Proof. By 3.3 the proof can be reduced to that of the following elementary fact: *If* (a_n) *is a non-decreasing sequence of positive numbers such that* $\sum_n a_n < \infty$, *then there are integers* $j_1 < j_2 < j_3 \ldots$ *with* $(j_n) \approx (a_n)$.

Remark. The method based on the invariance of diametral dimension and Lemma 3.1 is insufficient to settle conjecture 1 in the case of power series spaces of infinite type. In fact, if

$$a_n = \exp(\exp(\exp n)) \quad \text{and} \quad (b_n) = (a_2, a_2, a_4, a_4, a_6, a_6, \ldots),$$

then $(b_n) \prec (a_n)$, but $(b_n) \approx$ to no subsequence of (a_n).

Added in proof. A. Using the method of section 3 one can show that Conjecture 3.0 is true for all the spaces listed in Preliminaries, p. 427.

B. An interesting Dragilev-type theorem concerning the spaces $l^2[a_n^{-1/p}]$ (not necessarily nuclear) is given in [11].

References

[1] C. Bessaga and A. Pełczyński, *On embedding of nuclear spaces into the space of all infinitely differentiable functions on the real line* (Russian), Dokl. Akad. Nauk SSSR 134 (1960), p. 745-748.

[2] M. M. Dragilev, *On the canonical form of bases in the spaces of analytic functions* (Russian), Usp. Mat. Nauk 15 (2) (1960), p. 181-188.

[3] — *On regular bases in nuclear spaces* (Russian), Matem. Sbornik 68 (1965), p. 153-173.

[4] — A. S. Dynin and B. S. Mitiagin, *Criterion for nuclearity in terms of approximative dimension*, Bull. Acad. Polon. Sci. (Sér. sci. math.) 8 (1960), p. 535-540.

[5] B. S. Mitiagin, *The approximative dimension and bases in nuclear spaces* (Russian), Usp. Mat. Nauk 16 (4) (1961), p. 63-132.

[6] Z. Ogrodzka, *On simultaneous extensions of infinitely differentiable functions*, Studia Math. 28 (1967), p. 193-207.

[7] A. Pietsch, *Nukleare lokalkonvexe Räume*, Berlin 1965.

[8] S. Rolewicz, *On spaces of holomorphic functions*, Studia Math. 21 (1961), p. 135-160.

[9] H. G. Tillman, *Randverteilungen analytischer Funktionen und Distributionen*, Math. Zeitschrift 59 (1953), p. 61-83.

[10] V. P. Zaharynta, *On bases and isomorphism of the spaces of analytic functions on convex domains of many variables* (Russian), Теория функций и функциональный анализ 5 (1967), p. 5-12.

[11] — *On quasi-equivalence of bases in finite centers of Hilbertian scales*, Dokl. Akad. Nauk SSSR 130, No 4 (1968), p. 786-788.

Reçu par la Rédaction le 23. 1. 1968

STUDIA MATHEMATICA, T. XXXI. (1968)

A general maximum principle for optimization problems

by

M. ALTMAN (Warszawa)

Halkin and Neustadt [10] have presented a very general maximum principle for a mathematical programming problem over an arbitrary set. This principle includes and extends all the most important necessary conditions for optimization problems. Two basic features are characteristic for the Halkin-Neustadt principle. The first one is the "dual form" of the necessary criterion. Roughly speaking, by the "dual form" we mean some relation (equality or inequality) in a proper conjugate space. A classical example is Lusternik's theorem ([11], p. 339): The problem is to find an $\hat{x}$ such that $F(\hat{x}) = \min\{F(x) | P(x) = 0\}$, $F\colon X \to R^1$, $P\colon X \to Y$, X, Y being Banach spaces. The necessary condition in the "dual form" is $F'(\hat{x}) = \overline{P'(\hat{x})}l$ for some $l \in Y$, where $\overline{P'(\hat{x})}$ denotes the adjoint of the Fréchet derivative $P'(\hat{x})$ of P at $\hat{x}$. However, the essential role is played by the necessary condition in the "primary form" $\max\{F'(\hat{x})y | P'(\hat{x})y = 0, \|y\| = 1\} = 0$, which is more suitable for constructing computational algorithms (see [3]-[5]). This observation seems to be of general character. The second characteristic feature of the Halkin-Neustadt principle is that the assumptions as well as the additional equality constraints are closely connected with the method of proof based on Brouwer's fixed-point theorem. This method may be considered as a further development of the method of Canon, Cullum and Polak [6], and Halkin [9].

The purpose of this paper is to present a more general and stronger maximum principle. In comparison with the Halkin-Neustadt principle it has the following properties. The necessary condition has a "primary form" (which implies the "dual form") and is stronger than the Halkin-Neustadt necessary condition. The additional equality constraints are of a much more general nature. Finally, the abstract framework for the optimization process under consideration is much more general and does not depend on the application of Brouwer's fixed-point theorem. Moreover, it is impossible to apply Brouwer's fixed-point theorem under such general hypotheses. Besides, it should be emphasized that no continuity assumptions are made in the general case.

1. PROBLEM. Given a set L and real-valued functions $F = f_0, f_i$, $i = 1, \ldots, m$, defined on L, and a map $P: L \to Y$, Y being an arbitrary linear space. The problem is to find an element $z_0 \in L$ which minimizes F on the set of all $z \in L$ that satisfy the constraints $P(z) = 0$ and $f_i(z) \leqslant 0$ for $i = 1, \ldots, m$.

2. ASSUMPTIONS. There exist convex sets H_i $(i = 0, 1, \ldots, m+1)$ in a real linear space S, real-valued functions h_i $(i = 0, 1, \ldots, m)$ defined on the corresponding H_i and a linear (affine) map P': $H_{m+1} \to Y$ and a set $M \subset S$ with the following properties:

(i) $0 \in \prod_{i=0}^{m+1} H_i$ and $M \subset \prod_{i=0}^{m+1} H_i$;

(ii) M is convex;

(iii) the functions h_i $(i = 0, 1, \ldots, m)$ are convex and $h_i(0) = 0$ for each i;

(iv) for each y of M such that $P'y = 0$ there exist: a sequence of y_k of M, a sequence of positive numbers δ_k convergent to zero, and a sequence of elements $\theta(y_k, \delta_k) \in L$ such that

$$\text{(a)} \qquad \lim_{k\to\infty} \frac{f_i[\theta(y_k, \delta_k)] - f_i(z_0) - h_i(\delta_k y_k)}{\delta_k} \leqslant 0 \quad \text{for each } i = 0, 1, \ldots, m,$$

$$\text{(b)} \qquad \lim_{k\to\infty} \{h_i(y_k) - h_i(y)\} \leqslant 0 \quad \text{for each } i = 0, 1, \ldots, m,$$

$$\text{(c)} \qquad P[\theta(y_k, \delta_k)] = 0 \quad \text{for } k = 1, 2, \ldots$$

3. THE MAXIMUM PRINCIPLE (the primary form). *Let z_0 be a solution of the preceding problem satisfying the above hypotheses. Then*

$$(1) \qquad \inf\{\sigma | h_i(y) \leqslant \sigma,\ i = 0, 1, \ldots, m;\ P'y = 0,\ y \in M\} \geqslant 0,$$

or $\min\{\sigma | h_i(y) \leqslant \sigma,\ i = 0, 1, \ldots, m;\ P'y = 0, y \in M\} = 0$ *if* $0 \in M$ [1].

Proof. Suppose that there exist an element $y \in M$ and a negative σ such that $P'y = 0$ and $h_i(y) \leqslant \sigma$ for $i = 0, 1, \ldots, m$. Then there exist a sequence of elements y_k and a sequence of numbers δ_k satisfying the assumptions of section 2. Hence it follows from (b) that $h_i(y_k) \leqslant \sigma/2$ for sufficiently large k, and thus, by (i) and (ii), the inequality $\delta_k^{-1} h_i(\delta_k y_k) \leqslant \sigma/2$ holds also for sufficiently large k. In virtue of (a) we have

$$\frac{f_i[\theta(y_k, \delta_k)] - f_i(z_0)}{\delta_k} \leqslant \frac{\sigma}{4} \quad \text{for sufficiently large } k.$$

[1] This is equivalent to $\min\{h_0(y) | h_i(y) < 0,\ i = 1, \ldots, m;\ P'y = 0,\ y \in M\} = 0$, if $0 \in M$ and $[y\colon h_i(y) < 0,\ i = 1, \ldots, m;\ P'y = 0,\ y \in M]$ is not empty.

From the last inequality and from assumption (c) we conclude that z_0 is not an optimum solution.

4. THE MAXIMUM PRINCIPLE (the dual form). *In addition to the assumptions of section 2, let us assume that Y is a linear space such that the set of internal points of the image $P'(M)$ is not empty. If z_0 is a solution of the problem (section 1), then there exist a linear functional l defined on Y and real numbers $c_0, c_1, \ldots, c_m$ such that*

$$(\alpha) \qquad \sum_{i=0}^{m} c_i h_i(y) + l(P'y) \leqslant 0 \quad \textit{for all } y \in M;$$

$$(\beta) \qquad c_0, c_1, \ldots, c_m \textit{ and } l \textit{ do not vanish simultaneously};$$

$$(\gamma) \qquad c_i \leqslant 0 \quad \textit{for } i = 0, 1, \ldots, m;$$

$$(\delta) \qquad c_i f_i(z_0) = 0 \quad \textit{for } i = 1, \ldots, m.$$

In addition, if for a certain subset $\mathscr{J}$ of $\{0, 1, \ldots, m\}$ the point 0 is an internal point (see the Appendix) of H_i for all i of $\mathscr{J}$, and if $0 \in M$, then there exist linear functionals l_i $(i \in \mathscr{J})$ on S such that

$$(\varepsilon) \qquad \sum_{i\in\mathscr{J}} c_i l_i(y) + \sum_{\substack{i=0\\ i\notin\mathscr{J}}}^{m} c_i h_i(y) + l(P'y) \leqslant 0 \quad \textit{for all } y \textit{ of } M$$

and

$$(\eta) \qquad l_i(y) \leqslant h_i(y) \quad \textit{for all } y \textit{ of } H_i \textit{ and every } i \textit{ of } \mathscr{J}.$$

Moreover, if S is a topological linear space and, for some $i \in \mathscr{J}$, 0 is in the interior of H_i and h_i is continuous at 0, then l_i is continuous on S.

Proof. We shall show that relations (α), (β), (γ) and (δ) follow from relation (1), which holds also in the case where the indices $i = 1, \ldots, m$ are replaced by $i \in I(z_0) = [i\colon f_i(z_0) = 0]$. In the product space $R^{m+1} \times Y$ consider the set $W = \big(h_0(y) - \xi_0, h_1(y) - \xi_1, \ldots, h_m(y) - \xi_m, P'y\big)$ for y running over M. It follows from the convexity of h_i and from the linearity of P' that W is convex, since so is M. The set of internal points of W is not empty by assumption and the convex set W does not meet the convex cone $K = [(\xi_0, \xi_1, \ldots, \xi_m, 0)\colon \xi_i < 0, 0 \leqslant i \leqslant m]$ in $R^{m+1} \times Y$ $(0 \in Y)$. Hence, there is a hyperplane through 0 in $R^{m+1} \times Y$ separating W from K, i.e., there exist a linear functional l on Y and real numbers $c_0, c_1, \ldots, c_m$ satisfying the conditions (α), (β), (γ) and (δ). To prove statements (ε) and (η) we shall apply the corollary of the generalized Mazur-Orlicz theorem (see the Appendix). Let us show that if $0 \in M$ and $\mathscr{J}$ is as indicated in the theorem, then there exist linear functionals l_i $(i \in \mathscr{J})$ on S such that (ε) and (η) are satisfied. Let $\mathscr{J} = \{i_1, \ldots, i_k\}$ and put

$$g(y) = -\sum_{\substack{i=0\\ i\neq i_1}}^{m} c_{i_1}^{-1} c_i h_i(y) - c_{i_1}^{-1} l(P'y), \qquad y \in M,$$

assuming that $c_{i_1} \neq 0$. It is evident that $g(y)$ is concave and $g(y) \leqslant h_{i_1}(y)$. Hence, in virtue of the corollary (see the Appendix), there exists a linear functional l_1 on S such that $g(y) \leqslant l_1(y)$ for all y of M and $l_1(y) \leqslant h_{i_1}(y)$ for all y of H_{i_1}. Replacing $g(y)$ by $g_1(y)$, where

$$g_1(y) = \sum_{\substack{i=0\\ i\neq i_1,i_2}}^{m} c_{i_2}^{-1} c_i h_i(y) - c_{i_2}^{-1} l(P'y) - c_{i_2}^{-1} c_{i_1} l_1(y), \quad y \in M,$$

we can show just as before that there exists a linear functional l_2 on S such that $g_1(y) \leqslant l_2(y)$ for all y of M and $l_2(y) \leqslant h_{i_2}(y)$ for all y of H_{i_2}.

In this way the convex functionals h_{i_1} and h_{i_2} in relation (ε) have been replaced by the linear functionals l_1 and l_2, respectively. Successively repeating the above argument we arrive at relations (ε) and (η).

If S is a topological linear space and, for some i, h_i is continuous at 0, then it follows from (η) that $l_i(y) \leqslant 1$ in a neighbourhood of 0. Therefore ([3], p. 447, Lemma 7), l_i is continuous on S.

Remark 1. If Y is a locally convex linear topological space and the interior of the image $P'(M)$ is not empty, then l in relation (α) is a continuous linear functional on Y.

It is clear that the preceding results are valid also in the particular case where the mapping P is determined by n real-valued functions $f_{m+1}, \ldots, f_{m+n}$ and the corresponding linear mapping P' is defined by n linear real-valued functions $h_{m+1}, \ldots, h_{m+n}$ and Y is the linear n-space R^n, i.e.,

$$P(z) = \big(f_{m+1}(z), f_{m+2}, (z), \ldots, f_{m+n}(z)\big), \quad z \in L,$$

anp

$$P'y = \big(h_{m+1}(y), \ldots, h_{m+n}(y)\big), \quad y \in M.$$

We shall now discuss this case under hypotheses sufficient for the assumptions in section 2 to be satisfied.

5. Problem. The problem is the same as in section 1, where the constraint $P(z) = 0$ should be replaced by the equalities $f_j(z) = 0$, $j = m+1, \ldots, m+n$.

6. Assumptions (see [10]). There exist convex sets H_i ($i = 0, \ldots, m+n$) in a real linear space S, real-valued functions h_i ($i = 0, \ldots, m+n$) defined on the corresponding H_i, and a set $M \subset S$ with the following properties:

(i) $0 \in \bigcap_{i=0}^{m+n} H_i$ and $M \subset \bigcap_{i=0}^{m+n} H_i$.

(ii) M is convex.

(iii) The functions h_i are convex for $i \leqslant m$ and linear for $i > m$, and $h_i(0) = 0$ for each i.

(iv) Denote by M_0 the subset of elements y of M such that $h(y) = 0$,

where $h = (h_{m+1}, \ldots, h_{m+n})$, and for $y \in M_0$ let A_y be an n-simplex with vertices $y_j \in M$ ($j = 1, \ldots, n+1$) such that y is an interior point of A_y and $h(y_i)$ are in general position, i.e., the vectors $h(y_j - y_{n+1})$ for $j = 1, \ldots, n$, are linearly independent. The set of all such A_y for $y \in M_0$ will be denoted by $A(M_0)$. We shall now assume that for every y of M_0 there exist an A of $A(M_0)$, a set $D \subset (0, 1)$ and a mapping θ from $A \times D$ into L such that

(a) $$0 \in \overline{D},$$

(b)
$$\lim_{\substack{\delta\to 0\\ \delta\in D}} \frac{f_i[\theta(y, \delta)] - f_i(z_0) - h_i(\delta y)}{\delta} \begin{cases} = 0 & \text{for each } i = m+1, \ldots, m+n,\\ \leqslant 0 & \text{for each } i = 0, \ldots, m \end{cases}$$

uniformly over A, and

(c) for every $\delta \in D$ and every $j = m+1, \ldots, m+n$ the mappings $f_j[\theta(0, \delta)]$ are continuous on A.

The continuity on A is to be understood with respect to the ordinary finite-dimensional Euclidean topology on A.

7. The Maximum Principle (the primary form). *Let z_0 be a solution of the problem of section 5 satisfying the hypotheses of section 6. Then*

$$\inf\{\sigma | h_i(y) \leqslant \sigma,\ i = 0, \ldots, m;\ h_i(y) = 0,\ i = m+1, \ldots, m+n, y \in M\} \geqslant 0$$

or

(2) $$\min\{\sigma | h_i(y) \leqslant \sigma,\ i = 0, \ldots, m;\ h_i(y) = 0,\ i = m+1, \ldots, m+n, y \in M\} = 0 \text{ if } 0 \in M.$$

Proof. Using the results of section 4, it is sufficient to show that for each $\bar{y}$ of M_0 the assumptions of section 4 are satisfied, where the mapping P should be replaced by f, $P(z) = f(z) = \big(f_{m+1}(z), \ldots, f_{m+n}(z)\big) \in R^n$ and P' by $h = (h_{m+1}, \ldots, h_{m+n})$. First let us prove assumption (c) (section 2). Choose a sequence of positive numbers ε_k converging to 0 as k tends to infinity. Then it follows from (a) and (b) that there exists a sequence of positive numbers $\delta_k \in D$ converging to 0 such that for each $k = 1, 2, \ldots$ we have

(3) $$\left| \frac{f_i[\theta(y, \delta_k)] - f_i(z_0)}{\delta_k} - h_i(y) \right| \leqslant \frac{\varepsilon_k}{n} \quad \text{for } i = m+1, \ldots, m+n$$

uniformly over the n-simplex A corresponding to $\bar{y}$ of M_0. Put $K_k = [x\colon \|x\| \leqslant \varepsilon_k,\ x \in R^n]$. Since $\bar{y}$ is an interior point of the n-simplex A by assumption, 0 is an interior point of the n-simplex $h(A) \subset R^n$ with vertices $h(y_j)$, $j = 1, \ldots, n+1$; y_j are the vertices of A. Therefore,

$K_k \subset h(A) \subset R^n$ for almost all k. Let us define the mappings $z^k(x) = \theta(y, \delta_k)$ for $x = h(y)$, $y \in A$ $(k = 1, 2, \ldots)$ and the mappings

$$F_k(x) = \frac{1}{\delta_k} f[z^k(x)], \quad x = h(y), y \in A, f = (f_{m+1}, \ldots, f_{m+n}), \quad \text{for } k = 1, 2, \ldots$$

It follows from (c) that F_k is a continuous mapping from $h(A) \subset R^n$ into R^n and we have, by (3),

$$\|F_k(x) - x\| \leqslant \varepsilon_k \quad \text{for } x \in h(A),\ k = 1, 2, \ldots$$

Hence, for almost all k, $x - F_k(x)$ maps K_k into itself and by Brouwer's fixed-point theorem there exist $x_k \in K_k$ such that $F_k(x_k) = 0$, i.e., $f[\theta(y_k, \delta_k)] = 0$. Since $h(y_k) = x_k \to 0 = h(\bar{y})$ as $k \to \infty$, we have $y_k \to \bar{y}$. But $\bar{y}$ is an interior point of A and h_i $(0 \leqslant i \leqslant m)$ are convex functions; consequently, h_i are continuous at $\bar{y}$. Hence we obtain $h_i(y_k) \to h_i(\bar{y})$ as $k \to \infty$ for $i = 0, \ldots, m$. Thus we have shown that assumptions (b) and (c) of section 2 are satisfied.

Condition (a) of section 2 is already contained in condition (b) of section 6.

8. THE MAXIMUM PRINCIPLE (the dual form). *Let z_0 be a solution of the problem of section 5 satisfying the hypotheses of section 6. Then there exist real numbers $c_0, \ldots, c_{m+n}$ such that*

$$(\alpha) \qquad \sum_{i=0}^{m+n} c_i h_i(y) \leqslant 0 \quad \textit{for all } y \in M;$$

$$(\beta) \qquad \sum_{i=0}^{m+n} |c_i| > 0;$$

$$(\gamma) \qquad c_i \leqslant 0 \quad \textit{if } i \leqslant m;$$

$$(\delta) \qquad c_i f_i(z_0) = 0 \quad \textit{if } i = 1, \ldots, m.$$

In addition, if for some subset $\mathscr{J}$ of $\{0, \ldots, m\}$ the point 0 is an internal point of H_i for all i of $\mathscr{J}$ and if $0 \in M$, then there exist linear functionals l_i $(i \in \mathscr{J})$ on S such that

$$(\varepsilon) \qquad \sum_{i \in \mathscr{J}} c_i l_i(y) + \sum_{\substack{i=0 \\ i \notin \mathscr{J}}}^{m+n} c_i h_i(y) \leqslant 0 \quad \textit{for all } y \textit{ of } M,$$

and

$$(\eta) \qquad l_i(y) \leqslant h_i(y) \quad \textit{for all } y \textit{ of } M \textit{ and every } i \textit{ of } \mathscr{J}.$$

Moreover, if S is a topological linear space, for some i of $\mathscr{J}$, 0 is an interior point of H_i and h_i is continuous at 0, then l_i is continuous on S.

Proof. The proof follows from relation (2) exactly in the same way as the proof given in section 4.

Remark 2. Let us observe that Lusternik's necessary criterion for the conditional minimum mentioned in the introduction can also be included in the general scheme of the maximum principle considered above. For this aim it is sufficient to put $f_i \equiv 0$, $h_i \equiv 0$ for $i = 1, \ldots, m$. Besides, in the case where the constraint of the form $P(z) = 0$ is not involved in the optimization problem, one can put $P \equiv 0$, $P' \equiv 0$. In all these cases the general argument is still valid.

9. Regular directions. In some simple cases of non-linear programming it is convenient to use the notion of feasible directions (see [2]). For the conditional minimum of a functional Lusternik [11] has applied the notion of a tangent direction. Dubovitskiy and Milyutin [7] have used the notion of variations for more general optimization problems. Neustadt [14] and others have considered the first-order convex approximation of non-convex sets. We shall use the notion of a regular direction. This is actually a generalization of the feasible and tangent directions and can be used in place of variations as well as instead of the first-order convex approximation. All these notions have a common nature. However, the regular direction method seems to be more convenient in constructing computational algorithms, especially by using the primary form of the maximum principle in which the relations between the introduced constraints (for differentials or subdifferentials) and the original ones are better exhibited with regard to their role in the optimization process itself. Generalized gradient methods as in [3]-[5] can be applied also in the case of the primary form of the maximum principle. Given a set L, the vector y is said to be a *regular direction* for $z_0 \in L$ if there exist a set D of numbers $\delta \in D \subset (0, 1)$ and a set of elements $o(\delta)$, $\delta \in D$, such that

(r) $0 \in \bar{D}$; and $z_0 + \delta y + o(\delta) \in L$ for every δ of D and $o(\delta)/\delta \to 0$ as $\delta \to 0$.

It is clear that the set M of regular directions for z_0 is a cone and $0 \in M$. If $y \in M$ and $o(\delta) \equiv 0$, then y is called a *feasible direction*. Consider the problem of minimizing the functional f_0 on L subject to additional inequality constraints $f_i(z) \leqslant 0$, $z \in L$, $i = 1, \ldots, m$. Suppose that L has a convex set M of regular directions for the optimal solution z_0. In order to obtain the maximum principle for this problem, it is sufficient to have a differentiability notion in the sense of Milyutin and Dubovitskiy [7], i.e., there exist convex functionals h_i $(i = 0, \ldots, m)$ such that

$$(\mathrm{d}) \qquad \frac{f_i(z_0 + \delta z) - f(z_0)}{\delta} \xrightarrow[\substack{\delta \to 0+ \\ z \to y}]{} h_i(y)$$

or

$$\text{(d')} \qquad \lim_{\substack{\delta\to 0+\\ z\to y}} \frac{f_i(z_0+\delta z)-f(z_0)-h_i(\delta y)}{\delta} \leqslant 0.$$

Then the maximum principle in the primary form is

$$\text{(4)} \qquad \min\{\sigma \mid h_i(y) \leqslant \sigma,\ y \in M,\ i = 0, \ldots, m\} = 0.$$

Thus we see that having the notion of a regular direction one can put

$$\text{(5)} \qquad \theta(y, \delta) = z_0 + \delta y + o(\delta) \in L, \qquad \delta \in D,$$

as a special case of the function $\theta(y, \delta)$ involved in the preceding considerations, where the properties of $\theta(y, \delta)$ and of the constraints are combined (see (a), section 2.). In the case where relation (5) is supposed, assumptions (d) or (d′) concerning the constraints are made separately, i.e., independently of $\theta(y, \delta)$.

10. Dubovitskiy and Milyutin [7] have shown that the optimal control problem can be reduced to the problem of minimizing the functional

$$I(x, u) = \int_{t_0}^{t_1} F(x, u, t)\,dt,$$

subject to the constraints

$$\text{(1)} \qquad g(x) \leqslant 0,$$

$$\text{(2)} \qquad \varphi(u) \leqslant 0,$$

$$\text{(3)} \qquad \frac{dx}{dt} = f(x, u, t), \qquad x(t_0) = x_0,$$

$$\text{(4)} \qquad x(t_1) = x_1$$

under certain regularity hypotheses concerning the functions involved. It follows from [15] that the set L of (x, u) determined by the constraints (3) and (4) possesses a set of regular directions $(\bar{x}, \bar{u})$ defined as solutions of the system

$$\frac{d\bar{x}}{dt} = f_x\bar{x} + f_u\bar{u}, \qquad \bar{x}(t_0) = 0, \qquad \bar{x}(t_1) = 0,$$

provided that a non-degeneracy condition is satisfied.

It follows from [7] that the optimal control problem is contained in the scheme considered in section 9.

11. If there are additional equality constraints $f_i(z) = 0$, $i = m+1, \ldots, m+n$, in the problem considered in section 9, then the convex set M of regular directions should satisfy some additional hypotheses of uniformity replacing the corresponding assumptions in section 5; in relation (5) the convergence $o(\delta)/\delta \to 0$ as $\delta \to 0$ is uniform with respect to y on every n-simplex from the set of $A(M_0)$ defined in section 5. If we replace, in section 5, $\theta(y, \delta)$ by (5), the primary form of the maximum principle is given by (2) of section 7 and the dual form by (α)-(η) in section 8. Neustadt [14] reduces the optimal control problem to that formulated in section 5 by using the notion of the first-order convex approximation instead of the uniformly regular direction discussed above. He uses there the dual form of the maximum principle.

APPENDIX

Definition ([8], p. 410). If M is the subset of the linear space X, then a point $p \in M$ is called an *internal point* of M if, for each $x \in X$, there exists an $\varepsilon > 0$ such that $p + \delta x \in M$ for $|\delta| \leqslant \varepsilon$.

Basic separation theorem ([8], p. 412). *Let M and N be disjoint convex subsets of a linear space X, and let M have an internal point. Then there exists a non-zero linear functional f which separates M and N.*

The generalization of the Mazur-Orlicz theorem on inequalities. *Let W be a convex subset of the linear space X and let $0 \in W$ be an internal point of W. Given a subset Z of W and a real-valued function $c(x)$, $x \in Z$, defined on Z, and a real-valued convex function $p(x)$, $x \in W$, defined on W and $p(0) = 0$. In order that there exists a linear functional $f(x)$, $x \in X$, defined on X and such that $c(x) \leqslant f(x)$ for $x \in Z$ and $f(x) \leqslant p(x)$ for $x \in W$, the condition*

$$\text{(6)} \qquad \sum_{i=1}^{n} t_i c(x_i) \leqslant p\Big(\sum_{i=1}^{n} t_i x_i\Big)$$

for arbitrary $x_i \in Z$,

$$n = 1, 2, \ldots, t_i \geqslant 0 \ \text{with} \ \sum_{i=1}^{n} t_i \leqslant 1$$

is necessary and sufficient.

Proof. The necessity of condition (6) is obvious. We shall show that condition (6) is also sufficient. We define the following two subsets of $X \times R^1$:

$$U = [(x, \lambda):\ x \in W,\ \lambda \geqslant p(x)],$$

$$V = \Big[(z, \mu):\ z = \sum_{i=1}^{n} t_i x_i,\ \mu < \sum_{i=1}^{n} t_i c(x_i),\ x_i \in Z\Big],$$

where $n = 1, 2, \ldots$; x_i and t_i are arbitrary elements and numbers such that $x_i \in Z$, $t_i \geqslant 0$ and $\sum_{i=1}^{n} t_i \leqslant 1$.

It is easily verified that U and V are convex subsets and that $(x, \lambda) = (0, 1)$ is an internal point of U. Further, it follows from (6) that U and V are disjoint. Hence, in virtue of the basic separation theorem, there exist a real number c and a non-zero linear functional $l(x, \lambda)$ on $X \times R^1$ such that $l(z, \mu) \leqslant c \leqslant l(x, \lambda)$ whenever $(z, \mu) \in V$ and $(x, \lambda) \in U$. It is evident that $(tx, t\lambda) \in U$ (respectively V) whenever $(x, \lambda) \in U$ (respectively V) and $0 < t \leqslant 1$. From this it follows that $c = 0$. Further, $l(x, \lambda) = l_1(x) + a\lambda$, where l_1 is a linear functional on X and a is a real number. If $a = 0$, then $l_1(x) \geqslant 0$ for all x of W, and $l_1 \neq 0$. But this is impossible since $(0, 1)$ is an internal point of W. Thus, since $(0, 1) \in U$, we get $a > 0$. Let us set $f(x) = -a^{-1} l_1(x)$. Since $l_1(x) + ap(x) = l(x, p(x)) \geqslant 0$ for $x \in W$, we have $f(x) \leqslant p(x)$ for x of W. Suppose now that there exists an element x of Z such that $f(x) < c(x)$. Then there is a number b such that $f(x) < b < c(x)$. Hence, $(x, b) \in V$ and $l(x, b) = l_1(x) + ab \leqslant 0$. Thus, expressing $l_1(x)$ in terms of $f(x)$, we get $-af(x) + ab \leqslant 0$. Since a is positive, $b \leqslant f(x)$, which is a contradiction. Consequently, we have proved that $c(x) \leqslant f(x)$ for all x of Z.

Remark 3. In the Mazur-Orlicz theorem ([12], p. 147) the functional p is sublinear, i.e., subadditive and positive-homogeneous. Besides, condition (6) is supposed to be satisfied for arbitrary non-negative numbers t_i. Pertinent to this theorem is also a paper by Milman [13].

COROLLARY. *If the sets $Z \subset W$ are convex, $0 \in Z$ and 0 is an internal point of W, if $c(x)$ is concave on Z, $p(x)$ is convex on W and $c(x) \leqslant p(x)$ for all x of Z, then there exists a linear functional f on X such that $c(x) \leqslant f(x)$ for x of Z and $f(x) \leqslant p(x)$ for x of W.*

The proof of this corollary follows from the generalization of the Mazur-Orlicz theorem.

References

[1] M. Altman, *Stationary points in non-linear programming*, Bull. Acad. Pol. Sci., Série sci. math., astr. et phys., 12 (1964), p. 29-35.

[2] — *A feasible direction method for solving the nonlinear programming problem*, ibidem 12 (1964), p. 43-50.

[3] — *A generalized gradient method for the conditional minimum of a functional*, ibidem 14 (1966), p. 445-451.

[4] — *A generalized gradient method with self fixing step size for the conditional minimum of a functional*, ibidem 15 (1967), p. 19-24.

[5] — *A generalized gradient method for the conditional extremum of a functional*, ibidem 15 (1967), p. 177-183.

[6] M. Canon, C. Cullum and E. Polak, *Constrained minimization problems in finite-dimensional spaces*, SIAM J. on Control 4 (1966), p. 528-547.

[7] A. Ya. Dubovitskiy and A. A. Milyutin, *Extremum problems in the presence of constraints*, Zh. Vychisl. Mat. i Mat. Fiz. 5 (1965), p. 395-453 (in Russian).

[8] N. Dunford and I. T. Schwartz, *Linear operators*, Part I: *General theory*, New York 1958.

[9] H. Halkin, *Non-linear non-convex programming on an infinite-dimensional space*, Mathematical theory of control, edited by A. V. Balakrishnan and L. W. Neustadt, New York and London 1967.

[10] — and L. W. Neustadt, *General necessary conditions for optimization problems*, Proc. Nat. Acad. Sci. 56 (1966), p. 1066-1071.

[11] Л. А. Люстерник и В. И. Соболев, *Элементы функционального анализа*, Москва-Ленинград 1951.

[12] S. Mazur et W. Orlicz, *Sur les espaces métriques linéaires, II*, Studia Math. 13 (1953), p. 137-179.

[13] D. P. Milman, *Separators of non-linear functionals and their linear extensions*, Izv. Acad. Nauk SSSR, Ser. Mat., 27 (1963), p. 1189-1211 (in Russian).

[14] L. W. Neustadt, *An abstract variational theory with application to a broad class of optimization problems II*, SIAM J. on Control 5 (1967), p. 90-137.

[15] L. S. Pontryagin, V. G. Boltyanskiĭ, R. V. Gamkrelidze and E. F. Mischenko, *The mathematical theory of optimal processes*, New York 1962.

INSTYTUT MATEMATYCZNY POLSKIEJ AKADEMII NAUK
INSTITUTE OF MATHEMATICS OF THE POLISH ACADEMY OF SCIENCES

Reçu par la Rédaction le 28.2.1968