RM-77-13

ON THE THEORY OF MAX-MIN

A. Propoi

March 1977

PREFACE

Max-Min problems play an important role in the theory of nondifferentiable optimization methods. First, the solution of a Max-Min problem makes it possible to evaluate upper and/or lower bounds of the objective function for some optimization problem under uncertainty conditions and to elaborate the decision which guarantees the optimum objective function value in these uncertainty conditions. Second, dual methods of decomposition for solving large-scale optimization problems require the solution of a Max-Min problem. Third, many problems of game theory reduce to Max-Min (Min-Max) problems. In this paper the specific of Max-Min problems is investigated and the solution methods which realize the successive approximation of optimal solution, both for external and internal problems, are discussed.

On the Theory of Max-Min

Abstract

An approach to the solution of max-min problems which takes into account the peculiarities of both the external (max) and the internal (min) operations is considered.

The solution allows us to develop a set of methods for the solution of different kinds of max-min problems, including multi-stage max-min problems, max-min problems with linked constraints, etc.

1. *Introduction*

The theory of the Max-Min problem plays an important role in making optimal decisions under conditions of uncertainty [(1),(2)]. In a majority of cases, however, only a solution method for the "external" maximization problems were developed [(1),(3)].

This paper considers a class of methods for the solution of Max-Min problems, which realizes the successive approximation to optimal solutions both for "external" and "internal" problems and develops the approach introduced in [(4),(5)].

2. *Statement of the Problem*

We shall consider the following problem.

Problem 1.

Find $x^* \in X$, for which

$$\max_{x \in X} \min_{y \in Y} \phi(x,y) = \min_{y \in Y} \phi(x^*,y) = \omega^- , \qquad (1)$$

where X and Y are compact sets in the euclidean spaces R^n and R^m respectively, and the function $\phi(x,y)$ is supposed to be defined and continuous in D x Y, where D is some domain (open connected set $D \supset X$.

Along with problem (1), we shall consider the problem

$$\min_{y \varepsilon Y} \max_{x \varepsilon X} \phi(x,y) = \omega^{+} \tag{1a}$$

and a game $\Gamma(\phi,X,Y)$, where player I chooses $x \varepsilon X$, player II chooses $y \varepsilon Y$ and the payoff is the value of function $\phi(x,y)$.

Generally, $\omega^{-} \leq \omega^{+}$; the situation of equilibrium when $\omega^{-} = \omega^{+}$ is possible if the function ϕ and the sets X and Y possess some convexity properties [6].

In any case, the solution of Problem (1) and (finding) the optimal solution x* allows us to determine the low guaranteeing value of the objective function, that is

$$\phi(x^{*},y) \geq \omega^{-} \quad \text{for all } y \varepsilon Y \quad . \tag{2}$$

Further, only problem (1) will be considered here. (Problem 1a can be investigated in a similar way).

Let us introduce the function

$$\Phi^{*}(x) = \min_{y \varepsilon Y} \phi(x,y) \tag{3}$$

and the set

$$Y^{*}(x) = \{y | \phi(x,y) = \Phi^{*}(x), y \varepsilon Y)\} \quad , \tag{4}$$

which are defined for all $x \varepsilon D$.

In the majority of papers devoted to the solution methods of (1), only the maximization methods for the function $\Phi^{*}(x)$ are considered (see, e.g. [(1),(3)]. In these methods, it is necessary for each x^{ν} (ν is the number of an iteration, $\nu = 1,2,3,\ldots$) to determine either the whole set $Y^{*}(x^{\nu})$ [(1),(2)] or at least one element of this set [3]; that is, for each ν it is necessary to find the global minimum of the problem (2). In a general case, this requires a large amount of computation.

In this scheme, the successive approach to the solution of (1) is realized only on the variable **x**. Evidently, this is not the only way for solving (1). Thus, the problem of developing solution methods for problem (1) arises, which uses both the successive approach on vectors x^{ν} as well as the approach on "functions" (set mapping) $Y^{\nu}(x)$ (which will be defined later). The class of these methods are in some sense complete, including the approach to the solution of the problem (1), both for "internal" and "external" problems in (1).

3. *Extension of the Problem*

We shall replace the original Problem 1 by the following problem.

Problem 2.

Given: functional class $\mathcal{Y}$, find a vector x^* and a function $Y^*(x)$ (or the sequence of functions $\{Y_i^*(x)\}$) in $\mathcal{Y}$, which yields

$$\sup_{x \in X} \inf_{Y(x) \in \mathcal{Y}} \phi(x,Y(x)) = \phi(x^*,Y^*(x)) \text{ or } \lim_{i \to \infty} \phi(x^*,Y_i^*(x)) \tag{5}$$

Problem (5) needs some remarks.

Definition 1. The sequence of functions $Y_i^*(x) \in \mathcal{Y}$ $(i = 1,2,3,\ldots)$ is called the solution of the problem

$$\inf_{Y(x) \in \mathcal{Y}} \phi(x,Y(x)) \tag{6}$$

in a given class of functions $\mathcal{Y}$, if for any fixed $x \in X$ the limit of the sequence $\phi(x,Y_i^*(x))$, $i \to \infty$, exists and

$$\lim_{i \to \infty} \phi(x,Y_i^*(x)) = \Phi^*(x) \quad . \tag{7}$$

To illustrate Problem 2 and Definition 1, let us consider the problem

$$\inf_{Y(x) \in \mathcal{Y}} \int_X \phi(x,Y(x))dx \quad , \tag{8}$$

$\phi(x,Y(x)) \geq 0$ for all $x \in X$ and $Y(x) \in \mathcal{Y}$.

It is clear that if the sequence $\{Y_i^*(x)\}$ is a solution of (6) in the sense of Definition 1, then it is also a solution of (8):

The solution of (8) is defined in an ordinary way: the sequence of functions $Y^*_i(x) \in \mathcal{Y}$ $(i = 1,2,3,\ldots)$ is a solution of (8) in a given class of functions $\mathcal{Y}$, if

$$\lim_{i\to\infty} \int_X \phi(x, Y^*_i(x))dx \le \int_X \phi(x, Y(x))dx .$$

for all $Y(x) \in \mathcal{Y}$.

$$\lim_{i\to\infty} \int_X \phi(x, Y^*_i(x))dx = \inf_{Y(x)\in\mathcal{Y}} \int_X \phi(x, Y(x))dx \quad .$$

On the other hand, evidently if a sequence $\{Y^0_i(x)\}$, $Y^0_i(x) \in \mathcal{Y}$ $(i = 1,2,\ldots)$ is a solution of (8), then each function $Y^0_i(x)$ differs from $Y^*_i(x)$ only on a set of points $x \in X$ with measure zero.

Definition 1 immediately implies the following assertion:

<u>Lemma 1</u>. *The upper bound of Problem 2 is achieved and coincides with the upper bound of Problem 1:*

$$\sup_{x\in X} \inf_{Y(x)\in\mathcal{Y}} \phi(x, Y(x)) = \max_{x\in X} \Phi^*(x) = \omega^- .$$

In this sense, Problems 1 and 2 are equivalent. On the other hand, they differ in the solution of the "inner" problem; that is, in the case of Problem 2, the strategy of player II is evaluated not in separate points $x \in X$ (as in (1)), but is characterized on the whole, for all $x \in X$.

Such an extension of Problem 1 has some remarkable properties and allows us to simplify, in many cases, the solution of Problem 1 through taking more completely into account the specifics of the optimal strategy of $Y^*(x)$ of the player II.

In particular, Problem 2 possesses saddle-point properties without any convexity assumption, that is, the game $\Gamma(\phi, X, \mathcal{Y})$ with strategies $x \in X$, $Y(x) \in \mathcal{Y}$ has a saddle-point solution under only the continuity assumption [(6),(4)].

However, in this paper we shall not consider these properties of Problem 2, but shall investigate the interrelations between Problems 1 and 2.

4. *The Class of Feasible Functions*

Before introducing the definition of the class of feasible functions $\mathcal{Y}$, let us consider the properties of the original Problem 1.

Under the assumption given above, the following given assertions are true [(1),(2)]:

1. The function $\Phi^*(x)$ is continuous in D.

2. $Y^*(x)$ is an upper-semicontinuous point-to-set mapping, that is for any neighborhood $\omega(Y^*(x_0))$ of the set $Y^*(x_0)$ a positive $\delta > 0$ exists, such that $Y^*(x_0) \subset \omega(Y^*(x_0))$, if only

$$|x - x_0| \leq \delta \quad , \qquad x, x_0 \in D \quad .$$

If $Y^*(x)$ is a single valued function for the point $x = x_0$, e.g. the set $Y^*(x_0)$ contains only one element $\{y^*(x_0)\} = Y^*(x_0)$, then 2 implies the continuity of the function $y^*(x)$ at the point x_0.

3. Let the function $\phi(x,y)$ have a gradient $\partial\phi(x,y)/\partial x$ continuous with respect to x and y at the point x_0 for all y. Then the function $\Phi^*(x)$ has a directional derivative $g \in R^n$ at this point given by

$$\frac{\partial\Phi^*(x)}{\partial g} = \lim_{\varepsilon\to 0^+} \frac{\Phi^*(x_0+\varepsilon g)-\Phi^*(x_0)}{\varepsilon} = \min_{y\in Y^*(x)} \left(\frac{\partial\phi(x_0,y)}{\partial x}, g\right) \quad . \tag{9}$$

Bearing in mind these properties of $Y^*(x)$, let us introduce the set of feasible functions $\mathcal{Y}$.

<u>*Definition 2*</u>. [4,7] A finite set of domains $\{D_1,\dots,D_N\}$ defines the decomposition of the domain D, if

1. the boundary of each domain D_i $(i = 1,\dots,n)$ is piece-wise smooth (e.g. consists of a finite number of manifolds);
2. $D_i \cap D_j = \emptyset$, $i \neq j$;
3. $D_i \subseteq D$ $(i = 1,\dots,N)$;
4. $\bigcup_{i=1,\dots,N} \bar{D}_i \supseteq D$

   ($\bar{D}$ is the closure of D).

Let us denote

$$M_{i_1 \ldots i_r} = \bigcap_\nu \bar{D}_i \cap \bar{D} \qquad (\nu = 1,\ldots,r) \tag{10}$$

where $M_{i_1 \ldots i_r}$ is either an empty set or a connected smooth manifold (Fig. 1).

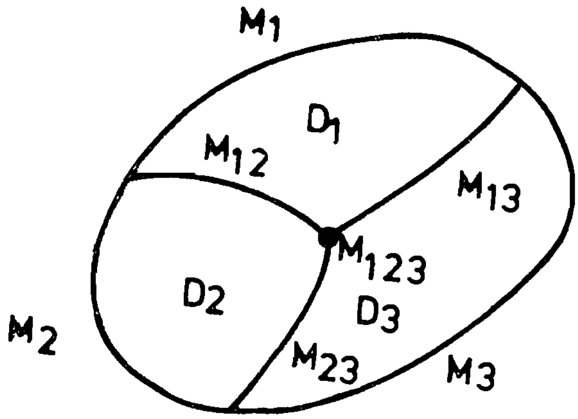

Figure 1.

Definition 3. [4] A multivalued function $Y(x)$ is feasible (that is, belongs to $\mathcal{Y}$), if

1. A decomposition of the domain D exists, given by the function $Y(x)$ in such a way that $Y(x)$ coincides on each D_i with a function $y_i(x)$, defined and continuously differentiable on a domain $\tilde{D}_i \supset \bar{D}_i$;

2. $Y(x) \subset Y$ for all $x \in D$;

3. For any point $x_0 \in M_{i_1 \ldots i_r}$, $Y(x_0)$ is the set of values of $y_{i_\nu}(x_0)$, where $y_{i_\nu}(x_0)$ is a limit of functions $y_{i_\nu}(x)$, when $x \to x_0$ within the domain D_{i_ν} $(\nu = 1,\ldots,r)$ e.g.

   $$y_{i_\nu}(x) \to y_{i_\nu}(x_0) \quad , \quad x \to x_0 \quad , \quad x \in D_{i_\nu} \quad ;$$

4. The following equalities are true:
   $$\phi(x_0, y_{i_\nu}(x_0)) = \text{const for all } \nu = 1,\ldots,r \tag{11}$$
   and $x_0 \in M_{i_1 \ldots i_r}$

Remarks: For some points $x_0 \in M_{i_1 \ldots i_r}$ it is possible that

$$\left| \frac{dy_i(x)}{dx} \right| \to \infty \quad , \qquad x \to x_0 \quad , \qquad x \in D_i \quad .$$

Then it is assumed that

$$\left| \left(\frac{\partial \phi(x, y_i(x))}{\partial y} \right)^T \frac{dy_i(x)}{dx} \right| \leq c < \infty \quad , \quad x \to x_0 \quad , \quad x \in D_i \quad ,$$

It is also possible that Y(x) is a continuous function y(x) on some manifolds $M_{i_1 \ldots i_r}$, that is

$$y_{i_\nu}(x_0) = y(x_0) \quad , \quad x_0 \in M_{i_1 \ldots i_r} \quad , \quad (\nu = 1, \ldots, r) \quad .$$

Then it is assumed that at point x_0 the derivative of $y(x_0)$ is discontinuous.

The Definitions 2,3 allows us to introduce a function

$$\Phi(x) = \phi(x, Y(x)) \quad , \qquad Y(x) \in Y \quad . \tag{12}$$

Let W be a domain in $R^n \times R^m$; $W \supset D \times Y$. The following properties of the function $\Phi(x)$ are true:

1. If $\phi(x,y)$ is continuous in W, then $\Phi(x)$ is a continuous function in D.
2. If $\phi(x,y)$ is continuously differentiable in W, then $\Phi(x)$ is continuously differentiable in D_i $(i = 1, \ldots, N)$ and

$$\frac{d\Phi(x_0)}{dx} = \frac{\partial \phi(x_0, y_i(x_0))}{\partial x} + \left[\frac{dy_i(x_0)}{dx} \right]^T \frac{\partial \phi(x_0, y_i(x_0))}{\partial y} \; , x_0 \in D_i \tag{13}$$

where

$$\frac{d\Phi}{dx} = \left[\frac{\partial \Phi}{\partial x^j} \right] \quad ; \qquad \frac{\partial \phi}{\partial y} = \left[\frac{\partial \phi}{\partial y^k} \right] \quad ; \qquad \frac{dy_i}{dx} = \left[\frac{dy_i^k}{dx^j} \right]$$

$(j = 1,\ldots,n; \quad k = 1,\ldots,m; \quad i = 1,\ldots,N)$

Let $x_0 \in M_{i_1 \ldots i_r}$ and define a cone at the point x_0:

$$K_i(x_0) = K(x_0,X) = \{g | g \in R^n;\ x_0 + \varepsilon g \in D_i;\ 0 < \varepsilon < \bar{\varepsilon}\} \quad , \quad (14)$$

where $\bar{\varepsilon}$ is some positive number depending on x_0 and D_i.

The closure $\bar{K}_i(x_0)$ of the set (14) is usually called the cone of feasible directions of the set D_i at the point x_0 (see, for example, [2]).

Definition 4. A vector g at a point $x_0 \in M_{i_1 \ldots i_r} = \bigcap_{\nu=1}^{n} \bar{D}_{i_\nu} \cap \bar{D}$ is directed to a domain D_i, if $g \in K_i(x_0)$.

If $x_0 \in D_i$, then $K_i(x_0)$ coincides with all of the space R^n and any vector $g \in R^n$ is directed to D_i at this point x_0.

Let us also define

$$T_{i_1 \ldots i_r}(x_0) = \bigcap_{\nu} \bar{K}_{i_\nu}(x_0) \quad (\nu = 1,\ldots,r) \quad , \quad (15)$$

$T_{i_r \ldots i_r}(x_0)$ is a hyperplane, tangent to $M_{i_1 \ldots i_r}$ at the point x_0.

3. If $x_0 \in M_{i_1 \ldots i_r}$, then the function $\Phi(x)$ is differentiable in any direction $g \in R^n$ and

$$\frac{\partial \Phi(x_0)}{\partial g} = \left(\frac{\partial \phi(x_0, y_i(x_0))}{\partial x}, g\right) + \left(\frac{\partial \phi(x_0, y_i(x_0))}{\partial y}, \frac{dy_i(x_0)}{dx} g\right) \quad (16)$$

$$g \in K_i(x_0), \ i \in \{i_1,\ldots,i_r\} .$$

If $x_0 \in M_{i_1 \ldots i_r}$ and $g \in T_{i_1 \ldots i_r}(x_0)$, then for all $\nu = 1,\ldots,r$:

$$\frac{\partial \Phi(x_0)}{\partial g} = \left(\frac{\partial \phi(x_0, y_{i_\nu}(x_0))}{\partial x}, g\right) + \left(\frac{\partial \phi(x_0, y_{i_\nu}(x_0))}{\partial y}, \frac{dy_{i_\nu}(x_0)}{dx} g\right) \quad (17)$$

The proof of the statements 1 - 3 follows directly from Definitions 2,3 and properties of the function $\phi(x,y)$.

As the equality (16) is also true for $g \in \bar{K}_i(x_0)$ (due to (11), (12) and the definition of $\bar{K}_i(x_0)$), then it follows from (16) and (17) that $\partial\Phi(x_0)/\partial g$ is a continuous function of g (x_0 is fixed) and may be a discontinuous function of x_0 (g is fixed).

Definition 5. Let functions $y_1(x)$ and $y_2(x)$ be defined and have continuous derivatives of orders k, $k = 0,1,\ldots,\ell$ in D. Let a function $\phi(x,y)$ be defined in $W \supset D \times Y$ and have continuous derivatives of orders s, $s = 0,1,\ldots,\ell'$, $\ell' \geq \ell$.

If the relation

$$\frac{d^k\Phi(x)}{dx^k} = \frac{d^k\phi(x,y_1(x))}{dx^k} = \frac{d^k\phi(x,y_k(x))}{dx^k}$$

holds for all $k = 0,1,\ldots,\ell$, and some $x \in D$, then the functions $y_1(x)$ and $y_2(x)$ are called ℓ-equivalent at the point $x \in D$.

The ℓ-equivalence of functions $y_i(x)$ in direction g (or in some domain D) is defined in a similar way.

Let $Y(x)$ be a function in $\mathcal{Y}$, then it follows from Definition 5 that any piecewise continuous functions $y(x)$, $y(x) \in Y(x)$, $x \in D$ are 0-equivalent in D (in accordance with (11)).

Functions $y_{i_\nu}(x)$ $(\nu = 1,\ldots,r)$ of Definition 3 are 1-equivalent in directions $g \in T_{i_1 \ldots i_r}(x_0)$ at $x_0 \in M_{i_1 \ldots i_r}$, because the equalities (16) hold true for them.

The idea of function equivalence permits us to simplify the class of function $\mathcal{Y}$ considerably and reduce it actually to the class of piecewise smooth functions. For example, the multivalued function $Y^1(x)$ on Fig. 1 is equivalent to the piecewise smooth function $Y^2(x)$, two-valued only at the point $x_0 = M_{12}$.

However, the question arises as to what measure Problems 1 and 2 are similar. The answer to this question has two approaches. The first consists of investigating the proximity of Problems 1 and 2, when the properties of the function ϕ and the sets X and Y are fixed. In the second approach, the conditions on $\{\phi,X,Y\}$ are searched for, such that $Y^*(x) \in \mathcal{Y}$, where $Y^*(x)$ is defined in (3).

We shall consider shortly both approaches. One can verify the following statement.

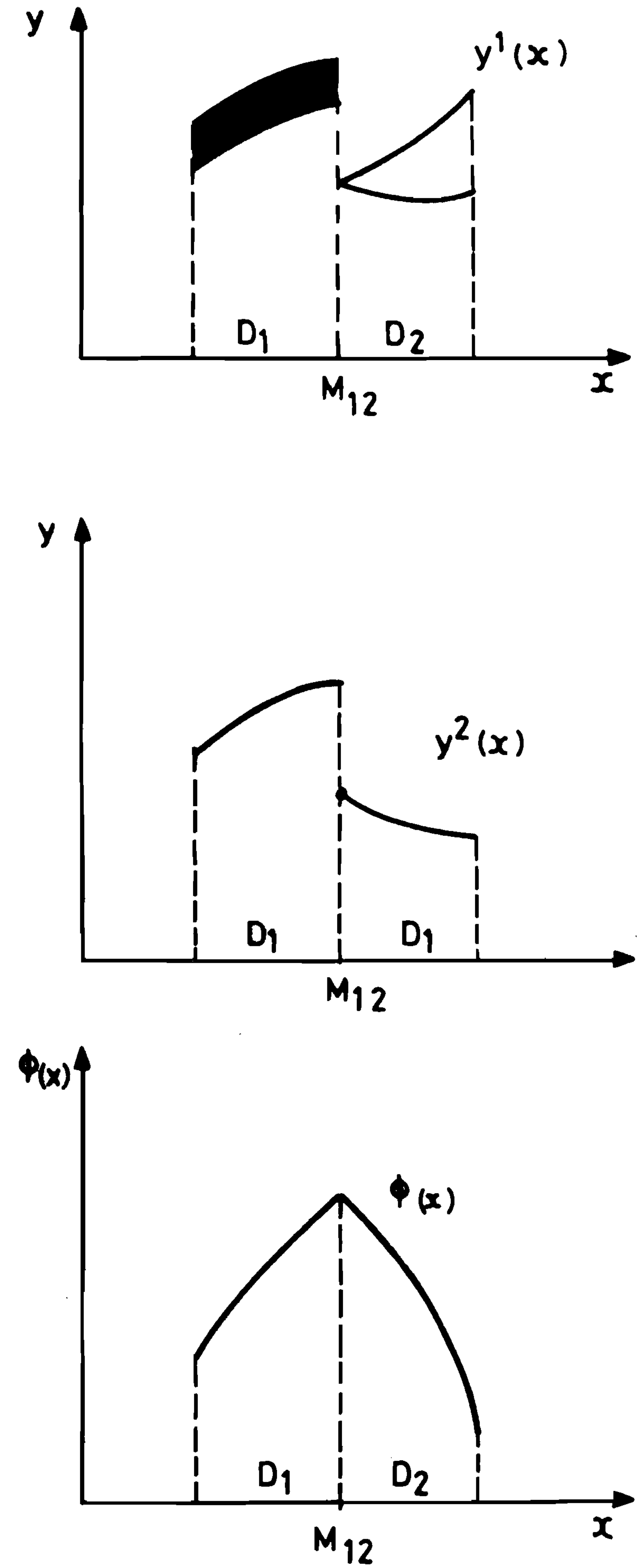

Figure 2.

Theorem 1. Let the function ϕ be continuous in $W \supset D \times Y$. Then a sequence of functions $Y_i^*(x) \subset \mathcal{Y}$ $(i = 1,2,\ldots)$ exists, such that for each $x \in D$:

$$\lim_{i\to\infty} \phi(x,Y_i^*(x)) = \Phi^*(x) .$$

Proof: Since ϕ is a continuous function with respect to $y \in Y$ for all $x \in D$ and Y is a compact set, a solution $Y^*(x)$ of the problem

$$\min_{y \in Y} \phi\ (x,y)$$

exists for each $x \in D$ (by Weierstrass's theorem).

Therefore, according to (2) and (3):

$$\Phi^*(x) = \phi(x,Y^*(x)) , \quad x \in D .$$

However, the multivalued function $Y^*(x)$ may not belong to the class of functions $\mathcal{Y}$; as only the property 2 of section 4 is true.

Theorem 1 states that the functions $Y(x)$ from the class of functions $\mathcal{Y}$ can approximate $Y^*(x)$ with an arbitrary accuracy.

For proving this, let us single out some subset $\tilde{D} \subset D$.

We assume that $\rho(\tilde{D}) > \varepsilon > 0$, where $\rho(\tilde{D})$ is the diameter of the set D.

By appropriate choice of $\tilde{D}$ we can always reduce the behavior of $Y^*(x)$ to the following four cases:

1) $Y^*(x) \in \mathcal{Y}$ for all $x \in \tilde{D}$;

2) $Y^*(x)$ is a multivalued function (point-to-set mapping) for each $x \in \tilde{D}$;

3) $Y^*(x) = y^*(x)$ is a single valued continuous function (point-to-point mapping), which has discontinuous derivative for all $x \in \tilde{D}$;

4) $Y^*(x)$ induces an infinite decomposition on $\tilde{D}$, that is, an infinite number of domains $D_i^* \subset D$ exists, such that $Y^*(x) = Y_i^*(x)$, $x \in D_i^*$, and $y_i^*(x)$ is a continuously differentiable function in D_i^*.

Let us consider each case separately.

1) If $Y^*(x) \in \mathcal{Y}$, then the assertion of the theorem reduces to

$$\phi(x, Y^*(x)) = \Phi^*(x), \quad x \in \tilde{D} .$$

2) If $Y^*(x)$ is a point-to-set mapping for all $x \in \tilde{D}$, then a function $y^*(x)$, defined and continuously differentiable on $\tilde{D}$, can be chosen, for which

$$y^*(x) \in Y^*(x) , \quad x \in \tilde{D} ;$$

$$\Phi^*(x) = \phi(x, Y^*(x)) = \phi(x, y^*(x)) , \quad x \in \tilde{D}$$

Evidently, $y^*(x) \in \mathcal{Y}$. In the case considered this can always be done due to the properties of $Y^*(x)$ and the class of functions $\mathcal{Y}$ (see also Definition 5).

3) If $Y^*(x) = y^*(x)$ is a continuous function for all $x \in \tilde{D}$, then one can choose smooth functions $y_i^*(x) \in \mathcal{Y}$, such that for any given $\varepsilon > 0$

$$|y^*(x) - y_i^*(x)| < \varepsilon , \quad i \geq N , \quad x \in \tilde{D} ,$$

holds, where $|y|$ is an appropriate norm of a function y.

Using the continuity properties of the function ϕ, one can easily obtain the statement of the theorem for this case.

4) If an infinite number of domains $D_j^* \subset \tilde{D}$ exists, we can define using the properties of $Y^*(x)$ for any given $\varepsilon > 0$ a set D_δ, such that

$$D_j^* \subset D_\delta \subset \tilde{D} , \quad j \geq M$$

$$\rho(D_\delta) < \delta$$

and

$$Y^*(x) \subset \omega_\varepsilon(Y^*(x_o)) , \quad x, x_o \in D_\delta ,$$

where $\omega_\varepsilon(Y^*)$ is an ε-neighborhood of the set Y^*.

Using the last relation, one can define a sequence of functions $Y_i^*(x) \in \mathcal{Y}$ on D_δ in such a way that, again,

$$|Y^*(x) - Y_i^*(x)| < \varepsilon , \quad i \geq N , \quad x \in D_\delta .$$

Since X is a compact set and $\rho(\tilde{D}) > \varepsilon > 0$, the finite number of such sets $\tilde{D}$ constitutes the covering of X. This completes the proof.

From this theorem it follows that for any $\varepsilon > 0$ we find a function $Y^*_\varepsilon(x) \in Y$ and a point $x \in D$, such that

$$0 \leq \phi(x, Y^*_\varepsilon(x)) - \min_{y \in Y} \phi(x,y) \leq \varepsilon \tag{18}$$

$$\phi(x_0, Y^*_\varepsilon(x_0)) = \min_{y \in Y} \phi(x_0, y) \tag{18a}$$

for all $x \in D$ and fixed $x_0 \in D$ (ε-optimal solution of Problem (6)).

Thus, if $\phi(x,y)$ is continuous in D, then Problem 2 approximates Problem 1 with respect to "internal" operation with arbitrary accuracy while the solutions of Problems 1 and 2 coincide with respect to "external" operations.

If stronger conditions on $\phi(x,y)$ are imposed, then $Y^*(x) \in Y$ and Problems 1 and 2 become completely equivalent. Thus, the following statement is true.

Theorem 2. *Let the second order derivatives*

$$\frac{\partial^2 \phi(x,y)}{\partial x \partial y} = \left[\frac{\partial^2 \phi(x,y)}{\partial x^s \partial y^j}\right]; \quad \frac{\partial^2 \phi(x,y)}{\partial y^2} = \left[\frac{\partial^2 \phi(x,y)}{\partial y^k \partial y^j}\right]$$
$$(k,j = 1,\ldots,m; \; s = 1,\ldots,n)$$

of the function ϕ *be continuous in* $W \supset D \times Y$ *and let the matrix*

$$\frac{\partial^2 \phi(x,y)}{\partial y^2}$$

be nonsingular at $x \in D_i$ $(i = 1,\ldots,N)$.

Then $Y^*(x) \in Y$ *and*

$$\frac{dy_i(x)}{dx} = -\left[\frac{\partial^2 \phi(x,y_i(x))}{\partial y^2}\right]^{-1} \frac{\partial^2 \phi(x,y_i(x))}{\partial x \partial y} \tag{19}$$

if $x \in D_i$, $Y^*(x) \in \overset{0}{Y}$ *(*$\overset{0}{Y}$ *is the interior of the set* Y*).*

The proof of the Theorem follows from the implicit function theorem.

If the conditions of Theorem 2 are not true, then, in general, $Y^*(x)$ doesn't belong to $\mathcal{Y}$. Let us consider two examples.

Example 1. Let $\phi(x,y) = y^3 - 3xy$; $x,y \in R^1$, $x \geq 0$, $y \geq 0$. Here $\phi''_{yy} = 0$ at point $(0,0)$ and

$$Y^*(x) = \sqrt{x}\ , \quad x \geq 0\ ; \quad \Phi^*(x) = -2x\sqrt{x}\ .$$

Evidently, $dY^*(x)/dx \to \infty$ if $x \to 0^+$, but $d\Phi^*(x)/dx \to 0$, $x \to 0^+$.

Supposing $Y^*(x) = y_1(x) = \sqrt{x}$, $x \in D_1 = \{x \geq 0\}$, $Y^*(x) = y_2(x) = -\sqrt{x}$, $x \in D_2 = \{x \leq 0\}$, one obtains $Y^*(x) \in \mathcal{Y}$.

Example 2. Let $\phi(x,y) = [2y - (y^2+3x^2)^{\frac{1}{2}}]^{2n}$, $x,y \in R^1$, $n \geq 1$. Here ϕ''_{xy}, ϕ''_{yy} are infinite at point $\{0,0\}$, but ϕ'_x, ϕ'_y exist and are continuous. For this example $Y^*(x) = |x|$ and is thus single-valued and continuous. At the same time, the function $dy^*(x)/dx$ is discontinuous at point $x = 0$. However, if we let $Y^*(x) = y_1(x) = x$, $x \in D_1 = \{x \geq 0\}$, $Y^*(x) = y_2(x) = -x$; $x \in D_2 = \{x \leq 0\}$, then $Y^*(x) \in \mathcal{Y}$.

The examples show, that for almost practically all interesting cases, the function $Y^*(x)$ will belong to the class $\mathcal{Y}$ (taking into account the remarks relating to Definition 3 and the Definition 5).

Let us make a last remark. It is known that any continuous functions can be uniformly approximated by infinitely differentiable function. As the function $\Phi^*(x) = \min_{y \in Y} \phi(x,y)$ is continuous, an infinitely differentiable function $\tilde{\Phi}(x)$, exists such that

$$\tilde{\Phi}(x) - \min_{y \in Y} \phi(x,y) \leq \varepsilon \text{ for all } x \in D\ .$$

If $\tilde{\Phi}(x)$, $\phi(x,y)$ are smooth functions, then it seems one can find a smooth function $\tilde{y}(x)$ in D, such that $\tilde{\Phi}(x) = \phi(x,\tilde{y}(x))$ for all $x \in D$.

This is not true. Let us consider an example.

Example 3. Find

$$\min_x \max_y (x-y)^2\ , \quad |x| \leq 1\ , \quad |y| \leq 1\ .$$

Evidently,

$$\Phi^*(x) = \max_{|y|\le 1} (x-y)^2 = \begin{cases} (x+1)^2 , & x \ge 0 \\ (x-1)^2 , & x \le 0 \end{cases}$$

$$Y^*(x) = \begin{cases} -1 , & x \ge 0 \\ 1 , & x \le 0 \end{cases} .$$

The function $\Phi^*(x)$ is continuous in its domain of definition, $Y^*(x)$ is multivalued at the point $x = 0$. The minimal value $\Phi^*(x)$, $|x| \le 1$ is equal to 1 and achieved at the point $x^* = 0$. At the same time, any continuous curve $y(x)$, which is defined in the square $|x| \le 1$, $|y| \le 1$, intersects the straight line $y = x$ and, consequently, $\min_{|x|\le 1} (x-y(x))^2 = 0$ for any continuous functions $y(x)$, $|y(x)| \le 1$, $|x| \le 1$ (Fig. 3).

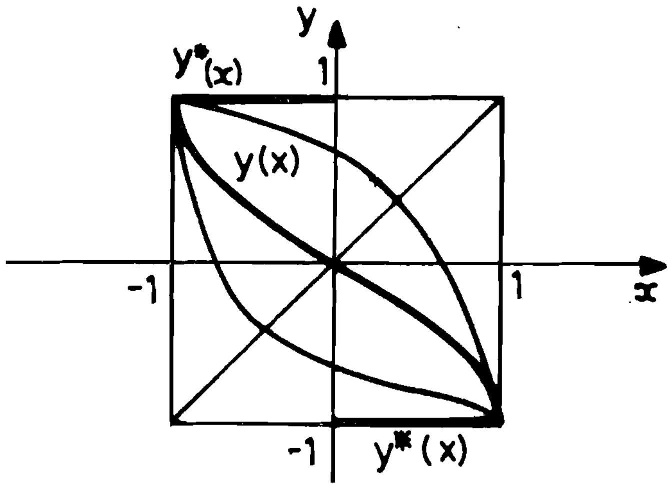

Figure 3.

Thus, the class Y of piecewise smooth functions, which have been introduced in this section, is not only sufficient in some sense, but also is necessary for replacing Problem 1 by Problem 2.

5. *Optimality Conditions*

Let us fix some function $Y(x) \in Y$ and consider the problem

$$\max_{x \in X} \Phi(x) = \max_{x \in X} \phi(x, Y(x)) \quad . \tag{20}$$

Let x* be a solution of (20); define the cone of feasible variations at the point x*.

$$K_0(x^*) = K(x^*,X) = \{g \mid x^* + \varepsilon g \in X,\ 0 < \varepsilon < \bar{\varepsilon}\} \tag{21}$$

Let

$$N_{i_1 \cdots i_2} = \bigcap_{\nu} M_{i_\nu} \cap M_{i_1 \cdots i_r} \qquad \nu = 1,\ldots,r \tag{22}$$

where M_{i_ν} is an external boundary of D_{i_ν} (Fig. 1).

Using conventional reasoning in the theory of mathematical programming, one can prove the following statement.

Theorem 3. *Let $x^* \in N_{i_1 \cdots i_r}$ be a solution of (20). Then*

$$\delta\Phi\left[\frac{\partial\phi}{\partial g}\right] = \left(\frac{\partial\phi(x^*,y_i(x^*))}{\partial x}, g\right) + \left(\frac{\partial\phi(x^*,y_i(x^*))}{\partial y}, \frac{dy_i(x^*)}{dx} g\right) \le 0 \tag{23}$$

for all $g \in \bar{K}_0(x^*) \cap \bar{K}_i(x^*)$, $i \in \{i_1,\ldots,i_r\}$.

Here $y_i(x^*)$ is the value of $Y(x)$, if $x \to x^*$, $x \in D_i$.

In particular, if x* is a point only of M_j, then (23) is true for all $g \in \bar{K}_0(x^*)$. If $x^* \in M_{i_1 \cdots i_r}$, then (23) is true for all $g \in \bar{K}_{i_\nu}(x^*)$ $(\nu = 1,\ldots,r)$, and if $x^* \in D_i$, then (23) is true for all $g \in R^n$ and

$$\frac{\partial\phi(x^*,y_i(x^*))}{\partial x} + \left[\frac{dy_i^*(x^*)}{dx}\right]^T \frac{\partial\phi(x^*,y_i(x^*))}{\partial y} = 0 \quad .$$

Now let x_0 be an arbitrary point of X and $Y^*(x) \in \mathcal{Y}$ is optimal (ε-optimal) for Problem (6).

Theorem 4. *If $\partial\phi(x,y)/\partial y$ exists and is continuous in $W \supset D \times Y$ and the boundary of the set Y is piecewise smooth, then*

$$\left(\frac{\partial\phi(x_0,y_i^*(x_0))}{\partial y}, \frac{dy_i^*(x_0)}{dx} g\right) = 0$$

where $x_0 \varepsilon M_{i_1 \ldots i_r}$; $g \varepsilon \bar{K}_i(x_0)$, $i \varepsilon \{i_1, \ldots, i_r\}$ *and* $y_i^*(x_0)$ *is the value of* $Y^*(x_0)$ *at the point* x_0, $x \to x_0$, $x \varepsilon D_i$.

Proof: In accordance with (18a), let us specify the (ε-optimal) solution $Y^*(x_0)$ of (6) in such a way that

$$\phi(x_0, y_i^*(x_0)) = \min_{y \varepsilon Y} \phi(x_0, y) \quad . \tag{24}$$

The point $y_i^*(x)$ can move, when $x \to x_0$, $x \varepsilon D_i$, either inside of the set Y, or on its boundary.

If $y_i^*(x_0)$ is interior to Y, then, by virtue of (24):

$$\frac{\partial \phi(x_0, y_i^*(x_0))}{\partial y} = 0 \tag{25}$$

If the point $y_i^*(x_0)$ is on the boundary of Y, then, evidently vectors

$$\delta y_i^*(x_0) = \frac{dy_i^*(x_0)}{dx} g,$$

for all feasible vectors g generate the linear manifold $T(y_i^*(x_0))$, tangent to Y at the point $y_i^*(x_0)$. In this case the gradient $\frac{\partial \phi(x, y_i(x_0))}{\partial y}$ will be orthogonal to $T(y_i^*(x_0))$ by virtue of the optimality of $y_i^*(x_0)$; that is

$$\left(\frac{\partial \phi(x_0, y_i^*(x_0))}{\partial y} , \delta y_i^*(x_0) \right) = 0 \tag{26}$$

for all $\delta y_i^*(x_0) \varepsilon T(y_i^*(x_0))$.

Finally, it is possible, that $T(y_i^*(x_0))$ reduces to a single point $y_i^*(x_0) = \text{const}$ at some vicinity of x_0. In that case,

$$\frac{dy_i^*(x_0)}{dx} = 0 \quad .$$

For a more formal proof of the last two equalities, let us suppose that the set Y is given by the system of inequalities:

$$f_j(y) \leq 0 \quad (j = 1,\ldots,s)$$

where the functions $f_j : R^m \to R^1$ are continuously differentiable.

Let $J(y)$ be the set of active constraints at the point y, i.e.

$$J(y) = \{j \mid f_j(y) = 0, j = 1,\ldots,s\}$$

and suppose that the gradients

$$\frac{\partial f_j(y^*)}{\partial y}, \qquad j \in J(y^*)$$

are linearly independent at the point y*. Then the optimality conditions for the problem (24) can be written as

$$\frac{\partial \phi(x_0, y_i^*(x_0))}{\partial y} = \sum_{j=1}^{s} \lambda_j(x_0) \frac{f_j(y_i^*(x_0))}{\partial y} \tag{27}$$

where the Lagrange multipliers $\lambda_j(x_0)$ satisfy the conditions

$$\lambda_j(x_0) f_j(y_i^*(x_0)) = 0 \quad ; \quad \lambda_j(x_0) \geq 0 \quad , \quad f_j(y_i^*(x_0)) \leq 0$$
$$(j = 1,\ldots,s)$$

If $J(y_i^*(x_0)) = \emptyset$, i.e. $y_i^*(x_0)$ is an interior point of the set Y, then all $\lambda_j(x_0) = 0$ and (25) is true.

Now let

$$f_j(y_i^*(x_0)) = 0 \quad , \quad j \in J(y_i^*(x_0)) \neq \emptyset \tag{28}$$

differentiate both parts of the equality (28) in the direction $g \in \bar{K}_i(x_0)$. It is assumed, that $x_0 \in M_{i_r \ldots i_r}$ or $x_0 \in D_i$

$i \in \{i_1 \cdots i_r\}$. In any case,

$$\left(\frac{\partial f_j(y_i^*(x_0))}{\partial y} , \frac{dy_i^*(x_0)}{dx} g \right) = 0 \quad , \qquad j \in J(y_i^*(x_0)) \quad . \qquad (29)$$

If the rank of the matrix

$$A(y_i^*(x_0)) \;=\; \left\| \frac{\partial f_j(y_i^*(x_0))}{\partial y_k} \right\| \quad ; \quad j \in J(y_i^*(x_0)), \; k = 1,\ldots,m$$

is equal to m, then from (29) it follows that

$$\frac{dy_i^*(x_0)}{dx} g = 0 \quad , \qquad g \in \bar{K}_i(x_0) \quad .$$

If the rank of the matrix $A(y_i^*(x_0))$ is less than m, then the system of equations (29) has a nonzero solution

$$\frac{dy_i^*(x_0)}{dx} g \neq 0 \qquad (30)$$

Multiplying both parts of the equality (27) from the right by the vector of (30) and considering (29), one can assume that in this case (26) holds. This completes the proof.

Theorems 3 and 4 specify the optimality conditions for problems (20) and (6). Now consider Problem 2.

Theorem 5. *Let $\{x^*, Y^*(x)\}$ be a solution of Problem 2. Then the inequality*

$$\left(\frac{\partial \phi(x^*, y_i^*(x^*))}{\partial x} , g \right) \leq 0 \qquad (31)$$

holds for all $g \in \bar{K}_0(x^*) \cap \bar{K}_i(x^*)$, *where* $y^*(x) \to y_i^*(x^*)$, $x \to x^*$, $x \in D_i$, $x^* \in M_{i_1 \ldots i_r}$ *or* $x^* \in N_{i_1 \ldots i_r}$, $i \in \{i_1, \ldots, i_r\}$.

If $x^* \in D_i$, then the inequality (31) is replaced by the equality.

The proof of this theorem follows from Theorems 3 and 4. To connect the optimality conditions of Problem 2 stated in Theorem 5, with optimality conditions of Problem 1 [(1),(2)], let us prove the following assertion.

Theorem 6. Let $x_0 \in \bar{D}_i$ *and* $Y(x)$ *be an arbitrary function of* Y. *Then*

$$\min_{y(x_0) \in Y(x_0)} \left[\left(\frac{\partial \phi(x_0, y(x_0))}{\partial x}, g \right) + \left(\frac{\partial \phi(x_0, y(x_0))}{\partial y}, \frac{dy(x_0)}{dx} g \right) \right] =$$

$$= \left(\frac{\partial \phi(x_0, y_i(x_0))}{\partial x}, g \right) + \left(\frac{\partial \phi(x_0, y_i(x_0))}{\partial y}, \frac{dy_i(x_0)}{dx} g \right) \qquad (32)$$

where $g \in \bar{K}_i(x_0)$, $Y(x) \to y_i(x_0)$, $x \to x_0$, $x \in D_i$, $y(x_0) \in Y(x_0)$.

Proof: When $x_0 \in D_i$ $(i = 1, \ldots, N)$ the equality (32) holds trivially, because in this case $Y(x_0) = y_i(x_0)$.

Now let $x_0 \in M_{ij}$. The manifold M_{ij} is the set of x, satisfying the relation

$$\phi(x, y_i(x)) = \phi(x, y_j(x)) \quad ,$$

or

$$F(x) \equiv \phi(x, y_i(x)) - \phi(x, y_j(x)) = 0 \quad .$$

The plane tangent to M_{ij} at the point x_0, is the set of g, satisfying

$$\left(\frac{dF(x_0)}{dx}, g \right) = 0$$

and the gradient $dF(x_0)/dx$ is directed to the side of increase of $F(x)$.

If $g \in K_j(x_0)$, then

$$\left(\frac{dF(x_0)}{dx}, g\right) > 0 \quad ; \tag{33}$$

if $g \in K_i(x_0)$, then

$$\left(\frac{dF(x_0)}{dx}, g\right) < 0 \quad . \tag{34}$$

From (33), (34) and (15), one can obtain (32). The case $x_0 \in M_{i_1 \cdots i_r}$ can be considered in a similar way.

Using Theorems 3, 6 and Definition 4, one can prove that the equality

$$\frac{\partial \Phi(x_0)}{\partial g} = \min_{y \in Y^*(x_0)} \left(\frac{\partial \phi(x_0, y)}{\partial x}, g\right) = \left(\frac{\partial \phi(x_0, y_i^*(x_0))}{\partial x}, g\right) \tag{35}$$

holds, where $Y^*(x_0)$ is defined in (3), $y_i^*(x)$ is any of the 1-equivalent values of $Y^*(x_0)$ in the domain D_i, $x_0 \in \bar{D}_i$.

Now Theorem 6 can be restated as follows.

<u>Theorem 7</u>. *Let x^*, $Y^*(x)$ be a solution of Problem 1. Then*

$$\min_{y \in Y^*(x^*)} \left(\frac{\partial \phi(x^*, y)}{\partial x}, g\right) \leq 0$$

for all $g \in \bar{K}_0(x^*)$.

The optimality conditions of Theorem 7 are given, for example, in [2].

6. *Conclusion*

The results given above allow us to develope methods which realize the approach $\{x^\nu, Y^\nu(x)\} \to \{x^*, Y^*(x^*)\}$, $\nu \to \infty$ both for external and internal operations of Max-Min problems. This permits us to take into consideration the specific features of the problems and thus to develop efficient methods of their solution. Evidently, these methods incorporate the usual scheme $\{x^\nu, Y^*(x^\nu)\} \to \{x^*, Y^*(x^*)\}$.

The author wishes to thank R. Mifflin for many helpful comments and suggestions on this subject.

References

[1] Danskin, J.M., The Theory of Max-Min, with Applications, *SIAM Journal on Applied Mathematics*, 14 (1966) 641-664.

[2] Demjanov, V.F. and V.N. Malozemov, Introduction to Min-Max, J. Wiley, New York (1974) (English translation).

[3] Shor, N.Z., Generalized Gradient Methods for Minimization of Nondifferentiable Functions and Its Applications to Mathematical Programming, *Ekonomika i Math Methody*, XII, 2(1976) (in Russian).

[4] Propoi, A.I., To the Theory of Max-Min Problems, *Zhurnal vychisl. math. i math. phyziki*, 11, 1(1971) (in Russian).

[5] Evtushenko, Yu.G., Some Local Properties of Min-Max Problems, *Zhurnal vychisl. math. i math. phyziki*, 14, 1(1974) (in Russian).

[6] Luce, R.D. and H. Raiffa, *Games and Decisions*, J. Wiley, New York (1957).

[7] Berkovitz, L.D., Necessary Conditions for Optimal Strategies in a Class of Differential Games and Control Problems, *J. SIAM Control*, 5, 1(1965).