JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 6, No. 2, 1970

# A Satisfactory Treatment of Equality and Operator Constraints in the Dubovitskii–Milyutin Optimization Formalism[1]

H. HALKIN[2]

**Abstract.** The formalism of Dubovitskii and Milyutin is very attractive but, up to now, it could not be applied to optimization problems involving equality and operator constraints. In the present paper, the formalism of Dubovitskii and Milyutin is extended to this more general situation. Theorem 2.1, the main result of the paper, is applied to the standard mathematical programming problem in normed linear space and an abstract maximum principle is obtained.

## 1. Introduction

During the last few years, many papers have been devoted to obtaining necessary conditions in mathematical programming which could be applied to optimal control problems (Refs. 1–8). In parallel with these papers, one must mention the papers of Gamkrelidze (Ref. 9) and Dubovitskii and Milyutin (Ref. 10). The relation between Gamkrelidze's approach and the mathematical programming approach is relatively clear and has been already commented on by Neustadt (Ref. 3) and Halkin (Ref. 5). The formalism of Dubovitskii and Milyutin is very attractive but, in its present stage of development, gives results which are weaker than those obtained in the papers mentioned above. In this paper, we extend the formalism of Dubovitskii and Milyutin to include these stronger results.

For the sake of simplicity, the results of the present paper are given within the framework of a normed linear space. The reader should be aware that these results can be extended to an arbitrary linear topological space or even a linear quasitopological space (see Lobry, Ref. 11).

[1] Paper received February 20, 1970. This research was sponsored by the Air Force Office of Scientific Research, Office of Aerospace Research, USAF, under Grant No. AFOSR-68-1529A. The author thanks Dr. K. Makowski for several valuable comments of an earlier draft of the present paper.

[2] Professor of Mathematics, University of California at San Diego, La Jolla, California.

## 2. Problem Statement and Theorem

In many optimization problems, one may express the optimality of a certain element by stating that a certain finite family $\{S_i : i \in I\}$ of subsets of a normed linear space $X$ has an empty intersection,[3] i.e.,

$$\bigcap_{i \in I} S_i = \varnothing. \tag{1}$$

The two essential parts of the formalism of Dubovitskii and Milyutin are these:

(a) To each set $S_i$, associate a convex set $\Omega_i$ which is such an appropriate approximation of $S_i$ that it is possible to prove that relation (1) implies

$$\bigcap_{i \in I} \Omega_i = \varnothing. \tag{2}$$

(b) From Relation (2), prove that the family $\{\Omega_i : i \in I\}$ is *separated* in the following sense: there exists a set of continuous linear functionals $\{\omega_i : i \in I\}$ over the normed linear space $X$ such that ($\alpha$) $\sum_{i \in I} \omega_i = 0$, ($\beta$) $\omega_i \neq 0$ for some $i \in I$, and ($\gamma$) $\omega_i(x) \geqslant 0$ whenever $i \in I$ and $x \in \Omega_i$.

In this paper, we shall consider situations in which we can prove that relation (1) implies that the family $\{\Omega_i : i \in I\}$ is separated, but[4] for which we could not prove that relation (1) implies relation (2).

Before stating Theorem 2.1, we shall give a few definitions.

**Definition 2.1** (Refs. 10–11). A subset $\Omega$ of a normed linear space $X$ is an *interior convex approximation* to a subset $S$ of $X$ if (i) $\Omega$ is open, (ii) $\Omega$ is convex, (iii) $0 \in \bar{\Omega}$, where $\bar{\Omega}$ is the closure of $\Omega$, and (iv) for all $\bar{x} \in \Omega$, there exists an $\epsilon > 0$ such that $| x - \bar{x} | < \epsilon$ and $\eta \in (0, \epsilon)$ imply $\eta x \in S$.

**Definition 2.2** (Ref. 3). A subset $\Omega$ of a normed linear space $X$ is a tangent approximation to a subset $S$ of $X$ if there exists a $\rho > 0$, a continuous real-valued function $\varphi$ defined[5] on $U = \{x: x \in X, | x | < \rho\}$, and a nonzero con-

[3] The empty set will de denoted by $\varnothing$.

[4] In the particular case treated by Dubovitskii and Milyutin, all but at most one of the sets $\Omega_i$ are open, and in that case one can prove (see Lemma 4.2) that relation (2) holds if, and only if, the family $\{\Omega_i : i \in I\}$ is separated. In this paper, we consider situations for which several of the sets $\Omega_i$ could fail to be open and, hence, for which we cannot say that relation (2) holds if the family $\{\Omega_i : i \in I\}$ is separated.

[5] If $x$ is an element of the normed linear space $X$, then $| x |$ will denote the norm of $x$.

tinuous linear function $h$ defined on $X$ such that (i) $\{x: x \in X, h(x) = 0\} = \Omega$, (ii) $\{x: x \in U, \varphi(x) = 0\} \subset S$, and (iii) for all $\bar{x} \in U$, we have

$$\lim_{x \to \bar{x}, \epsilon \to 0+} (1/\epsilon)[\varphi(\epsilon x) - h(\epsilon x)] = 0.$$

**Definition 2.3** (Refs. 2, 3, 6). If $k$ is a positive integer, we shall say that a subset $\Omega$ of a normed linear space $X$ is a *k-convex approximation* to a subset $S$ of $X$ if (i) $\Omega$ is convex, (ii) $0 \in \bar{\Omega}$, and (iii) for all sets $\{x_1, \ldots, x_l\}$ with $l \leqslant k$ elements in general position[6] in $\Omega$ and for all real numbers $\sigma > 0$, there exist a continuous function $\zeta$ from[7] $\mathrm{co}\{x_1, \ldots, x_l\}$ into $S$ and a real $\alpha > 1/\sigma$ such that $| \alpha\zeta(x) - x | \leqslant \sigma$ whenever $x \in \mathrm{co}\{x_1, \ldots, x_l\}$.

**Theorem 2.1.** If $I = \{-\mu, \ldots, m + 1\}$ and if $\{S_i: i \in I\}$ and $\{\Omega_i: i \in I\}$ are families of subsets of a normed linear space $X$ such that (i) $\bigcap_{i \in I} S_i = \varnothing$, (ii) for $i = -\mu, \ldots, 0$, the set $\Omega_i$ is an interior convex approximation to the set $S_i$, (iii) for $i = 1, \ldots, m$, the set $\Omega_i$ is a tangent approximation to the set $S_i$, and (iv) $\Omega_{m+1}$ is an $(m + 1)$-convex approximation to the set $S_{m+1}$, then the family $\{\Omega_i: i \in I\}$ is separated.

Dubovitskii and Milyutin have considered only the case $m = 0$, and in this case the proof of Theorem 2.1 is very easy. In order to apply Theorem 2.1 with $m = 0$ to the standard optimal control problem, Dubovitskii and Milyutin have thus been obliged (a) to lump equality constraints $x \in S_i$, $i = 1, \ldots, m$, and operator constraint $x \in S_{m+1}$ into a single constraint $x \in \bigcap_{i=1,\ldots,m+1} S_i$ and (b) to prove that, if $\omega^*$ is a continuous linear functional such that $\omega^*(x) \geqslant 0$ for all $x \in \bigcap_{i=1,\ldots,m+1} \Omega_i$, then $\omega^* = \omega_1 + \cdots + \omega_{m+1}$, where $\omega_1, \ldots, \omega_{m+1}$ are continuous linear functionals such that $\omega_i(x) \geqslant 0$ whenever $x \in \Omega_i$. This second step is justified by Dubovitskii and Milyutin (Ref. 10, p. 41) under some further assumptions which are not required in Theorem 2.1 given here. Moreover, even under these assumptions, this second step requires an (algebraic) topological argument which is not given in Ref. 10, but for which the reader is referred to the classical work of Pontryagin *et al.*

## 3. Properties of Convex Approximations

In this section, we shall give three lemmas concerning intersection, mapping, and topological properties of convex approximations.

[6] The set $\{x_1, \ldots, x_l\}$ is in general position if $x_2 - x_1, \ldots, x_l - x_1$ are linearly independent.

[7] The convex hull of $\{x_1, \ldots, x_l\}$ is denoted by $\mathrm{co}\{x_1, \ldots, x_l\}$.

**Lemma 3.1.** If $X$ is a normed linear space, if $\Omega_1 \subset X$ is an interior convex approximation to a set $S_1 \subset X$, if $\Omega_2 \subset X$ is a $k$-convex approximation to a set $S_2 \subset X$, and if $\Omega_1 \cap \Omega_2 \neq \varnothing$, then $\Omega_1 \cap \Omega_2$ is a $k$-convex approximation to the set $S_1 \cap S_2$.

**Proof.** Let $l \leqslant k$ and let $\{x_1, ..., x_l\} \subset (\Omega_1 \cap \Omega_2)$ be in general position. Since $\Omega_1$ is an interior convex approximation to the set $S_1$ and since the set $\text{co}\{x_1, ..., x_l\}$ is compact, we know that there exists an $\epsilon > 0$ such that $| x - \bar{x} | < \epsilon$ for some $\bar{x} \in \text{co}\{x_1, ..., x_l\}$ and $\eta \in (0, \epsilon)$ imply $\eta x \in S_1$. Let $\sigma > 0$ and let $\sigma^* = \min\{\sigma, \epsilon\}$. Since $\Omega_2$ is a $k$-convex approximation to the set $S_2$, we know that there exist a function $\zeta$ from $\text{co}\{x_1, ..., x_l\}$ into $S_2$ and a real $\alpha > 1/\sigma^*$ such that $| \alpha\zeta(x) - x | \leqslant \sigma^*$ whenever $x \in \text{co}\{x_1, ..., x_l\}$. Since $\sigma^* \leqslant \epsilon$, the function $\zeta$ maps $\text{co}\{x_1, ..., x_l\}$ into $S_1$, and Lemma 3.1 is proved.

**Lemma 3.2.** If $X$ and $Y$ are normed linear spaces, if $\Omega \subset X$ is a $k$-convex approximation to a set $S \subset X$, if $\rho > 0$, if $\varphi$ is a continuous mapping from $U = \{x: x \in X, | x | < \rho\}$ into $Y$, and if $h$ is a continuous linear mapping from $X$ into $Y$ such that, for all $\bar{x} \in U$, we have

$$\lim_{\epsilon \to 0, x \to \bar{x}} | \varphi(\epsilon x) - h(\epsilon x)|/\epsilon = 0,$$

then $h(\Omega)$ is a $k$-convex approximation to $\varphi(S \cap U)$.

**Proof.** Let $\{y_1, ..., y_l\}$ with $l \leqslant k$ be in general position in $h(\Omega)$, and let $\sigma > 0$. Let $x_1, ..., x_l \in \Omega$ such that $y_i = h(x_i)$ for $i = 1, ..., l$. The elements $x_1, ..., x_l$ are in general position. Since the set $\text{co}\{x_1, ..., x_l\}$ is compact, there is an $\eta > 0$ such that $| \epsilon x | < \rho$ and

$$| \varphi(\epsilon x) - h(\epsilon x)|/\epsilon \leqslant \sigma/4,$$

whenever $\epsilon \in (0, \eta)$ and $| x - \bar{x} | \leqslant \eta$ for some $\bar{x} \in \text{co}\{x_1, ..., x_l\}$. For every $y \in \text{co}\{y_1, ..., y_l\}$, let $\gamma(y) = (\gamma_1(y), ..., \gamma_l(y))$ be the barycentric coordinates of $y$ with respect to $\{y_1, ..., y_l\}$, that is, $\gamma(y)$ is the unique element of $R^l$ such that $\gamma_i(y) \geqslant 0$ for $i = 1, ..., l$; $\sum_{i=1}^{l} \gamma_i(y) = 1$ and $y = \sum_{i=1}^{l} \gamma_i(y)\, y_i$. Let $g(y) = \sum_{i=1}^{l} \gamma_i(y)\, x_i$. The function $g$ is an affine mapping from $\text{co}\{y_1, ..., y_l\}$ into $\text{co}\{x_1, ..., x_l\}$ such that $h(g(y)) = y$ for every $y \in \text{co}\{y_1, ..., y_l\}$ and $g(h(x)) = x$ for every $x \in \text{co}\{x_1, ..., x_l\}$. Let $L < +\infty$ be such that $| h(x)| \leqslant L \,| x |$ for all $x \in X$. Let $\sigma^* > 0$ be such that

$$(1 + L)\,\sigma^* < \sigma/4, \qquad \sigma^* < \eta, \qquad (\sigma^*)^2 < \rho/4, \qquad 4\sigma^* \max\{| x_i | : i = 1, ..., l\} < \rho.$$

We know that there is an $\alpha > 1/\sigma^*$ and a continuous function $\zeta^*$ from $\text{co}\{x_1, ..., x_l\}$ into $S$ such that $| \alpha\zeta^*(x) - x | \leqslant \sigma^*$ whenever $x \in \text{co}\{x_1, ..., x_l\}$. We now define a mapping $\zeta$ from $\text{co}\{y_1, ..., y_l\}$ into $Y$ by the relation $\zeta(y) = \varphi(\zeta^*(g(y)))$. We remark immediately that $\alpha > 1/\sigma^* \geqslant 1/\sigma$, that $\zeta$ is continuous (since $\zeta$ is the composition of continuous functions), and that $\zeta(y) \in \varphi(S \cap U)$ whenever $y \in \text{co}\{y_1, ..., y_l\}$, since

$$| \zeta^*(x)| \leqslant | \alpha\zeta^*(x) - x |/\alpha + | x |/\alpha \leqslant (\sigma^*)^2 + \rho/4 \leqslant \rho/2,$$

whenever $x \in \text{co}\{x_1, ..., x_l\}$. We conclude the proof of Lemma 3.2 by showing that $| \alpha\zeta(y) - y | \leqslant \sigma$ whenever $y \in \text{co}\{y_1, ..., y_l\}$ or, equivalently, that $| \alpha\varphi(\zeta^*(x)) - h(x)| \leqslant \sigma$ whenever $x \in \text{co}\{x_1, ..., x_l\}$. Indeed,

$$| \alpha\varphi(\zeta^*(x)) - h(x)| \leqslant | \alpha h(\zeta^*(x)) - h(x)| + | \alpha\varphi(\zeta^*(x)) - \alpha h(\zeta^*(x))| \leqslant L \, | \alpha\zeta^*(x) - x |$$
$$+ | \varphi((1/\alpha)(\alpha\zeta^*(x))) - h((1/\alpha)(\alpha\zeta^*(x))) | /(1/\alpha) \leqslant L\sigma^* + \sigma/4 \leqslant \sigma/4 + \sigma/4 < \sigma.$$

This concludes the proof of Lemma 3.2.

**Lemma 3.3.** If $\Omega \subset R^k$ is a $(k + 1)$-convex approximation to a subset $S$ of $R^k$ and if $0 \in \text{int}\, \Omega$, then $0 \in S$.

**Proof.** If $0 \in \text{int}\, \Omega$, then there is an $\epsilon > 0$ and a set $\{y_1, ..., y_{k+1}\} \subset \Omega$ in general position such that $| y | < \epsilon$ implies $y \in \text{co}\{y_1, ..., y_{k+1}\}$. Since $\Omega$ is a $(k + 1)$-convex approximation to the set $S$, then there exist a continuous function $\zeta$ from $\text{co}\{y_1, ..., y_{k+1}\}$ into $S$ and an $\alpha > 2/\epsilon$ such that

$$\sup_{y \in \text{co}\{y_1, ..., y_{k+1}\}} | \alpha\zeta(y) - y | \leqslant \epsilon/2.$$

We define a continuous function $h$ from $\text{co}\{y_1, ..., y_{k+1}\}$ into itself by the relation $h(y) = y - \alpha\zeta(y)$. Let $y^*$ be a fixed point of $h$ (Brouwer theorem); then, $\alpha\zeta(y^*) = 0$, that is, $\zeta(y^*) = 0$, which implies $0 \in S$. This concludes the proof of Lemma 3.3.

## 4. Some Consequences of the Hahn–Banach Theorem

In this section, we shall state and prove a few results related to the Hahn–Banach theorem.

**Hahn–Banach Theorem.** If $\Omega_1$ and $\Omega_2$ are disjoint, nonempty convex subsets of a normed linear space $X$ and if $\Omega_1$ is open, then there exists an affine, continuous nonconstant functional $\omega$ defined over $X$ such that $\omega(x) > 0$ for all $x \in \Omega_1$ and $\omega(x) \leqslant 0$ for all $x \in \Omega_2$.

**Lemma 4.1.** If $h$ is a concave function on a linear space $X$, $h(0) = 0$, $\Omega = \{x: x \in X, h(x) > 0\}$ is nonempty, and if $\omega$ is a linear function on $X$ such that $\omega(x) > 0$ for all $x \in \Omega$, then, for some $\lambda > 0$, we have $\lambda h(x) \leqslant \omega(x)$ for all $x \in X$. Moreover, if $h$ is linear, we have $\lambda h(x) = \omega(x)$ for all $x \in X$.

**Proof.** Let $S = \{(\omega(x), h(x) - t) : x \in X, t \geqslant 0\}$. The set $S$ is convex and $0 \notin \operatorname{int} S$; hence, there exist constants $\alpha$ and $\beta$, not both zero, such that $\alpha\omega(x) + \beta(h(x) - t) \geqslant 0$ for all $x \in X$ and all $t \geqslant 0$. This implies $\beta < 0$ and $\alpha\omega(x) + \beta h(x) \geqslant 0$ for all $x \in X$. There is an $x^* \in X$ such that $h(x^*)$ and $\omega(x^*) > 0$, that is, $\alpha\omega(x^*) \geqslant |\beta| h(x^*) > 0$, which implies $\alpha > 0$. If we let $\lambda = -\beta/\alpha > 0$, we have $\lambda h(x) \leqslant \omega(x)$ for all $x \in X$. This concludes the proof of Lemma 4.1.

**Lemma 4.2.**[8] If $\Omega_i$, $i = 0, 1,..., k$, are convex sets in a normed linear space $X$ such that (i) $0 \in \bar{\Omega}_i$ for $i = 0, 1,..., k$, (ii) $\Omega_i$ is open for $i = 1,..., k$, then the two following statements are equivalent: ($\alpha$) $\bigcap_{i=0,1,\ldots,k} \Omega_i = \varnothing$ and ($\beta$) the family $\{\Omega_i : i = 0,..., k\}$ is separated.

**Proof.** We begin by proving the easy implication $(\beta) \Rightarrow (\alpha)$. Let $\omega_0, ..., \omega_k$ be the continuous linear functionals such that (i) $\omega_0 + \cdots + \omega_k = 0$, (ii) not all $\omega_i$ are zero, and (iii) $i \in \{0,..., k\}$ and $x \in \Omega_i$ imply $\omega_i(x) \geqslant 0$.

From (i) and (ii), there are at least two indices, $i$ and $j$, $i \neq j$, in $\{0,..., k\}$ such that $\omega_i \neq 0$ and $\omega_j \neq 0$. There exists then at least one index $l$ in $\{1,..., k\}$ such that $\omega_l \neq 0$. If $x \in \bigcap_{i=0,\ldots,k} \Omega_i$, then $\omega_i(x) \geqslant 0$ for all $i = 0,..., k$ and $\omega_l(x) > 0$. We have then $\omega_0(x) + \cdots + \omega_k(x) > 0$, which contradicts (i).

We shall now prove the implication $(\alpha) \Rightarrow (\beta)$. Let

$$K_1 = \{(x_1, ..., x_k) : x_1 = \cdots = x_k \in \Omega_0\},$$

and let $K_2 = \Omega_1 \times \cdots \times \Omega_k$. The sets $K_1$ and $K_2$ are disjoint convex subsets of $X^k$ such that $0 \in \bar{K}_1$, $0 \in \bar{K}_2$ and $K_2$ is open. Hence, by the Hahn–Banach theorem, there exists a continuous, nonzero linear functional $\omega$ on $X^k$ such that $\omega(x_1, ..., x_k) > 0$ for all $(x_1, ..., x_k) \in K_2$ and $\omega(x_1, ..., x_k) \leqslant 0$ for all $(x_1, ..., x_k) \in K_1$. We have $\omega(x_1, ..., x_k) = \omega_1(x_1) + \cdots + \omega_k(x_k)$, where $\omega_1, ..., \omega_k$ are continuous linear functionals on $X$, not all zero. Let

[8] Lemma 4.2 is stated in Dubovitskii and Milyutin (Ref. 10) and proved in Lobry (Ref. 11). The proof given here is new.

$\omega_0 = -(\omega_1 + \cdots + \omega_k)$. The functionals $\omega_0, \omega_1, ..., \omega_k$ satisfy the three requirements (i) $\omega_0 + \omega_1 + \cdots + \omega_k = 0$, (ii) not all $\omega_i$ are zero, (iii) $i \in \{0, ..., k\}$ and $x \in \Omega_i$ imply $\omega_i(x) \geqslant 0$. This concludes the proof of Lemma 4.2.

**Lemma 4.3.** If $\Omega_i$, $i = 0, ..., k$, are convex sets in a normed linear space $X$ and if $\omega$ is a nonzero continuous linear functional on $X$ such that (i) $0 \in \bar{\Omega}_i$ for $i = 0, ..., k$, (ii) $\Omega_i$ is open for $i = 1, ..., k$, (iii) $\bigcap_{i=0,...,k} \Omega_i \neq \varnothing$, (iv) $\omega(x) \geqslant 0$ for all $x \in \bigcap_{i=0,...,k} \Omega_i$, then there exist continuous linear functionals $\omega_0, ..., \omega_k$ on $X$ such that ($\alpha$) $\omega = \omega_0 + \cdots + \omega_k$ and ($\beta$) $i \in \{0, ..., k\}$ and $x \in \Omega_i$ imply $\omega_i(x) \geqslant 0$.

**Proof.** Let $\omega_{k+1} = -\omega$ and $\Omega_{k+1} = \{x: x \in X, \omega_{k+1}(x) > 0\}$. By applying Lemma 4.2 [($\alpha$) $\Rightarrow$ ($\beta$)] to the sets $\Omega_0, ..., \Omega_{k+1}$, we obtain functionals $\omega_0^*, ..., \omega_{k+1}^*$ such that (i) $\omega_0^* + \cdots + \omega_{k+1}^* = 0$, (ii) not all of these functionals are zero, and (iii) $\omega_i^*(x) \geqslant 0$ whenever $i \in \{0, ..., k+1\}$ and $x \in \Omega_i$. We have $\omega_{k+1}^* \neq 0$, since otherwise, by Lemma 4.2 [($\beta$) $\Rightarrow$ ($\alpha$)], we would have $\bigcap_{i=0,...,k} \Omega_i = \varnothing$. If $\omega_{k+1}^* \neq 0$ and $\omega_{k+1}^*(x) \geqslant 0$ for all $x$ in the *open* set $\Omega_{k+1}$, then $\omega_{k+1}^*(x) > 0$ for all $x$ in $\Omega_{k+1}$; and, by Lemma 4.1, there exists a $\lambda > 0$ such that $\omega_{k+1}^* = \lambda\omega_{k+1}$. We conclude the proof of Lemma 4.3 by letting $\omega_i = (1/\lambda)\, \omega_i^*$ for $i = 0, ..., k$.

## 5. Proof of Theorem 2.1

By repeated applications of Lemma 3.1, we know that the set

$$\Omega = \Omega_{-\mu} \cap \Omega_{-\mu+1} \cap \cdots \cap \Omega_0 \cap \Omega_{m+1}$$

is an $(m+1)$-convex approximation of the set

$$S = S_{-\mu} \cap S_{-\mu+1} \cap \cdots \cap S_0 \cap S_{m+1}.$$

If $\Omega = \varnothing$, then Theorem 2.1 is a direct consequence of Lemma 4.2. If $m = 0$, then we have $S = \varnothing$, and hence $\Omega = \varnothing$. From now on, we shall assume that $m > 0$ and that $\Omega \neq \varnothing$.

For $i = 1, ..., m$, let $\rho_i$ be a positive number, and let $\varphi_i$ and $h_i$ be two functions satisfying the requirements stated in the definition of $\Omega_i$ as a tangent approximation to the set $S_i$; let $h^+ = (h_1, ..., h_m)$ and $\varphi^+ = (\varphi_1, ..., \varphi_m)$. By Lemma 3.2, we know that the set $h^+(\Omega)$ is an $(m+1)$-convex approximation of the set $\varphi^+(S \cap U)$, where $U = \{x: x \in X, |x| < \min_{i=1,...,m} \rho_i\}$.

We cannot have $0 \in \text{int}\, h^+(\Omega)$, since, by Lemma 3.3, we would then have $0 \in \varphi^+(S \cap U)$, which would contradict the assumption $\bigcap_{i=-\mu,...,m+1} S_i = \varnothing$. Since $h^+(\Omega)$ is a convex subset of $R^m$, we know that there is a vector $\lambda = (\lambda_1, ..., \lambda_m) \neq 0$ such that $\sum_{i=1}^m \lambda_i h_i(x) \geqslant 0$ for all $x \in \Omega$. For $i = 1,..., m$, we let $\omega_i = -\lambda_i h_i$. Since $h_i \neq 0$ for all $i \in \{1,..., m\}$ and since $\lambda_{i*} \neq 0$ for some $i^* \in \{1,..., m\}$, it follows that $\omega_{i*} \neq 0$. If $\sum_{i=1}^m \lambda_i h_i = 0$, we set $\omega_i = 0$ for $i = -\mu,..., 0, m + 1$; otherwise, the functional $\sum_{i=1}^m \lambda_i h_i$ is continuous, linear, nonzero, and nonnegative on $\Omega = \bigcap_{i=-\mu,...,0,m+1} \Omega_i$; hence, by Lemma 4.3, we have $\sum_{i=1}^m \lambda_i h_i = \omega_{-\mu} + \cdots + \omega_0 + \omega_{m+1}$, where $\omega_{-\mu}, ..., \omega_0, \omega_{m+1}$ are continuous linear functionals such that $\omega_i(x) \geqslant 0$ if $x \in \Omega_i$ and $i \in \{-\mu,..., 0, m + 1\}$. The functionals $\omega_{-\mu}, ..., \omega_{m+1}$ satisfy all the requirements stated in Theorem 2.1.

## 6. Abstract Maximum Principle

In this section, we recall a rather general form of the abstract maximum principle for mathematical programming problems and we prove that it can be derived from Theorem 2.1.

We are given nonnegative integers $\mu$ and $m$, a normed linear space $X$, a subset $L$ of $X$, and a function $\varphi = (\varphi_{-\mu}, ..., \varphi_{-1}, \varphi_0, \varphi_1, ..., \varphi_m)$ from $X$ into $R^{\mu+m+1}$. The problem is to find an $\hat{x} \in A$ such that $\varphi_0(\hat{x}) = \min_{x \in A} \varphi_0(x)$, where $A = \{x: \varphi_i(x) \leqslant 0$ for $i = -\mu,..., -1; \varphi_i(x) = 0$ for $i = 1,..., m$ and $x \in L\}$. An element $\hat{x}$ satisfying this condition will be called an optimal solution of the given problem.

We assume that the function $\varphi = (\varphi_{-\mu}, ..., \varphi_m)$ is continuous and that there exists a continuous function $h = (h_{-\mu}, ..., h_m)$ from $X$ into $R^{\mu+m+1}$ such that

(i) $\lim_{|x| \to 0} |\varphi(\hat{x}) + h(x) - \varphi(\hat{x} + x)|/|x| = 0;$

(ii) $h_i$ is convex for $i = -\mu,..., 0$ and linear for $i = 1,..., m$.

We assume, moreover, that there exists a set $M \subset X$ which is an $(m + 1)$-convex approximation to the set $L - \hat{x}$.[9]

**Abstract Maximum Principle.** If $\hat{x}$ is an optimal solution of the given problem, then there exist a nonzero vector $\lambda = (\lambda_{-\mu}, ..., \lambda_m) \in R^{\mu+m+1}$

[9] We denote by $L - \hat{x}$ the set $\{x - \hat{x}: x \in L\}$.

and a continuous linear function $l = (l_{-\mu}, ..., l_m)$ from $X$ into $R^{\mu+m+1}$ such that

(i) $\sum_{i=-\mu,\ldots,m} \lambda_i h_i(x) \leqslant \sum_{i=-\mu,\ldots,m} \lambda_i l_i(x) \leqslant 0$ for all $x \in M$;

(ii) $\lambda_i \leqslant 0$ for $i = -\mu, ..., 0$;

(iii) $\lambda_i \varphi_i(\hat{x}) = 0$ for $i = -\mu, ..., -1$;

(iv) for $i = -\mu, ..., 0$, we have $l_i(x) \leqslant h_i(x)$ for all $x \in X$;

(v) for $i = 1, ..., m$, we have $l_i = h_i$.

**Proof.** We define the sets

$$S_i = \{x - \hat{x} : x \in X, \ \ \varphi_i(x) \leqslant 0\} \quad \text{if} \quad i = -\mu, ..., -1,$$

$$S_0 = \{x - \hat{x} : x \in X, \ \ \varphi_0(x) < \varphi_0(\hat{x})\},$$

$$S_i = \{x - \hat{x} : x \in X, \ \ \varphi_i(x) = 0\} \quad \text{if} \quad i = 1, ..., m,$$

$$S_{m+1} = L - \hat{x}$$

and

$$\Omega_i = \{x : x \in X, \ \ \varphi_i(\hat{x}) + h_i(x) < 0\} \quad \text{if} \quad i = -\mu, ..., -1,$$

$$\Omega_0 = \{x : x \in X, \ \ h_0(x) < 0\},$$

$$\Omega_i = \{x : x \in X, \ \ h_i(x) = 0\} \quad \text{if} \quad i = 1, ..., m,$$

$$\Omega_{m+1} = M.$$

If, for some $i^* \in \{-\mu, ..., 0\}$, we have $\Omega_{i^*} = \varnothing$ or if, for some $i^* \in \{1, ..., m\}$, we have $h_{i^*} = 0$, we may set $\lambda_{i^*} = -1$, $\lambda_i = 0$ for all $i \in \{-\mu, ..., m\}$ with $i \neq i^*$; and, for each $i = -\mu, ..., m$, we may set $l_i$ to be a continuous linear support functional to $h_i$ (see Lemma 6.2 stated and proved below). From now on, we shall assume that $\Omega_i \neq \varnothing$ for all $i \in \{-\mu, ..., 0\}$ and that $h_i \neq 0$ for all $i \in \{1, ..., m\}$. From the stated assumptions, we know that, for every $i = 1, ..., m$, the set $\Omega_i$ is a tangent approximation to the set $S_i$ and that $\Omega_{m+1}$ is an $(m+1)$-convex approximation to $S_{m+1}$. Moreover, for every $i = -\mu, ..., -1$, the set $\Omega_i$ is an interior convex approximation to the set $S_i$ (apply Lemma 6.1 given below to the case $\varphi = \varphi_i$, $h = h_i$, $\alpha = 0$). The set $\Omega_0$ is an interior convex approximation to the set $S_0$ [apply Lemma 6.1 to the case $\varphi = \varphi_0$, $h = h_0$, $\alpha = \varphi_0(\hat{x})$].

If $\hat{x}$ is an optimal solution of the given problem, then $\bigcap_{i=-\mu,\ldots,m+1} S_i = \varnothing$

and, from Theorem 2.1, we know that the family $\{\Omega_i : i = -\mu,..., m+1\}$ is separated. We have then $\omega_{-\mu}(x) + ... + \omega_m(x) \leqslant 0$ for all $x \in \Omega_{m+1} = M$. We note immediately that, if $\varphi_i(\hat{x}) < 0$, for some $i \in \{-\mu,..., -1\}$, we have $0 \in \text{int}\, \Omega_i$, and hence $\omega_i = 0$. Let $J = \{i: i \in \{-\mu,..., m\}, \omega_i = 0\}$. If $i \in J$, let $\lambda_i = 0$ and let $l_i$ be a continuous linear support functional for $h_i$ (Lemma 6.2). If $i \notin J$, $i \in \{-\mu,..., 0\}$, we apply Lemma 4.1 to the case $h = -h_i$, $\omega = \omega_i$, and we obtain a $\lambda_i = -\lambda < 0$ such that $\lambda_i h_i(x) \leqslant \omega_i(x)$ for all $x \in X$. If $i \notin J$, $i \in \{1,..., m\}$, we apply Lemma 6.3 given below to the case $h = h_i$, $\omega = \omega_i$, and we obtain a $\lambda_i = \lambda \neq 0$ such that $\lambda_i h_i(x) = \omega_i(x)$. For all $i \notin J, i \in \{-\mu,..., m\}$, we let $l_i = \omega_i/\lambda_i$. This concludes the proof of the abstract maximum principle.

**Lemma 6.1.** If $\varphi$ and $h$ are continuous functionals on a normed linear space $X$, if $\hat{x} \in X$, if $\alpha \geqslant \varphi(\hat{x})$, if $h$ is convex and $h(0) = 0$, if

$$\lim_{|x| \to 0} | \varphi(\hat{x}) + h(x) - \varphi(\hat{x} + x)|/| x | = 0,$$

if $S = \{x: \varphi(x) \leqslant \alpha\}$, and if $\Omega = \{x: x \in X, \varphi(\hat{x}) + h(x) < \alpha\}$ is nonempty, then $\Omega$ is an interior convex approximation to the set $S - \hat{x}$.

**Proof.** If $\alpha > \varphi(\hat{x})$, then $\hat{x} \in \text{int}\, S$ and $0 \in \text{int}\, \Omega$, and $\Omega$ is (trivially) an interior convex approximation to the set $S$. Let us assume that $\alpha = \varphi(\hat{x})$. If $y \in \Omega$, then $h(y) < 0$. Let $\sigma = -h(y) > 0$. Since $h$ is continuous, there exists a $\rho > 0$ such that $| x - y | < \rho$ implies $h(x) < -\sigma/2$. Moreover, since $h$ is convex and $h(0) = 0$, we have $h(xt) < -(\sigma/2)t$ whenever $t \in (0, 1]$ and $| x - y | < \rho$. There is then a $\delta \in (0, 1]$ such that $t \in (0, \delta)$ and $| x - y | < \rho$ imply

$$| \varphi(\hat{x}) + h(tx) - \varphi(\hat{x} + tx)| \leqslant (\sigma/2)t,$$

that is,

$$\varphi(\hat{x} + tx) - \varphi(\hat{x}) \leqslant (\sigma/2)t + h(tx) \leqslant (\sigma/2)t - (\sigma/2)t = 0,$$

that is, $\hat{x} + tx \in S$. If we let $\epsilon = \min\{\delta, \rho\}$, we have then $tx \in S - \hat{x}$ whenever $| x - y | < \epsilon$ and $t \in (0, \epsilon)$. This concludes the proof of Lemma 6.1.

**Lemma 6.2.**[10] If $h$ is a continuous convex functional on a normed linear space $X$ with $h(0) = 0$, then there exists a continuous linear functional $l$ defined over $X$ such that $l(x) \leqslant h(x)$ for all $x \in X$. The functional $l$ is called a support functional of $h$.

[10] Lemma 6.2 is a well-known consequence of the Hahn–Banach theorem.

**Proof.** The subsets $\Omega_1 = \{(x, t): x \in X, t > h(x)\}$ and $\Omega_2 = \{0\}$ of $X \times R$ satisfy the conditions of the Hahn–Banach theorem; hence, there exists a nonzero continuous linear functional $\omega$ defined over $X \times R$ such that $\omega(x, t) > 0$ for all $(x, t) \in \Omega_1$. There exist a continuous linear functional $l_1$ defined over $X$ and a real number $\alpha$ such that $\omega(x, t) = l_1(x) + \alpha t$ for all $x \in X$ and all $t \in R$. If $l_1(x) + \alpha t > 0$ for all $x \in X$ and all $t > h(x)$, it follows that $\alpha > 0$. We have then $(1/\alpha)\, l_1(x) + h(x) \geqslant 0$ for all $x \in X$. By letting $l = -(1/\alpha)\, l_1$, we conclude the proof of Lemma 6.2.

**Lemma 6.3.** If $h$ and $\omega$ are nonzero linear functionals on a linear space $X$ such that $\omega(x) \geqslant 0$ for all $x \in X$ with $h(x) = 0$, then there is a $\lambda \neq 0$ such that $\lambda h(x) = \omega(x)$ for all $x \in X$.

**Proof.** Let $S = \{(\omega(x), h(x)): x \in X\}$. The set $S$ is convex and $0 \notin \operatorname{int} S$; hence, there exist constants $\alpha$ and $\beta$, not both zero, such that $\alpha\omega(x) + \beta h(x) \geqslant 0$ for all $x \in X$, that is, such that $\alpha\omega(x) + \beta h(x) = 0$ for all $x \in X$, since $X$ is a linear space. We have $\alpha$ *and* $\beta \neq 0$ since $\omega$ *and* $h \neq 0$. By letting $\lambda = -\beta/\alpha$, we conclude the proof of Lemma 6.3.

## References


1. NEUSTADT, L. W., *Optimal Control Problems as Extremal Problems in a Banach Space*, Proceedings of the Symposium on System Theory, Polytechnic Institute of Brooklyn, Brooklyn, New York, 1965.
2. HALKIN, H., *Finite Convexity in Infinite-Dimensional Spaces*, Proceedings of the Colloquium on Convexity, Edited by W. Fenchel, University of Copenhagen, Mathematical Institute, Copenhagen, 1965.
3. NEUSTADT, L. W., *An Abstract Variational Theory with Applications to a Broad Class of Optimization Problems 1: General Theory*, SIAM Journal on Control, Vol. 4, No. 3, 1966.
4. HALKIN, H., *An Abstract Framework for the Theory of Process Optimization*, Bulletin of the American Mathematical Society, Vol. 72, No. 4, 1966.
5. HALKIN, H., *Optimal Control as Programming in Infinite Dimensional Spaces*, CIME: Calculus of Variations, Classical and Modern, Edizioni Cremonese, Roma, 1966.
6. HALKIN, H., and NEUSTADT, L. W., *General Necessary Conditions for Optimization Problems*, Proceedings of the National Academy of Sciences, Vol. 56, No. 4, 1966.
7. CANON, M., CULLUM, C., and POLAK, E., *Constrained Minimization Problems in Finite Dimensional Spaces*, SIAM Journal on Control, Vol. 4, No. 3, 1966.

8. Halkin, H., *Nonlinear Nonconvex Programming in an Infinite-Dimensional Space*, Mathematical Theory of Control, Edited by A. V. Balakrishnan and L. W. Neustadt, Academic Press, New York, 1967.
9. Gamkrelidze, R. V., *On Some Extremal Problems in the Theory of Differential Equations with Applications to the Theory of Optimal Control*, SIAM Journal on Control, Vol. 3, No. 1, 1965.
10. Dubovitskii, A. Ya., and Milyutin, A. A., *Extremum Problems in the Presence of Restrictions*, USSR Computational Mathematics and Mathematical Statistics, Vol. 5, No. 1, 1965.
11. Lobry, C., *Étude Géométrique des Problèmes d'Optimisation en Présence de Constraintes*, Université de Grenoble, 1967.