}\|_{p^*}$$

as already observed.

Thus every element of the dual space of ℓ_p occurs in this fashion and the map $y \mapsto f_y$ is an isometry of ℓ_{p^*} onto ℓ_p^*. Thus, we can write

$$\ell_p^* \;=\; \ell_{p^*}$$

using this isometry.

Similarly, we can write

$$\ell_{p^*}^* \;=\; \ell_p.$$

It is easy to see that using these identifications of the dual spaces, the canonical isometry from ℓ_p into $\ell_p^{**} = \ell_p$ is nothing but the identity map, which is onto. Thus, the spaces ℓ_p, for $1 < p < \infty$ are all reflexive. ■

Remark 3.1.2 Though in the above example, we have not dealt with the case of real and complex sequence spaces separately, it is customary to identify the dual of the complex sequence space ℓ_p with ℓ_{p^*} via the following relation: if $y \in \ell_{p^*}$, then we define $f_y \in \ell_p^*$ by

$$f_y(x) \;=\; \sum_{i=1}^{\infty} x_i \overline{y_i}.$$

The mapping $y \in \ell_{p^*} \mapsto f_y \in \ell_p^*$ satisfies $f_{y+z} = f_y + f_z$ and $f_{\alpha y} = \overline{\alpha} f_y$. Such a mapping is called *conjugate linear.* This identification will be especially useful in the case of ℓ_2 which will be made clear when we study *Hilbert spaces* (cf. Chapter 7). ■

Example 3.1.2 Proceeding in a manner similar to that in the preceding example, we can show that $\ell_1^* = \ell_\infty$. In other words, given any continuous linear functional f on ℓ_1, we have

$$f(x) \;=\; \sum_{i=1}^{\infty} x_i f_i$$

for all $x = (x_i) \in \ell_1$, where $f_i = f(\mathsf{e}_i)$. Further, setting $\mathsf{f} = (f_i)$, we have

$$\|\mathsf{f}\|_\infty \;=\; \|f\|.$$

We now show that there is no similar identification of ℓ_∞^* with ℓ_1. Of course, if $y \in \ell_1$, the functional f_y defined by

$$f_y(x) = \sum_{i=1}^{\infty} x_i y_i$$

for all $x = (x_i) \in \ell_\infty$ is continuous and $\|f_y\| = \|y\|_1$. But there exist continuous linear functionals on ℓ_∞ which do not arise this way. So the identity mapping of ℓ_1 which is still the canonical imbedding of ℓ_1 into ℓ_1^{**}, is not surjective. Thus the space ℓ_1 is not reflexive.

To see this, let G be the subspace of all convergent sequences in ℓ_∞. For $x = (x_i) \in G$, define

$$g(x) = \lim_{i \to \infty} x_i.$$

Then $g : G \to \mathbb{R}$ is linear and $|g(x)| \leq \|x\|_\infty$. Thus, g is continuous as well and so, by the Hahn-Banach theorem, can be extended to a continuous linear functional f on ℓ_∞, preserving the norm. We claim that this continuous linear functional cannot be obtained from an element of ℓ_1 by the above outlined procedure.

Assume the contrary and let $y = (y_i) \in \ell_1$ be such that $f = f_y$. Consider the sequence $\{x^{(n)}\}$ in ℓ_∞ given by

$$x^{(n)} = \{0, 0, ..., 0, 1, 1, 1, ...\}$$

where the 1's start from the n-th entry. Then $\|x^{(n)}\|_\infty = 1$ and $x^{(n)} \in G$. We have

$$1 = f(x^{(n)}) = \sum_{i=n}^{\infty} y_i$$

which is impossible since $y \in \ell_1$ implies that

$$\sum_{i=n}^{\infty} |y_i| \to 0$$

as $n \to \infty$. ■

Remark 3.1.3 Since the canonical map $J : V \to V^{**}$ is an isometry, it is injective. Thus, if V is finite dimensional, then $\dim(V^{**}) = \dim(V^*) = \dim(V)$ and so, by dimension considerations, J has to be surjective as well. Thus, every finite dimensional space is automatically reflexive. In particular, for any positive integer N, we have that $(\ell_\infty^N)^* = \ell_1^N$, unlike what happens in the case of sequence spaces, as seen above. The interested reader can try to prove this directly. ■

3.2 Geometric Versions

In this section we will study the geometric versions of the Hahn-Banach theorem which concerns the separation of convex sets by means of hyperplanes.

Definition 3.2.1 *Let V be a real normed linear space. An* **affine hyperplane** *is a set of the form*

$$H = \{x \in V \mid f(x) = \alpha\},$$

denoted by $[f = \alpha]$, where f is a non-zero linear functional on V. ■

Proposition 3.2.1 *A hyperplane $[f = \alpha]$ is closed if, and only if, f is a continuous linear functional.*

Proof: Clearly, if f is continuous, then $[f = \alpha]$ is closed. Conversely, assume that the hyperplane H, given by $[f = \alpha]$, is closed. Then, its complement H^c is open and, since $f \not\equiv 0$, it is non-empty.(For, if $\alpha \neq 0$, then $\mathbf{0} \in H^c$; if $\alpha = 0$, and $f \not\equiv 0$, there exists $x \in V$ such that $f(x) \neq 0$ and so $x \in H^c$.)

Without loss of generality, assume that $x_0 \in H^c$ is such that $f(x_0) < \alpha$. Since H^c is open, there exists $r > 0$ such that the open ball centered at x_0 and of radius $2r$, denoted $B(x_0; 2r)$, is contained in H^c. Now, for all $x \in B(x_0; 2r)$, we have $f(x) < \alpha$. (If not, there exists $x_1 \in B(x_0; 2r)$ such that $f(x_1) > \alpha$. Let

$$t = \frac{f(x_0) - \alpha}{f(x_0) - f(x_1)}.$$

Then $0 < t < 1$ and, if $x_t = tx_1 + (1-t)x_0$, then $f(x_t) = \alpha$. But $B(x_0; 2r)$ is a convex set and so $x_t \in B(x_0; 2r)$, which is a contradiction.)

Thus, for any $z \in V$ such that $\|z\| \leq 1$, we get $f(x_0 + rz) \leq \alpha$ or, equivalently,

$$f(z) \leq \frac{\alpha - f(x_0)}{r}.$$

Thus the image of the unit ball is bounded and so f is continuous. ■

Proposition 3.2.2 *Let C be an open and convex set in a real normed linear space V such that $\mathbf{0} \in C$. For $x \in V$, set*

$$p(x) = \inf\{\alpha > 0 \mid \alpha^{-1}x \in C\}.$$

(The function p is called the Minkowski functional *of C.) Then, there exists $M > 0$ such that*

$$0 \le p(x) \le M\|x\| \tag{3.2.1}$$

for all $x \in V$. We also have

$$C \;=\; \{x \in V \mid p(x) < 1\}. \tag{3.2.2}$$

Further, p satisfies (3.1.1).

Proof: Since $\mathbf{0} \in C$ and C is open, there exists an open ball $B(\mathbf{0}; 2r)$, centered at $\mathbf{0}$ and of radius $2r$, contained in C. Now, if $x \in V$, we have $rx/\|x\| \in C$ and so, by definition, $p(x) \le \frac{1}{r}\|x\|$ which proves (3.2.1).

Let $x \in C$. Since C is open, and since $\mathbf{0} \in C$, there exists $\varepsilon > 0$ such that $(1+\varepsilon)x \in C$. Thus, $p(x) \le (1+\varepsilon)^{-1} < 1$. Conversely, let $x \in V$ such that $p(x) < 1$. Then, there exists $0 < t < 1$ such that $\frac{1}{t}x \in C$. Then, as C is convex, we also have $t\frac{1}{t}x + (1-t)\mathbf{0} \in C$, *i.e.* $x \in C$. This proves (3.2.2).

If $\alpha > 0$, it is easy to see that $p(\alpha x) = \alpha p(x)$. This is the first relation in (3.1.1). Now, let x and $y \in V$. Let $\varepsilon > 0$. Then

$$\frac{1}{p(x)+\varepsilon}x \in C \quad\text{and}\quad \frac{1}{p(y)+\varepsilon}y \in C.$$

Set

$$t \;=\; \frac{p(x)+\varepsilon}{p(x)+p(y)+2\varepsilon}$$

so that $0 < t < 1$. Then, as C is convex,

$$t\frac{1}{p(x)+\varepsilon}x + (1-t)\frac{1}{p(y)+\varepsilon}y \;=\; \frac{1}{p(x)+p(y)+2\varepsilon}(x+y) \in C$$

which implies that

$$p(x+y) \;\le\; p(x)+p(y)+2\varepsilon$$

from which the second relation in (3.1.1) follows since ε was chosen arbitrarily. ■

Proposition 3.2.3 *Let C be a non-empty open convex set in a real normed linear space V and assume that $x_0 \notin C$. Then, there exists $f \in V^*$ such that $f(x) < f(x_0)$ for all $x \in C$.*

Proof: Without loss of generality, we can assume that $\mathbf{0} \in C$. (If $\mathbf{0} \notin C$, let $x_1 \in C$. Then we consider the convex set $C - \{x_1\}$ which contains the origin and does not contain $x_0 - x_1$; if f is as in the proposition, we have $f(x - x_1) < f(x_0 - x_1)$ for all $x \in C$ which yields $f(x) < f(x_0)$ for all $x \in C$.)

Let W be the one-dimensional space spanned by x_0. Define $g : W \to \mathbb{R}$ by

$$g(tx_0) \;=\; t.$$

By definition of the Minkowski functional, since $\frac{1}{t}tx_0 = x_0 \notin C$, we have that

$$g(tx_0) \;=\; t \;\leq p(tx_0)$$

for $t > 0$. Since the Minkowski functional is non-negative, this inequality holds trivially for $t \leq 0$ as well. Thus, by the Hahn-Banach theorem (cf. Theorem 3.1.1), there exists a linear extension f of g to the whole of V such that, for all $x \in V$,

$$f(x) \;\leq\; p(x) \;\leq\; M\|x\|$$

(cf. (3.2.1)) which yields $|f(x)| \leq M\|x\|$, and so f is continuous as well. Now, if $x \in C$,

$$f(x) \;\leq\; p(x) \;<\; 1 \;=\; g(x_0) \;=\; f(x_0)$$

by (3.2.2) and this completes the proof. ■

Theorem 3.2.1 (**Hahn-Banach Theorem**) *Let A and B be two non-empty and disjoint convex subsets of a real normed linear space V. Assume that A is open. Then, there exists a closed hyperplane which separates A and B, i.e. there exists $f \in V^*$ and $\alpha \in \mathbb{R}$ such that*

$$f(x) \;\leq\; \alpha \;\leq\; f(y)$$

for all $x \in A$ and $y \in B$.

Proof: Let $C = A - B = \{x - y \mid x \in A, y \in B\}$. Since

$$C \;=\; \cup_{y \in B}(A - \{y\}),$$

we see immediately that C is both open and convex. Since A and B are disjoint, it also follows that $\mathbf{0} \notin C$. Hence, by the preceding proposition, there exists $f \in V^*$ such that $f(z) < 0$ for all $z \in C = A - B$. In other words, $f(x) < f(y)$ for all $x \in A$ and $y \in B$. Choose $\alpha \in \mathbb{R}$ such that

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

This completes the proof. ■

Theorem 3.2.2 (**Hahn-Banach Theorem**) *Let A and B be non-empty and disjoint convex sets in a real normed linear space V. Assume that A is closed and that B is compact. Then A and B can be separated* strictly *by a closed hyperplane, i.e. there exists $f \in V^*$, $\alpha \in \mathbb{R}$ and $\varepsilon > 0$ such that*

$$f(x) \ \leq \ \alpha - \varepsilon \ \text{ and } \ f(y) \ \geq \ \alpha + \varepsilon$$

for all $x \in A$ and $y \in B$.

Proof: Let $\eta > 0$. Then, if $B(\mathbf{0};\eta)$ is the open ball of radius $\eta > 0$ centered at $\mathbf{0}$, then $A+B(\mathbf{0};\eta)$ and $B+B(\mathbf{0};\eta)$ are non-empty, open and convex. Further, if $\eta > 0$ is sufficiently small, the two sets are disjoint as well. If not, there exists a sequence $\eta_n \to 0$ and $x_n \in A, y_n \in B$ such that $\|x_n - y_n\| \leq 2\eta_n$. Since B is compact, there exists a subsequence y_{n_k} which converges to $y \in B$. This implies then that $x_{n_k} \to y$ and, since A is closed, $y \in A$, *i.e.* $y \in A \cap B$, which is a contradiction.

Thus, we can choose $\eta > 0$ such that $A + B(\mathbf{0};\eta)$ and $B + B(\mathbf{0};\eta)$ are disjoint. Then, by the preceding theorem, there exists $f \in V^*$ and $\alpha \in \mathbb{R}$ such that for all $x \in A, y \in B$ and z_1 and z_2 in the closed unit ball, we have

$$f\left(x + \frac{\eta}{2}z_1\right) \ \leq \ \alpha \ \leq \ f\left(y + \frac{\eta}{2}z_2\right).$$

This implies that

$$f(x) + \frac{\eta}{2}\|f\| \ \leq \ \alpha \ \leq \ f(y) - \frac{\eta}{2}\|f\|.$$

This proves the result if we set $\varepsilon = \frac{\eta}{2}\|f\|$. ■

The following corollary is very useful in testing whether a given subspace of a normed linear space is dense or not.

Corollary 3.2.1 *Let W be a subspace of a real normed linear space V. Assume that $\overline{W} \neq V$. Then, there exists $f \in V^*$ such that $f \not\equiv 0$ and such that $f(x) = 0$ for all $x \in W$.*

Proof: Let $x_0 \in V \backslash \overline{W}$. Let $A = \overline{W}$ and $B = \{x_0\}$. Then A is closed, B is compact and they are non-empty and disjoint convex sets. Thus, there exists $f \in V^*$ and $\alpha \in \mathbb{R}$ such that for all $x \in \overline{W}$,

$$f(x) < \alpha < f(x_0).$$

Since W is a linear subspace, it follows that for all $\lambda \in \mathbb{R}$, we have $\lambda f(x) < \alpha$ for all $x \in W$. Now, since $\mathbf{0} \in W$, we have $\alpha > 0$. On the other hand, setting $\lambda = n$, we get that, for any $x \in W$,

$$f(x) < \frac{\alpha}{n}$$

whence we see that $f(x) \leq 0$ for all $x \in W$. Again, if $x \in W$, we also have $-x \in W$ and so $f(-x) \leq 0$ as well and so $f(x) = 0$ for all $x \in W$ and $f(x_0) > \alpha > 0$. ■

Remark 3.2.1 In case of normed linear spaces over $\mathbb{C}$, the conclusions of Theorems 3.2.1 and 3.2.2 hold with f being replaced by $\mathrm{Re}(f)$, the real part of f. This follows from Proposition 3.1.1. It is now easy to see that Corollary 3.2.1 is also valid for complex spaces. For another proof of this result, see Exercise 3.4. ■

Remark 3.2.2 A topological vector space is said to be *locally convex* if every point has a local basis made up of convex sets, *i.e.* every open neighbourhood of each point contains a convex open neighbourhood of that point.

The proofs of the geometric versions of the Hahn-Banach theorems go through *mutatis mutandis* in the case of locally convex spaces. In particular, Corollary 3.2.1 is also true for such spaces. For details, see Rudin [8]. ■

3.3 Vector Valued Integration

In this section we will apply the Hahn-Banach theorem to give a meaning to integration of vector valued functions.

Let us consider the unit interval $[0,1]$ endowed with the Lebesgue measure. Let V be a normed linear space over $\mathbb{R}$. Let $\varphi : [0,1] \to V$ be a continuous mapping. We would like to give a meaning to the integral

$$\int_0^1 \varphi(t)\ dt$$

as a vector in V in a manner that the familier properties of integrals are preserved.

Using our experience with the integral of a continuous real valued function, one could introduce a partition

$$0 = x_0 < x_1 < \cdots < x_n = 1$$

and form Riemann sums of the form

$$\sum_{i=1}^{n}(x_i - x_{i-1})\varphi(\xi_i)$$

where $\xi_i \in [x_{i-1}, x_i]$ for $1 \leq i \leq n$, and define the integral (if it exists) as a suitable limit of such sums. Assume that such a limit exists and denote it by $y \in V$. Let $f \in V^*$. Then, by the continuity and linearity of f, it will follow that $f(y)$ will be the limit of the Riemann sums of the form

$$\sum_{i=1}^{n}(x_i - x_{i-1})f(\varphi(\xi_i)).$$

But since $f \circ \varphi : [0,1] \to \mathbb{R}$ is continuous, the above limit of Riemann sums is none other than

$$\int_0^1 f(\varphi(t))\ dt.$$

Thus the integral of φ must satisfy the relation

$$f\left(\int_0^1 \varphi(t)\ dt\right) = \int_0^1 f(\varphi(t))\ dt \tag{3.3.1}$$

for all $f \in V^*$.

Notice that since V^* separates points of V (cf. Corollary 3.1.1 and Remark 3.1.1), such a vector, if it exists, must be unique. We use this to define the integral of a vector valued function.

Definition 3.3.1 *Let V be a real normed linear space and let $\varphi : [0,1] \to \mathbb{R}$ be a continuous mapping. The integral of φ over $[0,1]$, denoted*

$$\int_0^1 \varphi(t)\ dt,$$

is that vector in V which satisfies (3.3.1) for all $f \in V^*$. ■

Proposition 3.3.1 *Let $\varphi : [0,1] \to V$ be a continuous mapping into a real normed linear space V. Then the integral of φ over $[0,1]$ exists.*

Proof: Since $[0,1]$ is compact, the set $\overline{H}$ which is the closure (in V) of the set H which is the convex hull of $\varphi([0,1])$ (*i.e.* the smallest convex set containing $\varphi([0,1])$), is compact.

Let L be an arbitrary finite collection of continuous linear functionals on V. Define

$$E_L = \left\{ y \in \overline{H} \mid f(y) = \int_0^1 f(\varphi(t))\, dt \ \text{ for all } \ f \in L \right\}.$$

It is immediate to see that E_L is a closed set.

Step 1: For any such finite collection L of continuous linear functionals, $E_L \neq \emptyset$. To see this, let $L = \{f_1, \cdots, f_k\}$. Define $\mathcal{A} : V \to \mathbb{R}^k$ by

$$\mathcal{A}(x) = (f_1(x), \cdots, f_k(x)).$$

Then $\mathcal{A}$ is a continuous linear transformation and so $K = \mathcal{A}(\overline{H})$ is a compact and convex set. If $(t_1, \cdots, t_k) \notin K$, then, by the Hahn-Banach theorem (cf. Theorem 3.2.2), we can find constants $c_1, \cdots, c_k$ such that

$$\sum_{i=1}^{k} c_i u_i < \sum_{i=1}^{k} c_i t_i$$

for all $(u_1, \cdots, u_k) \in K$. In particular, for all $t \in [0,1]$, we have

$$\sum_{i=1}^{k} c_i f_i(\varphi(t)) < \sum_{i=1}^{k} c_i t_i.$$

Integrating this inequality over $[0,1]$, we get

$$\sum_{i=1}^{k} c_i m_i < \sum_{i=1}^{k} c_i t_i$$

where

$$m_i = \int_0^1 f_i(\varphi(t))\, dt.$$

In other words, if $(t_1, \cdots, t_k) \notin K$, then $(t_1, \cdots, t_k) \neq (m_1, \cdots, m_k)$. Thus, $(m_1, \cdots, m_k) \in K$. Thus, there exists $y \in \overline{H}$ such that, for $1 \leq i \leq k$, we have

$$m_i = f_i(y).$$

This means that $y \in E_L$, *i.e.* E_L is non-empty.

Step 2. Let I be a finite indexing set and let L_i be finite collections of elements in V^* for each $i \in I$. Then $L = \cup_{i\in I} L_i$ is still finite and further, since it is easy to see that

$$\cap_{i\in I} E_{L_i} = E_L,$$

it follows from the previous step that the class of closed sets

$$\{E_L \mid L \text{ a finite subset of } V^*\}$$

has finite intersection property. Since $\overline{H}$ is compact, it now follows that

$$\cap_{L,\ \text{finite subset of } V^*} E_L \neq \emptyset.$$

In particular, there exists y such that $y \in E_{\{f\}}$ for every $f \in V^*$, *i.e.* y satisfies

$$f(y) = \int_0^1 f(\varphi(t))\, dt$$

for every $f \in V^*$. Thus $y = \int_0^1 \varphi(t)\, dt$. This completes the proof. ■

Proposition 3.3.2 *Let V be a real normed linear space and let $\varphi : [0,1] \to V$ be continuous. Then*

$$\left\| \int_0^1 \varphi(t)\, dt \right\| \leq \int_0^1 \|\varphi(t)\|\, dt. \tag{3.3.2}$$

Proof: By Corollary 3.1.1, there exists $f \in V^*$ such that $\|f\| = 1$ and $f(y) = \|y\|$ where

$$y = \int_0^1 \varphi(t)\, dt.$$

Thus,

$$\begin{aligned} \left\| \int_0^1 \varphi(t)\, dt \right\| &= f\left(\int_0^1 \varphi(t)\, dt \right) = \int_0^1 f(\varphi(t))\, dt \\ &\leq \int_0^1 |f(\varphi(t))|\, dt \leq \int_0^1 \|\varphi(t)\|\, dt. \end{aligned}$$

This completes the proof. ■

Remark 3.3.1 Let $\varphi : [a,b] \to V$ be a continuous mapping. Define $\psi : [0,1] \to V$ by $\psi(t) = \varphi(a + t(b-a))$. Then we can define

$$\int_a^b \varphi(t)\, dt = (b-a) \int_0^1 \psi(t)\, dt.$$

It is easy to verify that for all $f \in V^*$, we have

$$f\left(\int_a^b \varphi(t)\ dt\right) = \int_a^b f(\varphi(t))\ dt.$$

Again the result of Proposition 3.3.2 remains valid in this case as well.

Assume now that $\varphi : [0, \infty) \to V$ is continuous. Assume further that the limit

$$\lim_{\lambda\to\infty} \int_0^\lambda \varphi(t)\ dt$$

exists, *i.e.*, for any sequence $\lambda_n \to \infty$ (as $n \to \infty$), we have that the limit

$$\lim_{n\to\infty} \int_0^{\lambda_n} \varphi(t)\ dt$$

exists and is independent of the sequence chosen. Then we define

$$\int_0^\infty \varphi(t)\ dt = \lim_{\lambda\to\infty} \int_0^\lambda \varphi(t)\ dt$$

and again it follows that, for all $f \in V^*$, we have

$$f\left(\int_0^\infty \varphi(t)\ dt\right) = \int_0^\infty f(\varphi(t))\ dt.$$

The result of Proposition 3.3.2 continues to hold. We can define integrals over other infinite intervals, if they exist, in a similar manner. ■

3.4 An Application to Optimization Theory

We conclude this chapter with an application of the Hahn-Banach theorem to optimization theory.

Definition 3.4.1 *A* **cone** *in a real vector space V is a set C such that:*
(i) $\mathbf{0} \in C$;
(ii) if $x \in C$ and $\lambda \geq 0$, then $\lambda x \in C$. ■

Lemma 3.4.1 *Let v_i, $1 \leq i \leq n$ be elements in a normed linear space V. Define*

$$C = \left\{\sum_{i=1}^n \lambda_i v_i \mid \lambda_i \geq 0,\ 1 \leq i \leq n\right\}.$$

Then C is a closed convex cone.

Proof: Step 1. Clearly, C is convex since, for any $0 < t < 1$, we have

$$t\sum_{i=1}^{n}\lambda_i v_i + (1-t)\sum_{i=1}^{n}\mu_i v_i \;=\; \sum_{I=1}^{n}(t\lambda_i + (1-t)\mu_i)v_i$$

and so, if $\lambda_i \geq 0$ and $\mu_i \geq 0$ for all i, we also have $t\lambda_i + (1-t)\mu_i \geq 0$. Also if $x \in C$, it is obvious that $\lambda x \in C$ for any $\lambda \geq 0$. Thus, C is a convex cone.

Step 2. Assume that the v_i are linearly independent. Let $\{w_n\}$ be a sequence in C and assume that $w_m \to w$ in V. If

$$W \;=\; \text{span}\{v_1, \cdots, v_n\},$$

then W is a finite dimensional subspace of V and is hence closed as well. Thus $w \in W$. If $w_m = \sum_{i=1}^{n}\lambda_i^m v_i$ with $\lambda_i^m \geq 0$ for all $1 \leq i \leq n$ and for all m, then $\lambda_i^m \to \lambda_i \geq 0$ for each $1 \leq i \leq n$ and $w = \sum_{i=1}^{n}\lambda_i v_i$. Thus $w \in C$ and so C is closed.

Step 3. Assume now that the v_i are linearly dependent. Then there exists a linear relation between them and so we can find scalars α_i such that $\sum_{i=1}^{n}\alpha_i v_i = \mathbf{0}$ and such that the set

$$J \;=\; \{i \mid 1 \leq i \leq n, \alpha_i < 0\}$$

is non-empty.

Let $v \in C$ be such that $v = \sum_{i=1}^{n}\lambda_i v_i$ with $\lambda_i \geq 0$ for all $1 \leq i \leq n$. Then $v = \sum_{i=1}^{n}(\lambda_i + t\alpha_i)v_i$ for any $t \in \mathbb{R}$. Define

$$t \;=\; \min_{i \in J}\left\{-\frac{\lambda_i}{\alpha_i}\right\} \;\geq\; 0.$$

Then, for all $1 \leq i \leq n$, we have $\lambda_i + t\alpha_i \geq 0$ and at least one of them must vanish. Then

$$C \;=\; \cup_{j\in I}\left\{v = \sum_{i\in I\backslash\{j\}}\lambda_i v_i \mid \lambda_i \geq 0 \;\text{ for all }\; i \in I\backslash\{j\}\right\}$$

where $I = \{1, 2, \cdots, n\}$. Each set in the union described on the right-hand side is a cone but generated by fewer elements from V. Iterating this procedure, we can ultimately write C as the finite union of cones each generated by a linearly independent set of vectors and hence, by the preceding step, each of these cones will be closed as well. Hence C, being the finite union of closed sets, is closed. This completes the proof. ■

Theorem 3.4.1 (**Farkas-Minkowski Lemma**) *Let V be a real reflexive Banach space and let $\{f_0, f_1, \cdots, f_n\}$ be elements of V^* such that if for some $x \in V$ we have $f_i(x) \geq 0$ for all $1 \leq i \leq n$, then $f_0(x) \geq 0$ as well. Then, there exists scalars $\lambda_i \geq 0$, $1 \leq i \leq n$ such that*

$$f_0 = \sum_{i=1}^{n} \lambda_i f_i.$$

Proof: Let

$$C = \left\{ \sum_{i=1}^{n} \lambda_i f_i \mid \lambda_i \geq 0, \ 1 \leq i \leq n \right\}$$

which is a closed convex cone in V^* by the preceding lemma. Assume that $f_0 \notin C$. Then, by the Hahn-Banach Theorem (cf. Theorem 3.2.2) there exist $\varphi \in V^{**}$ and $\alpha \in \mathbb{R}$ such that

$$\varphi(f_0) < \alpha < \varphi(f)$$

for all $f \in C$. Since $\mathbf{0} \in C$, it follows that $\alpha < 0$. Thus $\varphi(f_0) < 0$ as well.

Now, since V is reflexive, there exists $x \in V$ such that $\varphi = J_x$ and so $f_0(x) < 0$. On the other hand, since C is a cone, for all $\lambda > 0$, and for all $f \in C$, we have $\lambda f \in C$ and so $\varphi(\lambda f) > \alpha$ or, $\varphi(f) > \alpha/\lambda$ whence we deduce, on letting λ tend to infinity, that $\varphi(f) \geq 0$, *i.e.* $f(x) \geq 0$ for all $f \in C$. In particular $f_i(x) \geq 0$ for all $1 \leq i \leq n$ while $f_0(x) < 0$, which is a contradiction. Thus $f_0 \in C$ and the proof is complete. ■

The Farkas-Minkowski lemma is a key step in the proof of the *Kuhn-Tucker conditions* which play the same role in characterizing minima in the presence of constraints in the form of *inequalities* as that played by *Lagrange multipliers* in characterizing minima in the presence of constraints in the form of *equalities.* While the Kuhn-Tucker conditions are necessary in general 'nonlinear programming', they are necessary and sufficient in 'convex programming' *i.e.* when the functional to be minimized and the constraints are all given by convex functions.

Let V be a real normed linear space and let $J : V \to \mathbb{R}$ be a given functional. Let $K \subset V$ be a closed and convex subset. Then, if J attains a minimum over K at $u \in K$ and if J is differentiable at u, a necessary condition is that

$$J'(u)(v-u) \ \geq \ 0$$

for all $v \in K$ (cf. Exercise 2.44). We would like to generalize this to sets K which are not necessarily convex. To this end, we introduce the following definition.

Definition 3.4.2 *Let V be a real normed linear space and let $U \subset V$ be a non-empty subset. Let $u \in U$. Then, the* **tangent cone,** *denoted $C(u)$, at u is the union of the origin and the set of all vectors $w \in V$ such that*
(i) *there exists a sequence $\{u_k\}$ in U, $u_k \neq u$ for all $k \in \mathbb{N}$ and $\lim_{k\to\infty} u_k = u$;*
(ii)

$$\lim_{k\to\infty} \frac{u_k - u}{\|u_k - u\|} \ = \ \frac{w}{\|w\|}. \ \blacksquare$$

Remark 3.4.1 The second condition in the above definition may be written, in an equivalent fashion, as follows:

$$u_k \ = \ u + \|u_k - u\| \left(\frac{w}{\|w\|} + \delta_k \right)$$

where $\delta_k \to \mathbf{0}$ as $k \to \infty$. ■

Remark 3.4.2 It is clear that $C(u)$ is a cone (cf. Definition 3.4.1) since $w \in C(u)$ implies that $\lambda w \in C(u)$ as well, for any $\lambda > 0$ (with the same associated sequence $\{u_k\}$). This is a cone (which is not necessarily convex) with its vertex at the origin. Its translate

$$u + C(u) \ = \ \{u + w \mid w \in C(u)\},$$

is a cone with vertex at u. This translated cone contains the (half) tangents at u of all curves in U which pass through u, as it is easy to see from the Taylor expansion (at u) of the function describing the curve. ■

Proposition 3.4.1 *Let U be a non-empty subset of a real normed linear space V and let $u \in U$. Then, $C(u)$ is a closed cone.*

Proof: Let $w_n \in C(u)$ and let $w_n \to w$ in V. Without loss of generality, we can assume that $w \neq \mathbf{0}$ (since the origin is always in $C(u)$, by definition) and hence that $w_n \neq \mathbf{0}$ for all n. There exist $u_k^n \in U$ such that $u_k^n \neq u$ for all n and such that $u_k^n \to u$ as $k \to \infty$. Further,

$$u_k^n = u + \|u_k^n - u\| \left(\frac{w_n}{\|w_n\|} + \delta_k^n \right)$$

where $\delta_k^n \to \mathbf{0}$ as $k \to \infty$.

Choose a sequence $\{\varepsilon_n\}$ of positive reals which converges to zero as $n \to \infty$. Then, we can find positive integers $k(n)$ such that

$$\|u_{k(n)}^n - u\| < \varepsilon_n, \text{ and } \|\delta_{k(n)}^n\| \leq \varepsilon_n.$$

Consider the sequence $\{u_{k(n)}^n\}$. Clearly, $u_{k(n)}^n \to u$ and $u_{k(n)}^n \neq u$ for all n. Further,

$$u_{k(n)}^n = u + \|u_{k(n)}^n - u\| \left[\frac{w}{\|w\|} + \left(\delta_{k(n)}^n + \frac{w_n}{\|w_n\|} - \frac{w}{\|w\|} \right) \right]$$

which shows that $w \in C(u)$ since

$$\eta_n = \delta_{k(n)}^n + \left(\frac{w_n}{\|w_n\|} - \frac{w}{\|w\|} \right) \to \mathbf{0}$$

given that $w_n \to w$. This completes the proof. ∎

Proposition 3.4.2 *Let $J : U \subset V \to \mathbb{R}$ be a functional defined on a set U of a real normed linear space V. Assume that J attains a relative minimum at $u \in U$ and that J is differentiable at u. Then*

$$J'(u)(v - u) \geq 0$$

for all $v \in u + C(u)$.

Proof: Let $v \in u + C(u)$. Then $w = v - u \in C(u)$. Let $\{u_k\}$ be a sequence in U associated to w as in the definition of $C(u)$. Then, since J is differentiable at u, we have

$$\begin{aligned} J(u_k) - J(u) &= J'(u)(u_k - u) + \|u_k - u\|\varepsilon_k \\ &= \|u_k - u\| \left(\tfrac{1}{\|w\|} J'(u)w + J'(u)\delta_k + \varepsilon_k \right) \end{aligned}$$

where $\delta_k \to \mathbf{0}$ and $\varepsilon_k \to 0$ as $k \to \infty$. Setting $\eta_k = \|w\|(J'(u)\delta_k + \varepsilon_k)$, we see that $\eta_k \to 0$ as well. Since J attains a relative minimum at u, it follows that

$$0 \leq J(u_k) - J(u) = \frac{\|u_k - u\|}{\|w\|}(J'(u)w + \eta_k)$$

from which it immediately follows that $J'(u)w \geq 0$, which completes the proof. ■

The above result can be used to derive optimality conditions when a functional J is being minimized under constraints. We consider a finite set of functionals $\varphi_i : V \to \mathbb{R}$ for $1 \leq i \leq m$ and set

$$U \;=\; \{v \in V \mid \varphi_i(u) \leq 0,\ 1 \leq i \leq m\}. \tag{3.4.1}$$

Of particular interest is the case when the functionals φ_i are affine linear, *i.e.* there exist $f_i \in V^*$ and vectors $d_i \in V$ for $1 \leq i \leq m$ such that

$$\varphi_i(u) \;=\; f_i(u) + d_i \tag{3.4.2}$$

for $1 \leq i \leq m$. In this case, we have a simple characterization of the tangent cone.

Proposition 3.4.3 *Let U be as given by (3.4.1) and let the constraints φ_i be affine linear, given by (3.4.2). Then, for any $u \in U$,*

$$C(u) \;=\; \{w \in V \mid f_i(w) \leq 0,\ i \in I(u)\} \tag{3.4.3}$$

where

$$I(u) \;=\; \{i \mid 1 \leq i \leq m,\ \varphi_i(u) = 0\}.$$

Proof: Notice that (cf. Exercise 2.40) $\varphi_i'(u) = f_i$. If $i \in I(u)$, then φ_i attains its maximum over U at u. Then, by Proposition 3.4.2, $\varphi_i'(u)(w) \leq 0$ for all $w \in C(u)$.

Conversely, if $w \neq \mathbf{0}$ satisfies $f_i(w) \leq 0$ for all $i \in I(u)$, set $u_k = u + \varepsilon_k w$ where $\{\varepsilon_k\}$ is a sequence of positive reals converging to zero. then $u_k \neq u$ and $u_k \to u$. If $i \notin I(u)$, then $\varphi_i(u) < 0$, and so, by continuity, $\varphi_i(u_k) < 0$ for large enough k. If $i \in I(u)$, then $\varphi(u) = 0$ and so

$$\varphi_i(u_k) \;=\; f_i(u_k - u) \;=\; \varepsilon_k f_i(w) \;\leq\; 0.$$

Thus, for k large enough, we have that $u_k \in U$. Finally we see immediately that

$$\frac{u_k - u}{\|u_k - u\|} \;=\; \frac{w}{\|w\|}.$$

Thus, it follows that $w \in C(u)$. This completes the proof. ■

Theorem 3.4.2 (**Kuhn-Tucker Conditions**) *Let V be a real, reflexive Banach space. Let φ_i, $1 \leq i \leq m$ be as in (3.4.2) and let U be as in (3.4.1). Let $J : V \to \mathbb{R}$ be a functional which attains a relative minimum at $u \in U$. Assume that J is differentiable at u. Then, there exist constants $\lambda_i(u)$ such that*

$$\left.\begin{array}{rcl} J'(u) + \sum_{i=1}^m \lambda_i(u)\varphi_i'(u) & = & \mathbf{0} \\ \sum_{i=1}^m \lambda_i(u)\varphi_i(u) & = & 0 \\ \lambda_i(u) & \geq & 0, \ 1 \leq i \leq m. \end{array}\right\} \quad (3.4.4)$$

Proof: By Propositions 3.4.2 and 3.4.3, we have that for all w such that $\varphi_i'(u)w \leq 0$, $i \in I(u)$, we have $J'(u)w \geq 0$. Thus, by the Farkas-Minkowski lemma, there exist $\lambda_i(u) \geq 0$ for $i \in I(u)$ such that

$$J'(u) \; = \; -\sum_{i \in I(u)} \lambda_i(u)\varphi_i'(u).$$

Setting $\lambda_i(u) = 0$ for all $i \notin I(u)$, we get (3.4.4). This completes the proof. ∎

The above theorem can be generalized to cases when the φ_i are not affine. In this situation, in addition to differentiability at u, we need to assume another technical condition of 'admissibility' on the constraints at u. In particular, when the constraints φ_i, $1 \leq i \leq m$ are all convex, the admissibility condition reads as follows:

- either, all the φ_i are affine and the set U given by (3.4.1) is non-empty;
- or, there exists an element $v^* \in V$ such that $\varphi_i(v^*) \leq 0$ for all $1 \leq i \leq m$ and $\varphi_i(v^*) < 0$ whenever φ_i is not affine linear.

If J is differentiable at u and the constraints are differentiable and admissible (at u), then (3.4.4) is a necessary condition for u to be a relative minimum of J at u. In addition, if J and the constraints φ_i are all convex, then (3.4.4) is both necessary and sufficient. Interested readers can find further details in the book *Introduction à l'analyse numérique matricielle et à l'optimisation* by P.G. Ciarlet (Masson, Paris, France, 1982; English translation, Cambridge University Press, Cambridge, UK, 1989).

3.5 Exercises

3.1 Give an example to show that the functional whose existence is guaranteed by Corollary 3.1.1, is not unique.

3.2 A normed linear space V is said to be *strictly convex* if for x and $y \in V$ such that $x \neq y$, $\|x\| = \|y\| = 1$, we have

$$\left\|\frac{1}{2}(x+y)\right\| < 1.$$

Show that the spaces ℓ_1 and ℓ_∞ are not strictly convex, while ℓ_2 is strictly convex.

3.3 If V is a normed linear space such that V^* is strictly convex, show that given a subspace W of V and a continuous linear functional f on W, its Hahn-Banach extension to all of V is unique. In particular, show also that the functional whose existence is guaranteed by Corollary 3.1.1, is unique.

3.4 Use Corollary 3.1.1 to prove Corollary 3.2.1. (Hint: consider the quotient space $V/\overline{W}$.)

3.5 Let V be a normed linear space and let W be a subspace of V. Let X be a finite dimensional space and let $T : W \to X$ be a continuous linear transformation. Show that there exists a continuous linear transformation $\widetilde{T} : V \to X$ which extends T.

3.6 Let V be a normed linear space and let W be a finite dimensional subspace. Show that there exists a closed subspace Z of V such that

$$V = W \oplus Z.$$

3.7 (a) Let c_0 denote the space of all real sequences which converge to zero, equipped with the norm $\|.\|_\infty$. Prove that

$$c_0^* = \ell_1.$$

(b) Show that

$$\ell_1^* = \ell_\infty.$$

(c) For any positive integer N, show that

$$(\ell_\infty^N)^* = \ell_1^N.$$

3.8 Let V be a normed linear space and let W be a subspace of V. Define

$$W^{\perp} = \{g \in V^* \mid g(x) = 0 \text{ for all } x \in W\}.$$

(a) Show that $W^{\perp}$ is a closed subspace of V^*.
(b) Let $f \in V^*$. Show that

$$d(f, W^{\perp}) = \|f|_W\|_{W^*}$$

where $f|_W$ is the restriction of f to W and

$$d(f, W^{\perp}) = \inf_{g \in W^{\perp}} \|f - g\|_{V^*}.$$

(c) Let $f \in W^*$ and let $\widetilde{f} \in V^*$ be an extension of f preserving its norm. Define $\sigma : W^* \to V^*/W^{\perp}$ by

$$\sigma(f) = \widetilde{f} + W^{\perp}.$$

Show that σ is well defined and that it is an isometric isomorphism of W^* onto $V^*/W^{\perp}$.
(d) Let $\pi : V \to V/W$ be the canonical quotient map $x \mapsto x + W$. For $f \in (V/W)^*$, define $\tau(f) = f \circ \pi \in V^*$. Show that the range of τ is equal to $W^{\perp}$ and that the map $\tau : (V/W)^* \to W^{\perp}$ is an isometric isomorphism.

3.9 Let V be a normed linear space and let Z be a subspace of V^*. Define

$$Z^{\perp} = \{x \in V \mid g(x) = 0 \text{ for all } g \in Z\}.$$

(a) Show that $Z^{\perp}$ is a closed subspace of V.
(b) Show that, if W is a subspace of V, then

$$\left(W^{\perp}\right)^{\perp} = \overline{W}.$$

(c) If Z is a subspace of V^*, show that

$$\left(Z^{\perp}\right)^{\perp} \supset \overline{Z}.$$

(d) If V is reflexive and if Z is a subspace of V^*, show that

$$\left(Z^{\perp}\right)^{\perp} = \overline{Z}.$$

3.10 Let φ and ψ be continuous mappings from $[0,1]$ into a real normed linear space V. For arbitrary scalars α and β, show that

$$\int_0^1 (\alpha\varphi(t) + \beta\psi(t))\ dt \ = \ \alpha \int_0^1 \varphi(t)\ dt + \beta \int_0^1 \psi(t)\ dt.$$

3.11 Let V and W be normed linear spaces and let $A \in \mathcal{L}(V, W)$. Let $\varphi : [0,1] \to V$ be a continuous mapping. Show that

$$A\left(\int_0^1 \varphi(t)\ dt\right) \ = \ \int_0^1 A(\varphi(t))\ dt.$$

3.12 Let $\varphi : [a,b] \to V$ be a continuous mapping and let $c \in (a,b)$. Show that

$$\int_a^b \varphi(t)\ dt \ = \ \int_a^c \varphi(t)\ dt + \int_c^b \varphi(t)\ dt.$$

3.13 Let $\varphi : \mathbb{R} \to V$ be a continuous mapping. Let $a, b, h \in \mathbb{R}$. Show that

$$\int_a^b \varphi(t)\ dt \ = \ \int_{a+h}^{b+h} \psi(t)\ dt$$

where $\psi(t) = \varphi(t-h)$.

3.14 Let $\varphi : [0,1] \to V$ be a continuous mapping and let $t \in [0,1)$. Show that

$$\lim_{h\to 0} \frac{1}{h} \int_t^{t+h} \varphi(s)\ ds \ = \ \varphi(t).$$

3.15 Let V be a normed linear space and let $T \in \mathcal{L}(V)$ such that $\|T\| \leq 1$. Let $\lambda > 0$ and let $x \in V$. Define

$$R(x) \ = \ \int_0^\infty e^{-\lambda t} T(x)\ dt.$$

Show that $R \in \mathcal{L}(V)$ and that

$$\|R\| \ \leq \ \frac{1}{\lambda}.$$

3.16 (a) Let V be a real Banach space . Let $k > 0$. Define

$$X \ = \ \left\{u \in \mathcal{C}([0,\infty); V) \mid \sup_{t\geq 0} e^{-kt}\|u(t)\|_V < \infty\right\}.$$

Define

$$\|u\|_X = \sup_{t\geq 0} e^{-kt}\|u(t)\|_V.$$

Show that X is a Banach space with this norm.
(b) Let $f : V \to V$ be a mapping. Assume that there exists $L > 0$ such that

$$\|f(u) - f(v)\|_V \leq L\|u - v\|_V$$

for all u and $v \in V$. For $u \in X$, define

$$F(u)(t) = u_0 + \int_0^t f(u(s))\, ds$$

where $u_0 \in V$ is a fixed vector. Show that $F(u) \in X$ and that, for any u and $v \in X$, we have

$$\|F(u) - F(v)\|_X \leq \frac{L}{k}\|u - v\|_X.$$

(c) Deduce that, for any $u_0 \in V$, there exists a unique $u \in \mathcal{C}([0,\infty); V)$ such that

$$u(t) = u_0 + \int_0^t f(u(s))\, ds$$

which is also the solution of the initial value problem

$$\begin{aligned} \tfrac{du}{dt}(t) &= f(u(t)),\ t > 0 \\ u(0) &= u_0. \end{aligned}$$

3.17 Let V be any vector space and let $f_0, f_1, \cdots, f_n$ be linear functionals on V. Let $\text{Ker}(f_i)$ denote the kernel of f_i, $0 \leq i \leq n$. Assume that

$$\cap_{i=1}^n \text{Ker}(f_i) \subset \text{Ker}(f_0).$$

Show that there exist scalars α_i, $1 \leq i \leq n$ such that

$$f_0 = \sum_{i=1}^N \alpha_i f_i.$$

(Hint: Consider the image of the map $\mathcal{A}$ as defined in Section 3.3 and apply Corollary 3.2.1 to its image.)

3.18 Let V be a real normed linear space and let $K \subset V$ be a compact and convex subset. Let $C \subset V^*$ be a convex cone. Assume that for each $f \in C$, there exists a vector $x \in K$ (depending on f), such that $f(x) \geq 0$. Show that there exists $x \in K$ such that $f(x) \geq 0$ for all $f \in C$. (Hint: For $f \in C$, consider

$$K_f = \{x \in K \mid f(x) \geq 0\}$$

which is a non-empty closed set. Show that this collection of closed sets has finite intersection property.)