Optimization Methods and Software, 2001, vol. 16, pp. 193–230
Reprints available directly from the publisher
Photocopying permitted by license only

GENERAL LAGRANGE-TYPE FUNCTIONS IN CONSTRAINED GLOBAL OPTIMIZATION PART I: AUXILIARY FUNCTIONS AND OPTIMALITY CONDITIONS

[a] Yu.G. EVTUSHENKO[1], [b]A.M. RUBINOV and [a]V.G. ZHADAN

[a] Computing Centre of Russian Academy of Sciences,
40 Vavilov Str., 117967 Moscow GSP-1, Russia;
[b] School of Information Technology and Mathematical Sciences
University of Ballarat, Victoria 3353, Australia

The paper contains some new results and a survey of some known results related to auxiliary (Lagrange-type) functions in constrained optimization. We show that auxiliary functions can be constructed by means of two-step convolution of constraints and the objective function and present some conditions providing the validity of the zero duality gap property. We show that auxiliary functions are closely related to the so-called separation functions in the image space of the constrained problem under consideration. The second part of the paper (see Evtushenko et al., General Lagrange-type functions in constrained global optimization. Part II: Exact Auxiliary functions. Optimization Methods and Software) contains results related to exact auxiliary functions.

Keywords: Constrained optimization; Auxiliary function; Exact auxiliary function; Separation function; Convolution function; Zero duality gap

1 INTRODUCTION

In recent years global optimization has attracted attention of many researchers. One of the main reasons for this to happen is that the global optimization appears, in a huge variety of forms, in practical live and in pure theoretical study. In 1991 a new "Journal of Global Optimization" started to publish numerous results in this field. Among the published books we mention the first book "Towards global optimization" by Dixon and Szego [5]. An impressive number of books and papers have been published in this field later. The publications devoted to global optimization combined entirely new approaches with generalization of well-known local results. In this paper we develop the generalization of Lagrange saddle-point approach to global optimization. It contains new results and a survey of some results related to the so-called auxiliary (Lagrange-type) functions in mathematical programming.

Consider the following mathematical programming problem $P(f,g)$:

> *find the global minimum of a function f subject to inequality constraints $g(x) \le 0$, $x \in X$, where X is a metric space, $g(x) = (g^1(x), \ldots, g^m(x))$, f and g^i are real-valued functions defined on X.*

[1]Corresponding author.

(We can consider problems with equality constraints as a special case of problem $P(f,g)$. Indeed, each equality constraint $g^i(x)=0$ can be presented in the inequality form: $g^i(x)\le 0$, $-g^i(x)\le 0$.)

One of the main approaches in the study of the problem $P(f,g)$ consists in the construction of a unconstrained optimization problem, which is equivalent (in a certain sense) to the initial problem $P(f,g)$. The well-known example of such a problem of unconstrained minimization is the global minimization over X of the Lagrangian (Lagrange function) of $P(f,g)$:

$$\mathcal{L}(x,\omega) := f(x) + \sum_{i=1}^{m} \omega^i g^i(x), \tag{1}$$

where $\omega = (\omega^1,\dots,\omega^m)$ is a certain nonnegative vector. Unfortunately, the minimization of Lagrangian is equivalent to $P(f,g)$ only under suitable assumptions. The Lagrangian allows one to define the so-called dual objective function $\mathcal{Q}$ for the problem $P(f,g)$:

$$\mathcal{Q}(\omega) = \inf_{x\in X} \mathcal{L}(x,\omega),$$

and consider the following problem of unconstrained global maximization (dual problem to $P(f,g)$):

$$\mathcal{Q}(\omega) \longrightarrow \sup \ \text{ subject to } \omega = (\omega^i)_{i=1}^m \ge 0.$$

If the optimal value of this problem is equal to that of $P(f,g)$ (the zero duality gap property holds), then we can find a sequence of vectors ω_k such that the optimal values of the unconstrained problems $\mathcal{L}(x,\omega_k)$ tends to the optimal value of $P(f,g)$.

A similar approach is based on the penalty function of the form:

$$\mathcal{L}^+(x,\omega) := f(x) + \sum_{i=1}^{m} \omega^i g^i_+(x), \tag{2}$$

leads to exact penalization or to duality based on $\mathcal{L}^+$. (Here $a_+ := \max(a,0)$).

There exist many other functions, which can serve for constructing unconstrained problems, which are equivalent (in a certain sense) to $P(f,g)$. Such functions are very useful in the study (both theoretical and numerical) of problem $P(f,g)$. We shall call them *auxiliary* functions for $P(f,g)$.

First, we consider auxiliary functions which do not depend on a parameter. An auxiliary function M of a problem $P(f,g)$ can be constructed as a convolution of the objective function f and constraints $g^i : M(x) = h(f(x), g^1(x),\dots,g^m(x))$. In this paper we shall study mainly two-step convolution functions, that is, functions of the form $M(x) = \phi(f(x)-\eta, \psi(g(x))$. In other words, first we convolute given constraints g^i, $i=1,\dots,m$, to a single constraint $\psi(g(x))$ and then we convolute this new constraint and the objective function. Parameter η is an estimation of a global minimum of a function $f(x)$ on the feasible set.

There is a close connection between auxiliary functions and the so-called separation functions, which have been studied in [21] in the frameworks of the so-called Giannessi scheme. Let $\mathbb{R}^{1+m} = \mathbb{R}\times\mathbb{R}^m$ be the image space [12] of the problem $P(f,g)$. This problem generates the set $\mathcal{T} = \{(u,v)\in\mathbb{R}^{1+m} : u = f(x),\ v = g(x),\ x\in X\}$ (image of the problem $P(f,g)$) and its shifts $\mathcal{T}_\eta = \mathcal{T}(X) - \eta e_0$, where $e_0 = (1,0,\dots,0)$ is a unit vector. In order to describe sufficient (and sometimes necessary) conditions for global minimum of problem $P(f,g)$ it is sufficient to establish that the intersection of the set $\mathcal{T}_\eta$ with certain η and the set $\mathcal{H}^- = \{(u,v)\in\mathbb{R}^{1+m} : u<0,\ v\le 0\}$ is empty. The simplest way to prove the emptiness of the intersection is to find a function h, which separates sets $\mathcal{T}_\eta$ and $\mathcal{H}^-$, that is, $h(u,v) > h(u',v')$ for all $(u,v)\in\mathcal{T}_\eta$ and $(u',v')\in\mathcal{H}^-$. We show that auxiliary functions can be considered as a certain tool for constructing separation functions, hence they can serve for achieving optimality conditions.

Auxiliary functions

$$M(x,\omega) = q(\omega^0 f(x), \omega^1 g^1(x), \dots, \omega^m g^m(x))$$

depending on a certain parameter $\omega = (\omega^0, \omega^1, \dots, \omega^m)$ are very useful in the study of optimization problems. The simplest and well-known examples of auxiliary functions depending on a parameter are Lagrange function (1) (where $q(y^0, y^1, \dots, y^m) = \sum_{i=0}^{m} y^i$) and penalty function (2) ($q(y^0, y^1, \dots, y^m) = y^0 + \sum_{i=1}^{m} y^i_+$). We mention here also augmented Lagrangians (see [17] and references therein).

The auxiliary function $M(x,\omega)$ depending on a parameter ω leads to construction of the *dual function*

$$\mathcal{Q}(\omega) = \inf_{x \in X} M(x,\omega).$$

One of the main problems related to auxiliary functions is to establish the *zero duality gap property*, that is, to describe conditions under which

$$\sup_{\omega \in \Omega} \mathcal{Q}(\omega) = \inf_{x \in X_0} f(x), \tag{3}$$

where $X_0 = \{x \in X : g^1(x) \le 0, \dots, g^m(x) \le 0\}$ is the set of feasible points of $P(f,g)$. If (3) holds, then a solution of the problem $P(f,g)$ can be found by solving a sequence of unconstrained optimization problems. It is very interesting to describe auxiliary functions which provide the zero duality gap for broad classes of problems $P(f,g)$ without any assumptions, related to convexity and differentiability. Some of such classes were described in [19, 20, 13, 22].

In 1967 Eremin [6] and Zangwill [23] introduced a notion of exact penalty (EP) functions. Later, exact augmented Lagrangian (EAL) functions was proposed in [15]. EP and EAL functions proved to be a very valuable tool both in theoretical study of optimization problems and in the development of numerical algorithms because they gave an opportunity to solve nonlinear programming problems by means of a single minimization of an unconstrained function. Subsequently, the properties of these functions were studied in numerous publications. The interested reader is referred to [14, 3, 16, 4] and references therein for a detailed account of the developments in this field.

A quite natural generalization of the notions EP functions and EAL functions is the so-called exact auxiliary (EA) functions. These functions were introduced and studied in [7, 8, 9]. By definition, an EA function for $P(f,g)$ is an auxiliary function M, which enjoys the following property: the set of all unconstrained minima of M (under some rather simple conditions) coincides with the solution set to $P(f,g)$. If an auxiliary function M depends on a parameter ω, then this function is exact if there exists a set of values ω_0 of the parameter such that the auxiliary function $M(\cdot,\omega_0)$ is exact for any ω_0 from this set. It was shown in [7, 8, 9] that the class of EA functions is much broader than classes of EP functions and EAL functions and some of EA functions more convenient for applications than EP and EAL functions. Many new EA functions have been described in these papers. We mention here only exact interior penalty functions, exact Morrison's function, composite exact exterior-interior functions. The simplest EA function is a convolution of a vector consisting of penalization for constraints violation and parameterized objective function.

Like EP functions and EAL functions, general EA functions reduce the study of problem $P(f,g)$ to a single unconstrained minimization. Note that all considerations in [7, 8, 9] were carried out under the assumption that the Lagrangian of the original problem $P(f,g)$ has a saddle point. In this paper we demonstrate that this assumption can be replaced by a substantially weaker one.

Note that auxiliary functions (especially, EA functions) can serve as a base for classifying known numerical methods and designing new ones. However, in this paper we do not consider the numerical aspect of EA functions.

This paper is the first part of the work, consisting of two parts. This part is devoted to separation functions, convolution functions, auxiliary functions. We provide a classification of convolution and auxiliary functions and give several examples. We also consider auxiliary functions depending on the parameters and study the zero duality gap property. This first part contains both new results and a survey of known results (we give them mainly without proofs).

The EA functions are studied in the second part (see [10]).

2 PRELIMINARIES

We shall use the following notations:

- $\mathbb{R} = (-\infty, +\infty)$ is the real line,
- $\bar{\mathbb{R}} = [-\infty, +\infty]$ is the extended real line,
- $\mathbb{R}_{+\infty} = (-\infty, +\infty]$, $\mathbb{R}_{-\infty} = [-\infty, +\infty)$,
- $\mathbb{R}^n$ is the n-dimensional Euclidean space, equipped with the coordinate-wise order relation $\geq$,
- $\mathbb{R}^n_+ = \{x = (x^1, \dots, x^n) \in \mathbb{R}^n : x^i \geq 0 \text{ for all } i\}$ is the non-negative orthant,
- $\mathbb{R}^n_- = \{x = (x^1, \dots, x^n) \in \mathbb{R}^n : x^i \leq 0 \text{ for all } i\}$ is the non-positive orthant,
- $\mathbb{R}^n_{++} = \{x = (x^1, \dots, x^n) \in \mathbb{R}^n : x^i > 0 \text{ for all } i\}$ is the positive orthant,
- $\mathbb{R}^n_{--} = \{x = (x^1, \dots, x^n) \in \mathbb{R}^n : x^i < 0 \text{ for all } i\}$ is the negative orthant,
- $\mathcal{E}^n_+ = \mathbb{R}^n \setminus \mathbb{R}^n_{--}$ is the complement to the negative orthant $\mathbb{R}^n_{--}$,
- $\mathcal{E}^n_{++} = \mathbb{R}^n \setminus \mathbb{R}^n_-$ is the complement to the non-positive orthant $\mathbb{R}^n_-$.

Note that $\mathcal{E}^n_+$ is a closed set and $\mathcal{E}^n_{++}$ is an open set.

Let X be a metric space. Consider a function $f : X \to \mathbb{R}$, a mapping $g : X \to \mathbb{R}^m$ and the minimization problem $P(f,g)$, which is defined by the objective function f and the inequality constraint g:

$$P(f,g) : f(x) \longrightarrow \min \text{ subject to } g(x) \leq 0. \tag{4}$$

Let

$$X_0 := \{x \in X : g(x) \leq 0\} \tag{5}$$

be the set of feasible elements of $P(f,g)$. We assume that X_0 is non-empty. Denote by ρ the optimal value of $P(f,g)$:

$$\rho := \inf\{f(x) : x \in X_0\}.$$

We shall consider the function f and the mapping g on some subsets of X which contain X_0. Let

$$\mathcal{Y} := \{Y : X_0 \subseteq Y \subseteq X\}.$$

Consider the image space $\mathbb{R}^{1+m} = \mathbb{R}\times\mathbb{R}^m$ of $P(f,g)$ and the mapping $(f,g) : X \to\to \mathbb{R}^{1+m}$. This mapping maps a set $Y \in \mathcal{Y}$ onto the image set

$$\mathcal{T}(Y) := \{(u,v) \in \mathbb{R}^{1+m} : u = f(x),\ v = g(x),\ x \in Y\}.$$

The set $\mathcal{T}(Y)$ generates the family $(\mathcal{T}_\eta(Y))_{\eta\in\mathbb{R}}$ of its shifts. By definition

$$\mathcal{T}_\eta(Y) = \{(u-\eta, v) : (u,v) \in \mathcal{T}(Y)\} = \mathcal{T}(Y) - \eta e_0,$$

where $e_0 := (1,0,\ldots,0)$ is the unit vector. Clearly, we have

$$\mathcal{T}_\eta(X_0) \subseteq \mathcal{T}_\eta(Y) \tag{6}$$

for any $\eta \in \mathbb{R}$ and any $Y \in \mathcal{Y}$. Consider also the set

$$\mathcal{H}^- = \{(u,v) \in \mathbb{R}^{1+m} : u < 0,\ v \leq 0\}.$$

Obviously, $\operatorname{cl}\mathcal{H}^- = \mathbb{R}^{1+m}_-$.

Proposition 2.1. *Let $Y \in \mathcal{Y}$. The intersection $\mathcal{T}_\eta(Y) \cap \mathcal{H}^-$ is empty if and only if η is a lower estimate of the value ρ of the problem $P(f,g)$, i.e. $\eta \leq \rho$.*

Proof. Let $\mathcal{T}_\eta(Y)\cap\mathcal{H}^- = \emptyset$. It follows from (6) that $\mathcal{T}_\eta(X_0)\cap\mathcal{H}^- = \emptyset$ as well. Since $g(x) \leq 0$ for all $x \in X_0$, it follows that $f(x) - \eta \geq 0$ for all $x \in X_0$, so $\rho = \inf\{f(x) : x \in X_0\} \geq \eta$. Conversely, if $\rho \geq \eta$, then the inequality $g(x) \leq 0$ implies $f(x) - \eta \geq 0$, so $\mathcal{T}_\eta(X_0)$ does not intersect $\mathcal{H}^-$. Since $\mathcal{T}_\eta(Y) = \mathcal{T}_\eta(X_0) \cup \mathcal{T}_\eta(Y\backslash X_0)$ and $\mathcal{T}_\eta(Y\backslash X_0)$ does not intersect $\mathcal{H}^-$, we have $\mathcal{T}_\eta(Y) \cap \mathcal{H}^- = \emptyset$. $\square$

Corollary 2.1. *The following holds:*

$$\rho = \sup\{\eta : \mathcal{T}_\eta(Y) \cap \mathcal{H}^- = \emptyset\}.$$

Indeed, the optimal value of the problem coincides with the exact upper bound of its lower estimates.

Proposition 2.2. *Let $Y \in \mathcal{Y}$. Assume that $\mathcal{H}^- \cap \mathcal{T}_{\eta_*}(Y) = \emptyset$ and $\mathbb{R}^{1+m}_- \cap \mathcal{T}_{\eta_*}(Y) \neq \emptyset$. Then $\eta_* = \rho$.*

Proof. Due to Proposition 2.1, we have $\eta_* \leq \rho$. Assume that $\eta_* < \rho$. Since $\rho = \sup\{\eta : \mathcal{H}^- \cap \mathcal{T}_\eta(Y) = \emptyset\}$, it follows that there exists $\eta' > \eta_*$ such that $\mathcal{T}_{\eta'}(Y) \cap \mathcal{H}^- = \emptyset$. Let $\delta = \eta' - \eta_* > 0$. Since

$$\mathbb{R}^{1+m}_- \cap \mathcal{T}_{\eta_*}(Y) \neq \emptyset \ \text{ and } \ \mathcal{H}^- \cap \mathcal{T}_{\eta_*}(Y) = \emptyset,$$

we deduce that there exists a vector $\nu \in \mathbb{R}^m_-$ such that $(0,\nu) \in \mathcal{T}_{\eta_*}(Y)$. We have

$$(-\delta,\nu) \in \mathcal{T}_{\eta_*}(Y) - \delta e_0 = \mathcal{T}(Y) - \eta_* e_0 - (\eta' - \eta_*)e_0 = \mathcal{T}(Y) - \eta' e_0 = \mathcal{T}_{\eta'}(Y).$$

On the other hand, $(-\delta,\nu) \in \mathcal{H}^-$. Therefore, $\mathcal{H}^- \cap \mathcal{T}_{\eta'}(Y) \neq \emptyset$, which contradicts the definition of η'. Thus, the result follows. $\square$

In order to check that the intersection of two sets A and B is empty, it is convenient to find a function h (in general, nonlinear), which guarantees disjunctive separation of these sets in the following sense: for all $a \in A$ and for all $b \in B$ either $h(a) < 0 \leq h(b)$ or $h(a) \leq 0 < h(b)$. We assume in the sequel that the solution set

$$X_* := \{x \in X_0 : f(x) = \rho\}$$

of the problem $P(f,g)$ is nonempty.

Proposition 2.3. *Let $Y \in \mathcal{Y}$. Then $\mathbb{R}^{1+m}_{-} \cap \mathcal{T}_\rho(Y) = \{(0, g(x_*)) : x_* \in X_*\}$.*

Proof. Let x_* be a solution of the problem $P(f,g)$. Then $f(x_*) = \rho$, so $(f(x_*) - \rho, g(x_*)) = (0, g(x_*)) \in \mathbb{R}^{1+m}_{-}$. Clearly, $(f(x_*) - \rho, g(x_*)) \in \mathcal{T}_\rho(Y)$. Thus, $(0, g(x_*)) \in \mathbb{R}^{1+m}_{-} \cap \mathcal{T}_\rho(Y)$. Conversely, let $(u,v) \in \mathcal{T}_\rho(Y) \cap \mathbb{R}^{1+m}_{-}$. Since $(u,v) \in \mathcal{T}_\rho(Y)$, it follows that there exists $x_* \in X_0$ such that $v = g(x_*) \leq 0$ and $u = f(x_*) - \rho$. Since ρ is the value of $P(f,g)$, it follows that $u \geq 0$. On the other hand, since $(u,v) \in \mathbb{R}^{1+m}_{-}$, it follows that $u \leq 0$. Thus, $u = 0$, so $f(x_*) = \rho$. We have proved that $x_* \in X_*$, so $(u,v) \in \{(0, g(x_*)) : x_* \in X_*\}$. $\square$

3 OPTIMALITY CONDITIONS VIA SEPARATION FUNCTIONS

We now present conditions for minimum in terms of separation functions. First we consider necessary conditions. Recall that, due to Proposition 2.1, $\mathcal{H}^- \cap \mathcal{T}_\rho(Y) = \emptyset$ for any $Y \in \mathcal{Y}$.

Proposition 3.1. *Let $Y \in \mathcal{Y}$ and let η_* be a number such that:*

1) there exists a function h: $\mathbb{R}^{1+m} \to \bar{\mathbb{R}}$, which strictly separates $\mathcal{H}^-$ and $\mathcal{T}_{\eta_*}(Y)$, that is

$$h(u,v) < 0 \leq h(u',v') \ \ \text{for all} \ \ (u,v) \in \mathcal{H}^-, \quad (u',v') \in \mathcal{T}_{\eta_*}(Y); \tag{7}$$

2) $\mathbb{R}^{1+m}_{-} \cap \mathcal{T}_{\eta_*}(Y) \neq \emptyset$.

Let x_* be a solution of the problem $P(f,g)$. Then $f(x_*)\eta_*$. If h is lower semicontinuous, then $h(0, g(x_*)) = 0$.

Proof. It follows from (7) that $\mathcal{T}_{\eta_*}(Y) \cap \mathcal{H}^- = \emptyset$. Proposition 2.2 demonstrates now that $\eta_* = \rho$. Let x_* be a solution of the problem $P(f,g)$. Then $f(x_*) = \rho = \eta_*$. Assume now that h is lower semicontinuous. Consider a point $(0, g(x_*))$. Due to Proposition 2.3, we conclude that $(0, g(x_*)) \in \mathbb{R}^{1+m}_{-} \cap \mathcal{T}_\rho(Y)$. Since $(0, g(x_*)) \in \mathcal{T}_\rho(Y)$, we have $h(0, g(x_*)) \geq 0$. Since $(0, g(x_*)) \in \mathbb{R}^{1+m}_{-}$, we conclude, by applying lower semicontinuity of h, that $h(0, g(x_*)) \leq 0$. Hence, $h(0, g(x_*)) = 0$. $\square$

We now present a sufficient condition for a minimizer of the problem $P(f,g)$ in terms of separation functions.

Proposition 3.2. *Let $Y \in \mathcal{Y}$. Assume that there exist a function $h : \mathbb{R}^{1+m} \to \bar{\mathbb{R}}$ and a number η_* such that (7) holds. Let $x_* \in X_0$ be a point such that $f(x_*) = \eta_*$. Then x_* is a solution of the problem $P(f,g)$.*

Proof. Assume that there exists $\bar{x} \in X_0$ such that $f(\bar{x}) < f(x_*) = \eta_*$. Then $\bar{u} := f(\bar{x}) - \eta_* < 0$ and $\bar{v} := g(\bar{x}) \leq 0$, so $(\bar{u}, \bar{v}) \in \mathcal{H}^-$. Thus, $h(\bar{u}, \bar{v}) < 0$. On the other hand,

$$(\bar{u}, \bar{v}) \in \mathcal{T}_{\eta_*}(X_0) \subseteq \mathcal{T}_{\eta_*}(Y),$$

so $h(\bar{u}, \bar{v}) \geq 0$. We have a contradiction, which shows that $f(\bar{x}) \geq f(x_*)$ for all $\bar{x} \in X_0$. $\square$

The following version of Proposition 3.2 will be convenient for applications.

Proposition 3.3. *Let $h : \mathbb{R}^{1+m} \to \bar{\mathbb{R}}$ be a function such that $h(u,v) < 0$ for all $(u,v) \in \mathcal{H}^-$. Let $x_* \in X_0$, $\eta_* = f(x_*)$ and*

$$\inf_{x \in Y} h(f(x) - \eta_*, g(x)) \geq 0, \tag{8}$$

where $Y \in \mathcal{Y}$. Then x_* is a solution of the problem $P(f,g)$.

Proof. It follows from (8) that $(f(x)-\eta_*, g(x)) \notin \mathcal{H}^-$ for all $x \in Y$. Let $x \in X_0$, that is, $g(x) \le 0$. Since $(f(x)-\eta_*, g(x)) \notin \mathcal{H}^-$, it follows that $f(x) \ge \eta_* = f(x_*)$. □

Remark 3.1. *Let the function h be lower semicontinuous. Then $h(0, g(x_*)) = 0$. (In such a case the infimum in (8) is attained and equal to zero).*

Indeed, we have

$$(f(x_*)-\eta_*, g(x_*)) = (0, g(x_*)) \in \mathbb{R}^{1+m}_- \cap \mathcal{T}_{\eta_*}(X_0),$$

so $\mathbb{R}^{1+m}_- \cap \mathcal{T}_{\eta_*}(X_0) \ne \emptyset$. Thus, Proposition 3.1 demonstrates that $h(0, g(x_*)) = 0$.

Let h be a function with the following property:

$$u < 0 \Rightarrow h(u,v) < h(0,v) \quad \text{for all} \quad v \in \mathbb{R}^m_-. \tag{9}$$

Then instead of strict separation of sets $\mathcal{H}^-$ and $\mathcal{T}_{\eta_*}(Y)$ we can consider nonstrict separation of $\mathbb{R}^{1+m}_-$ and $\mathcal{T}_{\eta_*}(Y)$.

Proposition 3.4. *Let $Y \in \mathcal{Y}$. Let $h : \mathbb{R}^{1+m} \to \bar{\mathbb{R}}$ be a function such that (9) holds. Assume that*

$$\sup\{h(u,v) : (u,v) \in \mathbb{R}^{1+m}_-\} \le 0 \le \inf\{h(u',v') : (u',v') \in \mathcal{T}_{\eta_*}(Y)\}. \tag{10}$$

Let $x_* \in X_0$ be a point such that $f(x_*) = \eta_*$. Then x_* is a solution of the problem $P(f,g)$.

Proof. Due to Proposition 3.2, it is sufficient to show that $h(u,v) < 0$ for all $(u,v) \in \mathcal{H}^-$. Assume that there exists $(u,v) \in \mathcal{H}^-$ such that $h(u,v) = 0$. Since $u < 0$, it follows that $h(0,v) > h(u,v) = 0$, which is impossible. □

Condition (9) holds if the function $h(u,v)$ is strictly increasing in the first coordinate, that is, $u_1 < u_2 \Rightarrow h(u_1,v) < h(u_2,v)$ for all v. The function $h(u,v) = u + \psi(v)$ is the simplest example of the strictly increasing in the first coordinate function. We shall consider various examples of such functions later. Note that the function $h(u,v) = \max(u, v_1, \dots, v_m)$ does not possess the property (9). Convenient sufficient conditions can be given for problems $P(f,g)$, when a certain regularity condition holds. These conditions exclude constraints of the form $g^i(x) \le 0$, where g^i is a nonnegative function (in such a case the inequality constraints $g^i(x) \le 0$ is equivalent to the equality constraint $g^i(x) = 0$).

Let $\tilde{X}_0 = \{x \in X : g(x) \ll 0\}$. (The inequality $v \ll 0$ means that all coordinates of a vector v are negative.)

Regularity Condition. *A problem $P(f,g)$ is called* **regular** *if f is a lower semicontinuous function and the feasible set $X_0 := \{x \in X : g(x) \le 0\}$ is such that*

$$\operatorname{cl} \tilde{X}_0 = X_0. \tag{11}$$

Proposition 3.5. *Let regularity condition hold. Assume that there exist a function h : $\mathbb{R}^{1+m} \to \bar{\mathbb{R}}$ and a number η_* such that*

$$h(u,v) < 0 \quad \textit{for all} \quad (u,v) \ll 0 \tag{12}$$

and

$$0 \le h(u',v') \quad \textit{for all} \quad (u',v') \in \mathcal{T}_{\eta_*}(Y) \quad \textit{with} \quad Y \in \mathcal{Y}. \tag{13}$$

Let x_* be a point such that $f(x_*) = \eta_*$. Then x_* is a solution of the problem $P(f,g)$.

Proof. Since f is lower semicontinuous and $\operatorname{cl} \tilde{X}_0 = X_0$, it follows that

$$\inf\{f(x) : x \in \tilde{X}_0\} = \min\{f(x) : x \in X_0\}. \tag{14}$$

Assume that x_* is not a solution of $P(f,g)$. Then, due to (11) and (14), there exists a point $\bar{x} \in \tilde{X}_0$ such that $f(\bar{x}) < f(x_*) = \eta_*$. We have $\bar{u} := f(\bar{x}) - \eta_* < 0$ and $\bar{v} := g(\bar{x}) \ll 0$, so $(\bar{u}, \bar{v}) \ll 0$. It follows from (12) that $h(\bar{u}, \bar{v}) < 0$. On the other hand, $(\bar{u}, \bar{v}) \in \mathcal{T}_{\eta_*}(X_0) \subseteq \mathcal{T}_{\eta_*}(Y)$, so (13) implies $h(\bar{u}, \bar{v}) \geq 0$. We have a contradiction. Thus, the result follows. $\Box$

We now present two more sufficient conditions for optimality. In contrast with Proposition 3.3, we do not assume now that x_* is a feasible element.

Proposition 3.6. *Let $Y \in= \mathcal{Y}$ and let $h : \mathbb{R}^{1+m} \to \bar{\mathbb{R}}$ be a function such that $h(u,v) > 0$ for all $(u,v) \notin \mathbb{R}^{m+1}_-$. Assume that there exist $\eta_* \in (-\infty, \rho]$ and $x_* \in Y$ such that*

$$h(f(x_*) - \eta_*, g(x_*)) = \min_{x \in Y} h(f(x) - \eta_*, g(x)) = 0. \tag{15}$$

Then x_* is a solution of $P(f,g)$.

Proof. We have from (15) that $(u,v) = (f(x_*) - \eta_*, g(x_*)) \in \mathbb{R}^{m+1}_-$. Thus, $x_* \in X_0$ and $f(x_*) - \eta_* \leq 0$. Since $\eta_* \leq \rho$ and $f(x) \geq \rho$ for all $x \in X_0$, it means that $f(x_*) = \rho$. $\Box$

Proposition 3.7. *Let regularity condition hold and $Y \in \mathcal{Y}$. Let also h be a function such that $h(u,v) < 0$ for all $(u,v) \ll 0$ and $h(u',v') > 0$ for all $(u',v') \notin \mathbb{R}^{m+1}_-$. Assume that there exist a number η_* and $x_* \in Y$ such that (15) holds. Then x_* is a solution of $P(f,g)$.*

Proof. The same argument as in the proof of Proposition 3.6 demonstrates that $x_* \in X_0$ and $f(x_*) \leq \eta_*$. If x_* is not a solution of $P(f,g)$, then there exists a point $\bar{x} \in \tilde{X}_0$ such that $f(\bar{x}) < f(x_*) \leq \eta_*$. Hence, $h(f(\bar{x}) - \eta_*, g(\bar{x})) < 0$, which contradicts (15). $\Box$

We now give necessary and sufficient condition for optimality.

Proposition 3.8. *Let $h : \mathbb{R}^{1+m} \to \bar{\mathbb{R}}$ be a function such that*

$$h(u,v) < 0 \Longleftrightarrow (u,v) \in \mathcal{H}^-; \tag{16}$$

and

$$h(0,v) = 0 \Longleftrightarrow v \leq 0. \tag{17}$$

Let $x_* \in Y$, where $Y \in \mathcal{Y}$ and $f(x_*) = \eta_*$. Then x_* is a solution of $P(f,g)$ if and only if

$$\min_{x \in Y} h(f(x) - \eta_*, g(x)) = h(0, g(x_*)) = 0. \tag{18}$$

Proof.

1) Assume that (18) holds. Since $h(0, g(x_*)) = 0$, we conclude, by applying (17), that $g(x_*) \leq 0$. Thus, $x_* \in X_0$. Proposition 3.3 demonstrates that x_* is a solution of $P(f,g)$.

2) Let x_* be a solution of $P(f,g)$. Since $g(x_*) \leq 0$, it follows from (17) that $h(0, g(x_*)) = 0$, so we only need to check that $h(f(x) - \eta_*, g(x)) \geq 0$ for all $x \in Y$. First, consider $x \in X_0$. Then $f(x) - f(x_*) = f(x) - \eta_* \geq 0$, hence $(f(x) - \eta_*, g(x)) \notin \mathcal{H}^-$. It follows from (16) that $h(f(x) - \eta_*, g(x)) \geq 0$.

Consider now a point $x \in Y$ such that $g(x) \not\leq 0$. Then again $(f(x) - \eta_*, g(x)) \notin \mathcal{H}^-$, so $h(f(x) - \eta_*, g(x)) \geq 0$. $\Box$

Remark 3.2. *We can assume, in particular, that $Y = X$ in Proposition 3.8. Then a separation function with properties indicated in this proposition allows us to establish that an arbitrary point $x_* \in X$ such that (18) holds is a solution of $P(f,g)$ (in particular, x_* is a feasible element of $P(f,g)$).*

Remark 3.3. *Assume that h is lower semicontinuous and* (16) *holds. Then* (17) *follows from the condition similar to* (9)*:*

$$u < 0 \Rightarrow h(u,v) < h(0,v) \quad \text{for all} \quad v \in \mathbb{R}^m.$$

Indeed, let $v \le 0$ and let $u < 0$. Then $(u,v) \in \mathcal{H}^-$, so $h(u,v) < 0$. Since h is lower semicontinuous, it follows that $h(0,v) \le 0$. On the other hand, since $(0,v) \notin \mathcal{H}^-$, it follows that $h(0,v) \ge 0$. Thus, $h(0,v) = 0$.

Assume now that $h(0,v) = 0$. Then $h(u,v) < h(0,v) = 0$ for $u < 0$. Hence $(u,v) \in \mathcal{H}^-$, which implies $v \le 0$.

Remark 3.4. *It follows from* (16) *that the level set $\{(u,v) : h(u,v) < 0\}$ of the function h from Proposition* 3.8 *is not open, so this function is not upper semicontinuous, hence discontinuous.*

The following function h satisfies conditions of Proposition 3.8: $h(u,v) = u + \delta(v)$, where δ is the indicator function of the cone $\mathbb{R}^m_-$:

$$\delta(v) = \begin{cases} 0, & \text{if } \ v \in \mathbb{R}^m_-, \\ +\infty, & \text{if } \ v \notin \mathbb{R}^m_-. \end{cases}$$

We can get more interesting results if we consider a separation functions depending on a parameter (see Section 4). Separation functions in a parametric form first were considered by Giannessi (see [11, 12, 21] and references therein). In the next section we shall present main ideas of Giannessi scheme.

4 GIANNESSI SCHEME AND CONDITIONS FOR OPTIMALITY IN A PARAMETRIC FORM

The approach, which was presented in previous sections, can be considered as a modification of the approach, which was proposed by Giannessi (see [11] and also [21] and references therein). We indicate some ideas behind the original Giannessi scheme conformably to the problem $P(f,g)$. As it was mentioned in [11], the separation of two sets by a not necessary linear function plays a key role in the study of optimality conditions.

Giannessi proposed to study the separation of the set $\mathcal{H}^-$, which we exploit here, and the following set:

$$\mathcal{K}_{\bar{x}} = \{(u,v) \in \mathbb{R}^{1+m} : u = f(x) - f(\bar{x}), \;\; v = g(x), \;\; x \in X\},$$

where $\bar{x} \in X$. He also considered a separation by a family of separation functions $h(u,v;\omega)$ depending on a parameter $\omega \in \Omega$. In particular, the following definition from [21] is very useful.

Definition 4.1. *A function $h : \mathbb{R}^{1+m} \times \Omega \to \bar{\mathbb{R}}$ is called a regular weak separation (RWS) function if $h(u,v;\omega) < 0$ for all $(u,v) \in \mathcal{H}^-$ and all $\omega \in \Omega$; for each $(u,v) \notin \mathcal{H}^-$ there exists $\omega \in \Omega$ such that $h(u,v;\omega) \ge 0$.*

The main tool in the study of optimality conditions by means of RWS functions is the so-called theorem of the alternative [11].

Proposition 4.1 (weak alternative). *Let $h : \mathbb{R}^{1+m} \times \Omega \to \bar{\mathbb{R}}$ be a RWS function and let $\bar{x} \in X$. Then the following assertions are not simultaneously true:*

1) for each $x \in X$ there exists $\omega \in \Omega$ such that: $h(f(x) - f(\bar{x}), g(x); \omega) \ge 0$;

2) $\mathcal{H}^- \cap \mathcal{K}_{\bar{x}} \neq \emptyset$.

Proof.
1) Assume that there exists $(u,v) \in \mathcal{H}^- \cap \mathcal{K}_{\bar{x}}$. Since $(u,v) \in \mathcal{K}_{\bar{x}}$, it follows that there exists $x \in X$ such that $u = f(x) - f(\bar{x})$, $v = g(x)$. Since $(u,v) \in \mathcal{H}^-$ and h is a RWS function, it follows that $h(f(x) - f(\bar{x}), g(x); \omega) < 0$ for all $\omega \in \Omega$, so Assertion 2) does not hold.
2) Assume now that $\mathcal{H}^- \cap \mathcal{K}_{\bar{x}} = \emptyset$. Then for each $x \in X$ there exists $\omega \in \Omega$ such that $h(f(x) - f(\bar{x}), g(x); \omega) \geq 0$, so Assertion 1) is valid. □

We now present a simple example of a RWS function. Let $\Omega = (0, +\infty)$ and

$$h(u, v; \omega) = \phi(u) + \omega \max(0, v^1, \dots, v^m),$$

where $\phi(u) < 0 = \phi(0)$ for $u < 0$ and $\phi(u) \geq 0$ for $u > 0$. We now check that h is a RWS function. Indeed, let $(u,v) \in \mathcal{H}^-$, that is, $u < 0$ and $v^i \leq 0$ for all i. Then $h(u,v;\omega) = \phi(u) < 0$ for all $\omega > 0$. Assume now that $(u,v) \notin \mathcal{H}^-$, then either $u \geq 0$, or $v \notin \mathbb{R}^m_-$. If $u \geq 0$ and $v \leq 0$, then $h(u,v;\omega) = \phi(u) \geq 0$ for all $\omega > 0$. If $v \notin \mathbb{R}^m_-$, then $\max(0, v^1, \dots, v^m) > 0$, so $h(u,v;\omega) > 0$ for sufficiently large ω.

Results of previous subsections can be easily reformulated in terms of RWS functions. We present only a reformulation of Proposition 3.1 and Proposition 3.3 (see Proposition 4.2 and Proposition 4.3, respectively).

Proposition 4.2. *Let* $h : \mathbb{R}^{1+m} \times \Omega \to \bar{\mathbb{R}}$ *be a RWS function such that the function* $(u,v) \mapsto h(u,v;\omega)$ *is lower semicontinuous for each* $\omega \in \Omega$. *Let* $Y \in \mathcal{Y}$ *and let* η_* *be a number such that*

1) for all $(u',v') \in \mathcal{T}_{\eta_*}$ *there exists* $\omega \in \Omega$ *such that* $h(u',v';\omega) \geq 0$;

2) $\mathbb{R}^{1+m}_- \cap \mathcal{T}_{\eta_*}(Y) \neq \emptyset$.

Let x_* *be a solution of the problem* $P(f,g)$. *Then* $f(x_*) = \eta_*$ *and there exists* $\omega' \in \Omega$ *such that* $h(0, g(x_*); \omega') = 0$.

Proposition 4.3. *Let* $h : \mathbb{R}^{1+m} \times \Omega \to \bar{\mathbb{R}}$ *be a RWS function. Let* $x_* \in X_0$, $\eta_* = f(x_*)$ *and there exists* $\omega' \in \Omega$ *such that*

$$\inf_{x \in Y} h(f(x) - \eta_*, g(x); \omega') \geq 0, \tag{19}$$

where $Y \in \mathcal{Y}$. *Then* x_* *is a solution of* $P(f,g)$. *If the function* $(u,v) \mapsto h(u,v;\omega')$ *is lower semicontinuous, then*

$$h(0, g(x_*); \omega') = \inf_{x \in Y} h(f(x) - \eta_*, g(x); \omega') = 0.$$

5 CONVOLUTION FUNCTIONS

The results obtained in Section 3 demonstrate that the problem $P(f,g)$ may be reduced to minimization of the function $x \mapsto h(f(x) - \eta, g(x))$ over a set $Y \in \mathcal{Y}$ (in particular, over the entire space X). Here h is a separation function (or similar to a separation function as in Proposition 3.6). The class of separation functions $h(u,v)$ is rather broad. Many separation functions may be constructed by the composition of some special functions, which are called convolution functions, since they are used for the convolution of the objective function f and coordinate functions $g^1, \dots, g^m$ of the constraint mapping g. As a rule, we shall consider a two-step convolution. First, we shall convolute all constraints to a single constraint. Second, we shall convolute the objective function and the obtained single constraint. In terms of separation it means that a separation function $h(u, v)$ has the form

$$h(u, v) = q_1(u, q_2(v)), \tag{20}$$

where $q_1 : \mathbb{R}^2 \to \mathbb{R}$ and $q_2 : \mathbb{R}^m \to \bar{\mathbb{R}}$ are convolution functions. A function h defined by (20) is a separation function only if convolution functions q_1 and q_2 possess some properties. In this section we shall formulate some of these properties for a functions defined on the space $\mathbb{R}^n$, provide a classification of convolution functions and give some examples of such functions.

We define continuity for functions mapping into the extended real line $\bar{\mathbb{R}}$ by the usual way: a function $q : \mathbb{R}^n \to \bar{\mathbb{R}}$ is called continuous on $\mathbb{R}^n$ if $\lim\limits_{k\to\infty} q(z_k) = q(z)$ for any $z \in \mathbb{R}^n$ and any sequence $z_k \to z$.

Definition 5.1. *A continuous function $q : \mathbb{R}^n \to \bar{\mathbb{R}}$ is called a* **convolution function** *if q strictly separates $\mathbb{R}^n_{--}$ and $\mathbb{R}^n_{++}$, that is, $q(z_1) < q(z_2)$ for any $z_1 \in \mathbb{R}^n_{--}$ and $z_2 \in \mathbb{R}^n_{++}$.*

Note that, due to continuity, any convolution function strictly separates the origin from one of the orthants $\mathbb{R}^n_{--}$ or $\mathbb{R}^n_{++}$, that is, either $q(0) > q(z)$ for any $z \in \mathbb{R}^n_{--}$ or $q(0) < q(z)$ for any $z \in \mathbb{R}^n_{++}$. In the case of *increasing convolution* function we shall use abbreviation IC function. (A function q defined on $\mathbb{R}^n$ is called increasing if $z_1 \leq z_2$ implies $q(z_1) \leq q(z_2)$. Note also that each continuous strictly increasing function $q : \mathbb{R}^n \to \mathbb{R}$ is an IC function (q is strictly increasing if $z_1 \ll z_2$ implies $q(z_1) < q(z_2)$).

Definition 5.2. *A convolution function $q(z)$ is called interior if there exists a constant $\gamma \in \mathbb{R}_{+\infty}$ such that*

$$q(z_1) < \gamma \ \ \textit{for all} \ \ z_1 \in \mathbb{R}^n_{--} \ \ \textit{and} \ \ q(z_2) = \gamma \ \ \textit{for all} \ \ z_2 \in \mathbb{R}^n_{+}.$$

If

$$q(z_1) < \gamma \ \ \textit{for all} \ \ z_1 \in \mathbb{R}^n_{--} \ \ \textit{and} \ \ q(z_2) = \gamma \ \ \textit{for all} \ \ z_2 \in \mathcal{E}^n_{+},$$

then the convolution function $q(z)$ is called **strictly interior**.

Definition 5.3. *A convolution function $q(z)$ is called exterior if there exists a constant $\gamma \in \mathbb{R}_{-\infty}$ such that*

$$q(z_1) = \gamma \ \ \textit{for all} \ \ z_1 \in \mathbb{R}^n_{-} \ \ \textit{and} \ \ q(z_2) > \gamma \ \ \textit{for all} \ \ z_2 \in \mathbb{R}^n_{++}.$$

If

$$q(z_1) = \gamma \ \ \textit{for all} \ \ z_1 \in \mathbb{R}^n_{-} \ \ \textit{and} \ \ q(z_2) > \gamma \ \ \textit{for all} \ \ z_2 \in \mathcal{E}^n_{++},$$

then the convolution function $q(z)$ is called **strictly exterior**.

Definition 5.4. *A convolution function $q(z)$ is called common if there exist a constant $\gamma \in \mathbb{R}$ such that*

$$q(z_1) < \gamma < q(z_2) \ \ \textit{for all} \ \ z_1 \in \mathbb{R}^n_{--} \ \ \textit{and} \ \ z_2 \in \mathbb{R}^n_{++}.$$

If

$$q(z_1) < \gamma < q(z_2) \ \ \textit{for all} \ \ z_1 \in \mathbb{R}^n_{--} \ \ \textit{and} \ \ z_2 \in \mathcal{E}^n_{++},$$

then the convolution function is called **strictly common**.

If $q(z)$ is strictly interior, strictly exterior or strictly common convolution function, then, due to continuity, it has the same value γ on the boundary of the orthant $\mathbb{R}^n_{-}$, where γ is the constant from the definitions of these functions. Note that unlike to an ordinary convolution function the common convolution function necessarily strictly separates the origin from the orthants $\mathbb{R}^n_{--}$ and $\mathbb{R}^n_{++}$ simultaneously. For strictly common convolution function we have

$$\mathbb{R}^n_{--} = \{z : q(z) < \gamma\}, \qquad \mathbb{R}^n_{-} = \{z : q(z) \leq \gamma\}.$$

Definition 5.5. *The convolution function $q(z)$ is called proper if $-\infty < \gamma < +\infty$, otherwise it is called* **improper**.

Below we restricted ourselves mainly by application of proper convolution functions. We assume for simplicity that the constant γ for any proper convolution function is equal to zero. The following implications follow directly from the definitions:
1) if q is a proper strictly interior convolution function, then

$$q(z) < 0 \Longleftrightarrow z \in \mathbb{R}^n_{--};$$

2) if q is a proper strictly exterior convolution function, then

$$q(z) > 0 \Longleftrightarrow z \notin \mathbb{R}^n_{-}; \tag{21}$$

3) if q is a proper strictly common convolution function, then (see (22)):

$$q(z) < 0 \Longleftrightarrow z \in \mathbb{R}^n_{--}, \qquad q(z) > 0 \Longleftrightarrow z \notin \mathbb{R}^n_{-}. \tag{22}$$

Thus, a strictly interior convolution function strictly separates the open set $\mathbb{R}^n_{--}$ and the closed set $\mathcal{E}^n_{+}$. A strictly exterior convolution function strictly separates the closed set $\mathbb{R}^n_{-}$ and the open set $\mathcal{E}^n_{++}$. A strictly common convolution function strictly separates $\mathbb{R}^n_{--}$ from $\mathcal{E}^n_{+}$ and $\mathbb{R}^n_{-}$ from $\mathcal{E}^n_{++}$ simultaneously.

Note that any finite convolution function is a proper convolution function. We shall describe some other properties of finite convolution functions.

Proposition 5.1. *Let $q_1(z)$ and $q_2(z)$ be finite convolution functions, and let $c > 0$, $a \in \mathbb{R}^n_{++}$. Then the following functions are also finite convolution functions:*

$$\begin{aligned} q(z) &= cq_1(z), \\ q(z) &= q_1(z) + q_2(z), \\ q(z) &= \min\{q_1(z), q_2(z)\}, \\ q(z) &= \max\{q_1(z), q_2(z)\}, \\ q(z) &= q_1(D(\mathbf{a})z), \end{aligned} \tag{23}$$

where $D(\mathbf{a}) = \operatorname{diag}(\mathbf{a})$ is a diagonal matrix with a vector $\mathbf{a}$ on the diagonal. Moreover, if both functions $q_1(z)$ and $q_2(z)$ are simultaneously interior, exterior or common convolution functions, then the resulting function $q(z)$ is also a convolution function of the same type.

Proof. It is clear that any function $q(z)$ from (23) is continuous. The inequality $q(z_1) < q(z_2)$ for $z_1 \in \mathbb{R}^n_{--}$ and $z_2 \in \mathbb{R}^n_{++}$ follows directly from corresponding inequalities for functions $q_1(z)$ and $q_2(z)$ and from the fact that $D(\mathbf{a})z_1 \in \mathbb{R}^n_{--}$, $D(\mathbf{a})z_2 \in \mathbb{R}^n_{++}$ for any $z_1 \in \mathbb{R}^n_{--}$, $z_2 \in \mathbb{R}^n_{++}$. The proof of the second conclusion is similar. $\square$

Proposition 3.3 demonstrates that sufficient optimality conditions can be obtained by exploiting of a separation function h defined on $\mathbb{R}^{1+m} = \mathbb{R}^1 \times \mathbb{R}^m$ with the property:

$$h(u, v) < 0 \;\text{ for all }\; (u, v) \in \mathcal{H}^- = \mathbb{R}^1_{--} \times \mathbb{R}^m_{-}.$$

Proper convolution functions of various kinds will allow us to construct a function with such a property.

A lot of various convolution functions can be proposed. For some applications it is important to have functions, which do not destroy the convexity. Otherwise, in the case of convex programming we lose opportunity to use the local minimization techniques. Note that IC functions are more preferable from this point of view. In particular, if $q(z)$ is a convex IC function, then the convolution of convex functions by means of $q(z)$ is a convex function as well.

Let $\mathcal{S} := \{\alpha \in \mathbb{R}^n_+ : \sum_{i=1}^{n} \alpha^i = 1\}$ be the unit simplex. For each $\alpha \in \mathcal{S}$ set $I(\alpha) = \{i : \alpha^i \neq 0\}$. For $p \neq 0$ and $\alpha \in \mathcal{S}$ we define a function $\|z\|_{p,\alpha} : \mathbb{R}^n \to \mathbb{R}$. If $p > 0$, then

$$\|z\|_{p,\alpha} = \left(n \sum_{i \in I(\alpha)} \alpha^i |z^i|^p \right)^{1/p}. \tag{24}$$

If $p < 0$, then $\|z\|_{p,\alpha}(z) = 0$ if $z^i = 0$ for at least one index $i \in I(\alpha)$, otherwise, $\|z\|_{p,\alpha}(z)$ is defined by (24). It follows from the definition that $\|\cdot\|_{p,\alpha}$ is positively homogeneous of degree one function.

We define also the function $\|\cdot\|_{p,\alpha}$ for $p = 0$ by the following way

$$\|z\|_{0,\alpha} = \sqrt{n} \prod_{i \in I(\alpha)} |z^i|^{\alpha^i}. \tag{25}$$

Limit cases with $p = +\infty$ or $p = -\infty$ lead to the functions:

$$\|z\|_{+\infty,\alpha} = \max_{i \in I(\alpha)} |z^i|, \qquad \|z\|_{-\infty,\alpha} = \min_{i \in I(\alpha)} |z^i|.$$

The function $\|z\|_{p,\alpha}$ is convex for $p \geq 1$. Clearly, this function is a seminorm and it is a norm if $\alpha \gg 0$. In particular, if $p \neq +\infty$, then $\|z\|_{p,\alpha}$ is a scaled Hölder norm. If $p < 1$, then $\|z\|_{p,\alpha}$ is concave on any orthant of $\mathbb{R}^n$ and, strictly speaking, the notation $\|\cdot\|$, which means a norm, cannot be used (norm is obligatory a convex function). Nevertheless we use this notation understanding this violation. In what follows we shall omit the parameter a in notation of the function $\|z\|_{p,\alpha}$, if $\alpha = n^{-1}e$, where e is a vector of ones.

Let $z = (z^1, \ldots, z^n) \in \mathbb{R}^n$. Denote the positive part and the negative part of z by z_+ and z_-, respectively. By definition,

$$z_+ = (z^1_+, \ldots, z^n_+), \qquad z_- = (z^1_-, \ldots, z^n_-),$$

where $a_+ = \max(a, 0)$, $a_- = \min(a, 0)$ for $a \in \mathbb{R}$.

The following classes of IC functions are widely used in constrained optimization.

1) The class of exterior IC functions

$$q(z) = \|z_+\|_{p,\alpha}, \qquad -\infty \leq p \leq +\infty. \tag{26}$$

If $p > 0$ and $\alpha \gg 0$, then the function (26) is a strictly exterior IC function.

2) The class of interior IC functions

$$q(z) = -\|z_-\|_{p,\alpha}, \qquad -\infty \leq p \leq +\infty. \tag{27}$$

If $p \leq 0$ and $\alpha \gg 0$, then the function (27) is a strictly interior IC function.

3) The third class is a class of common IC function consisting of sums of the functions (26) and (27):

$$q(z) = \|z_+\|_{p_1,\alpha} - \|z_-\|_{p_2,\alpha}, \tag{28}$$

where $-\infty \le p_1 \le +\infty$ and $-\infty \le p_2 \le +\infty$.

Note that if $p_1 = p_2 = 1$, then (28) becomes a linear function

$$q(z) = \sum_{i=1}^{n} a^i z^i, \tag{29}$$

with a vector $a = (a^i) = n\alpha$. If $p = p_1 \ge 1$ and $p_2 = -p_1$, then (28) has a form

$$q(z) = \|z_+\|_{p,\alpha} - \|z_-\|_{-p,\alpha}. \tag{30}$$

This convex function is a strictly common IC function if $\alpha \gg 0$.

Let $I = \{1, \dots, n\}$. If $p = +\infty$ and $\alpha \gg 0$, then (30) becomes the strictly common IC function

$$q(z) = \max_{i \in I} z^i. \tag{31}$$

The common (but not strictly common) IC function

$$q(z) = \min_{i \in I} z^i \tag{32}$$

can be derived from (30) with $p = -\infty$ and $\alpha \gg 0$. Observe that all functions (26) – (32) are proper IC functions. The following two functions are the simplest examples of unproper strictly interior and strictly exterior IC functions, respectively:

$$q(z) = \begin{cases} \|z_-\|_{-p}^{-p}, & z \in \mathbb{R}^n_{--}, \\ +\infty, & z \in \mathcal{E}^n_+, \end{cases} \qquad q(z) = \begin{cases} -\|z_+\|_p^{-p}, & z \in \mathcal{E}^n_{++}, \\ -\infty, & z \in \mathbb{R}^n_-, \end{cases}$$

where $p > 0$.

We now describe another set of IC functions, which is based on the notion of IPH functions. Let T coincides with either the space $\mathbb{R}^n$ or the cone $\mathbb{R}^n_+$. A function $q : T \to \mathbb{R}$ is called IPH if q is *increasing* and *positively homogeneous* of degree one ($q(\lambda x) = \lambda q(x)$ for $\lambda > 0$). Examples of IPH functions are given by (26) – (32). We present some other examples of IPH functions. The following functions defined on the entire space $\mathbb{R}^n$ are IPH:

$$q(z) = \max_{i \in I} a^i z^i \quad \text{with} \quad a = (a^i) \in \mathbb{R}^n_+ \backslash \{0\}. \tag{33}$$

Note that $q(z) > 0$ for all $z \in \mathbb{R}^n_{++}$ and $q(z) < 0$ for all $z \in \mathbb{R}^n_{--}$. If $a^i > 0$ for all $i \in I$, then $q(z) > 0$ for all $z \in \mathcal{E}^n_{++}$.

$$q(z) = \min_{i \in I} a^i z^i \quad \text{with} \quad a = (a^i) \in \mathbb{R}^n_+ \backslash \{0\}. \tag{34}$$

We have $q(z) > 0$ for all $z \in \mathbb{R}^n_{++}$ and $q(z) < 0$ for all $z \in \mathbb{R}^n \backslash \mathbb{R}^n_+$.

$$q(z) = \left(\sum_{i \in I} a^i (z^i)^p \right)^{1/p} \quad \text{with} \quad a^i \ge 0 \quad \text{and} \quad p = \dots, \frac{1}{5}, \frac{1}{3}, 1, 3, 5, \dots. \tag{35}$$

If $a^i > 0$ for all $i \in I$, then $q(z) < 0$ for all $z \in \mathbb{R}^n_- \backslash \{0\}$ and $q(z) > 0$ for all $z \in \mathbb{R}^n_+ \backslash \{0\}$.

Note that (33) and (34) are generalizations of (31) and (32), respectively. The following functions are IPH as well:

$$q_1(z) = \max_{\ell \in V} \langle \ell, z \rangle, \qquad q_2(z) = \min_{\ell \in V} \langle \ell, z \rangle, \tag{36}$$

where $\langle \ell, z\rangle$ is the inner product of vectors ℓ and z. Here $V \subset \mathbb{R}^n_+$ is a compact set. Clearly, (33) and (34) are special cases of q_1 and q_2, respectively.

If q is an IPH function defined on $\mathbb{R}^n$, then its restriction to $\mathbb{R}^n_+$ is again IPH, so functions defined on $\mathbb{R}^n_+$ by (33) and (34) are IPH. The following functions defined on $\mathbb{R}^n_+$ are IPH as well:

$$q(z) = \left(\sum_{i\in I} a^i (z^i)^p\right)^{1/p}, \qquad (z \in \mathbb{R}^n_+), \tag{37}$$

where $a = (a_i) \in \mathbb{R}^n_+\setminus\{0\}$, $p > 0$. If $a = n\alpha$, where $\alpha = (\alpha^1, \dots, \alpha^n) \in \mathcal{S}$, then (37) coincides with the restriction to $\mathbb{R}^n_+$ of functions $\|\cdot\|_{p,\alpha}$ defined by (24). If $p = \dots, 1/3, 1, 3, \dots$, these functions coincide with the restriction to $\mathbb{R}^n_+$ of functions defined by (35).

Let us give one more example of an IPH function defined on $\mathbb{R}^n_+$. Let $\alpha \in \mathcal{S}$ and $\alpha^i > 0$ for all $i \in I$. Let $c > 0$. Then the function

$$q(z) = c z_1^{\alpha^1} \dots z_n^{\alpha^n}, \qquad (z \in \mathbb{R}^n_+), \tag{38}$$

is IPH. If $c = \sqrt{n}$, then (38) coincides with restriction of (25) on $\mathbb{R}^n_+$. (In economics the function (38) is called *Cobb–Douglas function.*)

Let q be an IPH function defined on either $\mathbb{R}^n$ or $\mathbb{R}^n_+$. Then $q(0) = q(2\cdot 0) = 2q(0)$, hence either $q(0) = 0$ or $q(0) = \pm\infty$, so, if q is a proper function, then $q(0) = 0$. Assume now that q is a proper function defined on $\mathbb{R}^n$. Since q is increasing, it follows that $q(z) \le 0$ for $z \in \mathbb{R}^n_-$ and $q(z) \ge 0$ for $z \in \mathbb{R}^n_+$.

Proposition 5.2. *Let $q : \mathbb{R}^n \to \bar{\mathbb{R}}$ be a continuous IPH function such that $q(0) = 0$. Assume that there exist $z_1 \in \mathbb{R}^n_{--}$ and $z_2 \in \mathbb{R}^n_{++}$ such that $-\infty < q(z_1) < 0$ and $0 < q(z_2) < +\infty$. Then q is a finite common IC function.*

Proof. Since $z_2 \gg 0$, it follows that for each $z \in \mathbb{R}^n$ there exists $\lambda > 0$ such that $z \le \lambda z_2$. We have $q(z) \le q(\lambda z_2) = \lambda q(z_2) < +\infty$. If $z \gg 0$, then there exists $\mu > 0$ such that $z \gg \mu z_2$, hence $q(z) > \mu q(z_2) > 0$. The similar argument shows that $q(z) > -\infty$ for all $z \in \mathbb{R}^n$ and $q(z) < 0$ for $z \in \mathbb{R}^n_{--}$. $\square$

Proposition 5.3. *Let $q_1 : \mathbb{R}^n_+ \to \bar{\mathbb{R}}$ be a continuous IPH function such that $q_1(0) = 0$ and $q_1(z) > 0$ for all $z \in \mathbb{R}^n_+\setminus\{0\}$. Assume that there exists $z_1 \in \mathbb{R}^n_{++}$ such that $q_1(z_1) < < +\infty$. Then the function*

$$q(z) = q_1(z_+) \qquad (z \in \mathbb{R}^n) \tag{39}$$

is a finite strictly exterior IC function.

Proof. For each $z \in \mathbb{R}^n$ there exists $\lambda > 0$ such that $z_+ \le \lambda z_1$, therefore, $q_1(z) \le \lambda q_1(z_1) < +\infty$. If $z \in \mathbb{R}^n_-$, then $z_+ = 0$, so $q(z) = 0$. If $z \notin \mathbb{R}^n_-$, then $z_+ \in \mathbb{R}^n_+$ and $z_+ \ne 0$, so $q(z) > 0$. Since $z^1_+ \le z^2_+$ for all $z^1 \le z^2$, the function (39) is increasing. $\square$

Proposition 5.4. *Let $q_1 : \mathbb{R}^n_+ \to \bar{\mathbb{R}}$ be a continuous IPH function such that $q_1(z) = 0$ for all boundary points z of $\mathbb{R}^n_+$. Assume that there exists $z_1 \gg 0$ such that $q_1(z_1) > 0$. Then the function*

$$q(z) = -q_1(-z_-) \qquad (z \in \mathbb{R}^n) \tag{40}$$

is a finite strictly interior IC function.

Proof. The same argument as under proof of Proposition 5.2 shows that q_1 is finite and $q_1(z) > 0$ for all $z \in \mathbb{R}^n_{++}$. If $z \in \mathbb{R}^n_{--}$, then $-z_- \in \mathbb{R}^n_{++}$, so $q(z) = -q_1(-z_-) < 0$. If $z \notin \mathbb{R}^n_{--}$, then $-z_- \notin \mathbb{R}^n_{++}$, so $-z_-$ is a boundary point of $\mathbb{R}^n_+$. Hence $q(z) = 0$. The function (40) is increasing because of the function $-q_1(z)$ is decreasing and $z^1_- \le z^2_-$ for all $z^1 \le z^2$. $\square$

Remark 5.1. *Conditions of Proposition* 5.4 *are valid for the Cobb–Douglas function defined by* (38).

6 AUXILIARY FUNCTIONS

Consider the problem $P(f, g)$ defined by (4). We shall use the optimality conditions obtained in Section 3 for examination of this problem. For this purpose we shall construct separation functions of the form (20) by a two step convolution procedure. First we shall convolute constraints $g = (g^1, \dots, g^m)$ by means of a convolution function $\psi : \mathbb{R}^m \to \mathbb{R}$. Second, we shall convolute the objective function f and the new constraint $\psi(g(\cdot))$ by means of a convolution function $\varphi : \mathbb{R}^2 \to \bar{\mathbb{R}}$. Thus, we shall consider functions $h :: \mathbb{R}^{1+m} \to \bar{\mathbb{R}}$ of the form

$$h(u, v) = \varphi(u, \psi(v)). \tag{41}$$

As a rule, we assume that φ and ψ are proper IC functions and, moreover, the function $\varphi(\cdot, 0)$ possesses the following property:

$$u < 0 \Longrightarrow \varphi(u, 0) < 0 = \varphi(0, 0). \tag{42}$$

Note that (42) can be expressed in the following form: the function $\varphi(\cdot, 0)$ is either strictly interior or strictly common convolution function defined on $\mathbb{R}$.

We shall often assume that ψ is either proper strictly exterior or proper strictly common convolution function. It easy to check that in both cases

$$g(x) \le 0 \Longleftrightarrow \psi(g(x)) \le 0. \tag{43}$$

Indeed, if ψ is strictly exterior, then (43) follows directly from (21), if ψ is strictly common, then (43) directly follows from (22). Hence, the feasible set X_0 can be described by a single constraint $\psi(g)$, that is,

$$X_0 = \{x \in X : \psi(g(x)) \le 0\}. \tag{44}$$

The function h, defined on the image space $\mathbb{R}^{1+m}$ by (41), allows us to define the following function $M : X \times \mathbb{R} \to \bar{\mathbb{R}}$:

$$M(x, \eta) = \varphi(f(x) - \eta, \psi(g(x))). \tag{45}$$

A variable η can be considered as an estimation (lower or upper) of the optimal value ρ of the problem $P(f, g)$.

We are interested in *auxiliary functions* for problem $P(f, g)$, that is, functions M such that their unconstrained minimization on the set $Y \in \mathcal{Y}$ is equivalent (in a certain sense) to solving problem $P(f, g)$. Here we shall study only auxiliary functions of the form (45).

We shall show in next Proposition that, if ψ is a proper convolution function and φ is a proper IC function, then (41) is a separation function, so it is possible to reformulate the sufficient conditions obtained in Section 3 in terms of convolution functions φ and ψ and corresponding auxiliary function M. Recall that $\mathcal{Y} = \{Y \subseteq X : X_0 \subset Y\}$.

Proposition 6.1. *Let $Y \in \mathcal{Y}$. Let ψ be a proper convolution function and φ be a proper IC function such that* (42) *holds. Assume that $x_* \in X_0$ and $\eta_* \in \mathbb{R}$ possess the following properties:*

$$\begin{aligned} f(x_*) &= \eta_*, & (46)\\ \psi(g(x_*)) &= 0, & (47)\\ M(x_*, \eta_*) &= \min_{x \in Y} M(x, \eta_*). & (48) \end{aligned}$$

Then x_* is a solution of the problem $P(f,g)$.

Proof. Consider the function $h(u,v) = \varphi(u, \psi(v))$. We shall check that all conditions of Proposition 3.3 hold for this function. It follows from (46) – (48) that

$$h(f(x_*) - \eta_*, g(x_*)) = \min_{x \in Y} h(f(x) - \eta_*, g(x)) = 0. \tag{49}$$

Since $\psi(v)$ is a proper convolution function, it follows that $\psi(v) \leq 0$ for all $v \in \mathbb{R}^m_-$. Moreover, since φ is a proper IC function, we have $\varphi(u, \psi(v)) \leq \varphi(u, 0)$ for $v \in \mathbb{R}^m_-$ and any $u \in \mathbb{R}$. Applying (42), we conclude that

$$h(u,v) = \varphi(u, \psi(v)) \leq \varphi(u, 0) < \varphi(0,0) = 0$$

for all $(u,v) \in \mathcal{H}^-$. $\square$

Remark 6.1. *Let ψ be a strictly exterior or strictly common convolution function. Then (see (44)) $\psi(g(x)) = 0$ implies $x \in X_0$, so, due to (47), we can omit the assumption $x_* \in X_0$.*

Proposition 6.2. *Let regularity condition hold. Let also φ be a proper strictly interior convolution function and ψ be a proper interior convolution function. If there exist $x_* \in X_0$ and $\eta_* \in \mathbb{R}$ such that $f(x_*) = \eta_*$ and*

$$M(x_*, \eta_*) = \min_{x \in X_0} M(x, \eta_*), \tag{50}$$

then $x_* \in X_*$.

Proof. Since $\psi(v)$ is a proper interior convolution function, we have $\psi(v) < 0$ for all $v \in \mathbb{R}^m_{--}$. Similarly, since $\varphi(u, \psi(v))$ is a proper strictly interior convolution function, it follows that $\varphi(u, \psi(v)) < 0$ for any $(u,v) \in \mathbb{R}^2_{--}$. These relations imply the inequality $h(u,v) < 0$ for all $(u,v) \in \mathbb{R}^{m+1}_{--}$, where h is the function defined by (41). Moreover, it follows from (50) that

$$h(f(x_*) - \eta_*, g(x_*)) = \min_{x \in X_0} h(f(x) - \eta_*, g(x)).$$

Since $f(x_*) - \eta_* = 0$ and $\psi(g(x_*)) \leq 0$, we have for the proper strictly interior convolution function φ that

$$h(f(x_*) - \eta_*, g(x_*)) = \varphi(f(x_*) - \eta_*, \psi(g(x_*))) = 0.$$

Therefore, $h(u,v) \geq 0$ for all $(u,v) \in \mathcal{T}_{\eta_*}(X_0)$. Applying Proposition 3.5, we conclude that $x_* \in X_*$. $\square$

Note that (46) and (47) imply that

$$M(x_*, \eta_*) = 0. \tag{51}$$

However, in general two equalities (46) and (47) cannot be replaced by one equality (51). Indeed, if ψ is a strictly exterior or strictly common convolution function, then $\psi(g(x_*)) = 0 \Longrightarrow g(x_*) \leq 0$, hence x_* is a feasible point. We cannot provide the feasibility of x_* using only (51). Nevertheless, we can replace (46) and (47) by (51) under some additional assumptions.

Proposition 6.3. *Let $Y \in \mathcal{Y}$ and let φ and ψ be proper strictly exterior convolution functions. If there exist $x_* \in Y$ and $\eta_* \leq \rho$ such that (48) and (51) hold, then x_* is a solution of $P(f,g)$.*

Proof. Since φ and ψ are proper strictly exterior convolution functions, it follows that $h(u,v) > 0$ for all $(u,v) \in \mathcal{E}^{m+1}_{++}$. (Here h is defined by (41)). We also have from (48) and (51) that (15) holds. Thus, by Proposition 3.6, $x_* \in X_*$. $\square$

Proposition 6.4. *Let regularity condition hold and $Y \in \mathcal{Y}$. Let also φ and ψ be proper strictly common convolution functions. If there exist $x_* \in Y$ and $\eta_* \in \mathbb{R}$ such that* (48) *and* (51) *hold, then x_* is a solution of $P(f,g)$.*

Proof. Let conditions in the proposition hold. Then the function (41) possesses the following properties:

$$\begin{array}{rcl} h(u,v) & < & 0 \text{ for all } (u,v) \in \mathbb{R}^{m+1}_{--} \text{ and} \\ h(u',v') & > & 0 \text{ for all } (u',v') \notin \mathbb{R}^{m+1}_{-}. \end{array}$$

Thus, the result follows from Proposition 3.7. □

Sufficient conditions given by Propositions 6.1 – 6.4 are based on the auxiliary unconstrained optimization problem

$$M(x,\eta) \longrightarrow \min \quad \text{subject to} \quad x \in Y, \tag{52}$$

which we denote by $P_Y(f,g;\eta)$. We have to select functions φ and ψ such that a solution of $P_Y(f,g;\eta)$ exists and, moreover, this solution is feasible for $P(f,g)$.

The simplest example of a convolution function φ is the linear function

$$\varphi(u,w) = u + w. \tag{53}$$

In such a case $M(x,\eta) = f(x) - \eta + \psi(g(x))$. Clearly, the variable η does not affect on the solution set of the minimization problem $P_Y(f,g;\eta)$, so we can exclude it from $M(x,\eta)$. Thus, having the convolution function (53), we can consider the following auxiliary function:

$$M(x) = f(x) + \psi(g(x)), \qquad (x \in X). \tag{54}$$

Remark 6.2. *A well-known example of an auxiliary function* (54) *is Augmented Lagrangian (see [17] and references therein). The following assertion holds:*

Proposition 6.5. *Let $Y \in \mathcal{Y}$ and let M be an auxiliary function defined by* (54)*. Let ψ be a proper convolution function. If there exists $x_* \in X_0 \cap Y$, such that*

1) x_* is a minimizer of the function (54) *on the set Y;*

2) $\psi(g(x_*)) = 0$.

Then $x_* \in X_*$.

Proof. The proof follows directly from Proposition 6.1. □

The solution set of (52) with the objective function (54) coincides with the solution set of $P(f,g)$ only under some rather restrictive assumptions. Thus, we need to consider more general functions $h(u,v)$, which are determined by nonlinear IC function φ.

7 AUXILIARY FUNCTIONS DEPENDING ON A PARAMETER

Consider the separation function h depending on a parameter $\omega \in \Omega$, where Ω is an arbitrary set. Let $Y \in \mathcal{Y}$ be a set such that the objective function f of the problem $P(f,g)$ is bounded from below on Y and let $\eta < \gamma := \inf_{x \in Y} f(x)$. Then $f(x) - \eta > 0$ for all $x \in Y$. Without loss of generality we assume in the current section that f is positive on Y and $\inf_{x \in Y} f(x) > 0$, otherwise, we can consider the function $f-\eta$. Introduce an auxiliary function depending on the parameter ω:

$$M(x,\omega) = h(f(x), g(x); \omega),$$

and consider the following problem $P_Y(f,g;\omega)$ of unconstrained minimization

$$M(x,\omega) \longrightarrow \min \quad \text{subject to} \quad x \in Y. \tag{55}$$

We assume in the sequel that this problem has a solution. It is interesting to find conditions such that

$$\sup_{\omega\in\Omega} \min_{x\in Y} M(x,\omega) = \min_{x\in X_0} f(x), \tag{56}$$

where $X_0 = \{x \in X : g(x) \leq 0\}$ is the set of feasible elements of the problem $P(f,g)$. The equality (56) is usually called the zero duality gap property. If the zero duality gap property holds, then the problem $P(f,g)$ can be reduced to a sequence of unconstrained optimization problems.

In this section we consider only separation functions $h(u,v;\omega)$ with the following property:

$$(u > 0,\ v \leq 0) \Longrightarrow h(u,v;\omega) = u \quad \text{for all} \quad \omega \in \Omega. \tag{57}$$

Lemma 7.1. *$M(x,\omega) = h(f(x), g(x); \omega) = f(x)$ for all $x \in X_0$ and $\omega \in \Omega$.*

Proof. The result follows directly from (57), since $f(x) > 0$ and $g(x) \leq 0$ for all $x \in X_0$. $\square$

Let (x_*, ω_*) be a saddle point of the function $M(x,\omega)$. Assume that (57) holds and

$$(u > 0,\ h(u,v;\omega) \leq h(u,v;\omega_*) \Longrightarrow v \leq 0.$$

Then x_* is a solution of $P(f,g)$. Indeed, we have

$$h(f(x_*), g(x_*); \omega) \leq h(f(x_*), g(x_*); \omega_*) \qquad (\omega \in \Omega),$$

hence $g(x_*) \leq 0$, that is, $x_* \in X_0$. On the other hand, keeping in mind Lemma 7.1, we have for $x \in X_0$

$$f(x) = h(f(x), g(x); \omega_*) \geq h(f(x_*), g(x_*); \omega_*) = f(x_*).$$

Let $\Omega = \mathbb{R}^m_+$ and $h : \mathbb{R}^{1+m} \times \mathbb{R}^m_+ \to \mathbb{R}$ be a linear function of the form

$$h(u,v;\omega) = u + \sum_{i=1}^{m} \omega^i v^i.$$

Then the auxiliary function $M(x,\omega)$ coincides with the Lagrange function of the problem $P(f,g)$. It is well-known that the zero duality gap property holds in such a case only under some additional assumptions. Thus, if we want to provide the existence of the zero duality gap property for a broad class of problems, we need to suggest some conditions, which exclude linearity. We now present one of a possible set of such conditions. We shall use the approach developed by Andramonov [1].

Consider the problem $P(f,g)$. Let $\alpha > 0$ be a sufficiently small number. Define the following set-valued mapping $G : [0,\alpha] \to X$:

$$G(\varepsilon) = \{x \in Y : g^j(x) \leq \varepsilon,\ j = 1,\dots,m\}.$$

Clearly, $G(0) = X_0$. First we need the following assumption, related to the problem $P(f,g)$.

Assumption 7.1.

1) *The mapping G is upper semicontinuous at the origin, that is, for each $\mu \in [0, \alpha]$ there exists $\delta > 0$ such that $G(\lambda) \subset G(0) + \mu B$ for all $\lambda \in (0, \delta)$. Here $B = \{x : \|x\| \le 1\}$.*

2) *The function f is uniformly continuous on a set $Y \in \mathcal{Y}$ and*
$$\gamma := \inf_{x \in Y} f(x) > 0. \tag{58}$$

It is well-known that a set-valued mapping F is upper semicontinuous if its graph $\mathrm{gr}\, F$ is a compact set (see, for example, [2]). Note that

$$\begin{aligned} \mathrm{gr}\, G \;&:=\; \{(x, \varepsilon) \in Y \times [0, \alpha] : g^j(x) \le \varepsilon,\ j = 1, \dots, m,\ 0 \le \varepsilon \le \alpha\} \subset \\ &\subset\; \{x \in Y : g^j(x) \le \alpha,\ j = 1, \dots, m\} \times [0, \alpha]. \end{aligned}$$

Thus, the mapping G is upper semicontinuous if the set $\{x \in Y : g^j(x) \le \alpha,\ j = 1, \dots, m\}$ is compact.

Now we present a set of assumptions related to a set of parameters Ω and to a separation function $h(u, v; \omega)$.

Assumption 7.2. *A set of parameters Ω is a subset of the space $\mathbb{R}^k$ and $\Omega \supset R_e$, where $R_e = \{(\lambda, \dots, \lambda) : \lambda > 0\}$ is the open ray starting from the origin and passing through $e = (1, \dots, 1)$.*

Assumption 7.3. *The separation function $h : \mathbb{R}^{1+m} \times \Omega$ enjoys the following properties:*

1) *$h(u, v; \omega) = h(u, v_+; \omega)$ for all $u > 0$, $v \in \mathbb{R}^m$ and $\omega \in \Omega$;*

2) *$h(u, 0; \omega) = u$ for all $u > 0$ and $\omega \in \Omega$;*

3) *$h(u, v; \omega) \to +\infty$ as $\min_i \omega^i \to +\infty$ for all $u \ge 0$ and $v \in \mathbb{R}^m_+ \setminus \{0\}$;*

4) *the function $(u, v) \mapsto h(u, v; \omega)$ is increasing on the set $\{(u, v) : u > 0,\ v \ge 0\}$ for all $\omega \in \Omega$.*

Note that (57) follows from 1) and 2). We now present some examples of the function h with properties 1) – 4):

$$\begin{aligned} &1)\quad \Omega = \mathbb{R}^m_{++}, \quad h(u, v; \omega) = u + \sum_{i=1}^m \omega^i v^i_+; \\ &2)\quad \Omega = \mathbb{R}_+, \quad\;\;\; h(u, v; \omega) = u + \omega \max(0, v^1, \dots, v^m) \equiv u + \omega \max_j v^j_+; \\ &3)\quad \Omega = \mathbb{R}^m_+, \quad\;\, h(u, v; \omega) = \left(|u|^p + \sum_{j=1}^m (\omega^j |v^j_+|)^p, \right)^{1/p}, \quad p > 0. \end{aligned}$$

Remark 7.1. *Consider a stronger version of Assumption 7.3, where items 1)–4) hold for all $u \in \mathbb{R}$, not only for $u > 0$. Then h is a RWS function (see Section 4). Indeed, let $(u, v) \in \mathcal{H}^-$. Then $u < 0$, $v < 0$, hence $h(u, v; \omega) = h(u, 0; \omega) = u < 0$ for all $\omega \in \Omega$. Consider now a point $(u, v) \notin \mathcal{H}^-$. Then either $u > 0$, $v \le 0$ or $v \notin \mathbb{R}^m_-$. If the former holds, then $h(u, v; \omega) = u \ge 0$ for all $\omega \in \Omega$. If the latter is valid, then, due to 3), there exists $\omega' \in \Omega$ such that $h(u, v; \omega') > 0$. Note that the functions h from Examples 1) and 2) above possess properties, mentioned here.*

Let $\omega \in \Omega$. Consider the problem $P_Y(f, g; \omega)$, defined by (55). Let $x_*(\omega)$ be a solution of this problem (we assume that such a solution exists).

Lemma 7.2. *For any $\delta > 0$ there exists $\ell = \ell(\delta) > 0$ such that $x_*(\omega) \in G(\delta)$ if $\min_i \omega^i \geq \ell$.*

Proof. Assume, in the contrary, that there exists $\delta' > 0$ such that for any positive integer $\ell > 0$ it is possible to find $\omega_\ell \in \Omega$ with the following properties:

1) $\min_i \omega_\ell^i \geq \ell$;

2) $x_*(\omega_\ell) \notin G(\delta')$, that is, there exists j such that $g^j(x_*(\omega_\ell)) > \delta'$.

Assume, without loss of generality, that the index j does not depend on ℓ. Consider the vector $e_j = (e^i)$ where $e^i = 0$ for $i \neq j$ and $e^j = \delta'$. Let $u_\ell = f(x_*(\omega_\ell))$ and $v_\ell = g(x_*(\omega_\ell))$. Then $(v_\ell)_+ \geq e_j$ and $u_\ell \geq \gamma$, where γ is defined by (58). Due to properties 2) and 4) of separation function h (see Assumption 7.3), we have

$$h(u_\ell, v_\ell; \omega_\ell) = h(u_\ell, (v_\ell)_+; \omega_\ell) \geq h(\gamma, e_j, \omega_\ell).$$

It follows from the property 3) of the function h that

$$h(u_\ell, v_\ell; \omega_\ell) \to +\infty \quad \text{as} \quad \ell \to +\infty. \tag{59}$$

Consider now a solution x_* of $P(f, g)$. Let $u_* = f(x_*)$, $v_* = g(x_*) \in \mathbb{R}^m_-$. Since $x_*(\omega_\ell)$ is a solution of $P_Y(f, g; \omega_\ell)$ and $x_* \in X_0 \subset Y$, we have, by applying Lemma 7.1,

$$h(u_\ell, v_\ell; \omega_\ell) = h(f(x_*(\omega_\ell)), g(x_*(\omega_\ell)); \omega_\ell) = \min_{x \in Y} h(f(x), g(x); \omega_\ell) \leq h(u_*, v_*; \omega_\ell) = u_*,$$

which contradicts (59). $\square$

Theorem 7.1. *Assume that Assumptions 7.1–7.3 hold. Then*

$$\sup_{\omega \in \Omega} \min_{x \in Y} h(f(x), g(x); \omega) = \rho,$$

where ρ is the value of $P(f, g)$.

Proof. Let $\varepsilon > 0$. Since the function f is uniformly continuous on Y, there exists $\mu > 0$ such that $\|x - x'\| \leq \mu$ implies $|f(x) - f(x')| < \varepsilon$ for all x, $x' \in Y$. Since G is upper semicontinuous, there exists $\delta > 0$ such that $G(\lambda) \subset X_0 + \mu B$ if $\lambda \leq \delta$. It follows from Assumption 7.2 and Lemma 7.2 that there exists $\omega' \in \Omega$ such that $x_*(\omega') \in G(\delta)$, hence $x_*(\omega') \in X_0 + \mu B$. Let $x(\omega') \in X_0$ be an element such that $\|x_*(\omega') - x(\omega')\| \leq \mu$. Then

$$\rho \leq f(x(\omega')) \leq f(x_*(\omega')) + \mu = \min_{x \in Y} h(f(x), g(x); \omega') + \mu.$$

Since μ is an arbitrary positive number, we have

$$\rho \leq \sup_{\omega \in \Omega} \min_{x \in Y} h(f(x), g(x); \omega).$$

We now check that the opposite inequality is valid. Let x_* be a solution of $P(f, g)$. Since $x_* \in Y$, we have, by applying Lemma 7.1, for each $\omega \in \Omega$

$$\rho = f(x_*) = h(f(x_*), g(x_*); \omega) \geq \min_{x \in Y} h(f(x), g(x); \omega).$$

Thus,

$$\rho \geq \sup_{\omega \in \Omega} \min_{x \in Y} h(f(x), g(x); \omega).$$

□

We now present a different set of conditions, which guarantees the zero duality gap property. Let $q : \mathbb{R}^{1+m} \to \mathbb{R}$ be a continuous increasing function with the following properties:

1) there exist positive numbers $a_1, \dots, a_m$ such that

$$q(u, y^1, \dots, y^m) \geq \max(u, a_1 y^m, \dots, a_m y^m), \qquad u \geq 0, \quad (y^1, \dots, y^m) \in \mathbb{R}^m;$$

2) $q(u, 0, \dots, 0) = u$ for all $u \geq 0$.

Let $\Omega = \mathbb{R}^m_+$. Define $h : \mathbb{R}^{1+m} \times \Omega$ by

$$h(u, v; \omega) = q(u, \omega^1 v^1, \dots, \omega^m v^m), \qquad (u, v) \in \mathbb{R}^{1+m}, \quad \omega \in \Omega.$$

The following result holds [19].

Theorem 7.2. *Let X be a finite-dimensional space. Consider the problem $P(f, g)$ with continuous f and g. Assume that the set X_0 of feasible elements is compact. Let $Y \in \mathcal{Y}$ and (if Y is unbounded),*

$$\lim_{\|x\| \to +\infty, \ x \in Y} f(x) = +\infty.$$

Then the zero duality gap property holds:

$$\sup_{\omega \in \Omega} \min_{x \in Y} h(f(x), g(x); \omega) = \min_{x \in X_0} f(x).$$

Consider now separation functions h of the form (41): $h(u, v) = \varphi(u, \psi(v))$. Assume that h depends on a parameter ω. First we consider the case, where only the outer convolution function φ depends on ω and give necessary and sufficient conditions for optimality.

Proposition 7.1. *Let Ω be a set of parameters such that Assumption 7.2 holds. Consider a mapping $\varphi : \mathbb{R}^2 \times \Omega \to \mathbb{R}$ and a mapping $\psi : \mathbb{R}^m \to \mathbb{R}$ such that:*

$$\varphi(u, 0; \omega) = u \ \ \textit{for all} \ \ u > 0 \ \ \textit{and} \ \ \omega \in \Omega; \tag{60}$$

$$\begin{aligned} &\varphi(u, w; \omega) \to +\infty \ \textit{as} \ \min_i \omega^i \to +\infty \qquad && \textit{for all } u > 0; \\ &\psi(v) = \psi(v_+) && \textit{for all } v \in \mathbb{R}^m. \end{aligned} \tag{61}$$

Assume also that $\varphi(u, 0)$ is increasing on $(0, +\infty)$ and ψ is increasing on $\mathbb{R}^m_+$. Let

$$h(u, v; \omega) = \varphi(u, \psi(v); \omega).$$

Then Assumption 7.3 holds for the function h.

Proof. The proof is immediate. □

Consider now a separation function

$$h(u, v; \omega) = \varphi(u, \psi(v; \omega)), \qquad (u, v) \in \mathbb{R}^{1+m}, \quad \omega \in \Omega, \tag{62}$$

where Ω is a set of parameters such that Assumption 7.2 is valid.

Proposition 7.2. *Let $\varphi : \mathbb{R}^2 \to \mathbb{R}$ and $\psi : \mathbb{R}^m \times \Omega$ be mappings with the following properties:*

1) $\varphi(u,0) = u$ for $u > 0$, φ is an increasing function on $\{(u,w) : u > 0,\ w \geq 0\}$;

2) $\psi(v,\omega) = \psi(v_+\omega)$ for all $v \in \mathbb{R}^m$ and $\omega \in \Omega$; $\psi(\cdot,\omega)$ is an increasing function on $\mathbb{R}^m_+$ for all $\omega \in \Omega$.

Then Assumption 7.3 *holds for the function h defined by* (62).

Proof. The proof is immediate. □

It follows from Proposition 7.1 and Proposition 7.2 that the zero duality gap property holds for Problem $P(f,g)$ such that Assumption 7.2 holds and a separation function h possesses properties described in one of these propositions.

8 AUXILIARY FUNCTIONS WITH SPECIAL OUTER CONVOLUTION FUNCTIONS

In this section we present a survey of results obtained in [20]. We consider an auxiliary function $M(x;\omega) = \varphi(f(x), \psi(g(x)); \omega)$ which is defined by functions φ and ψ with some special properties. Assume that $\Omega = \mathbb{R}_+$ and $\varphi(u,w;\omega) = q(u, \omega w_+)$, where q is an IPH function defined on $\mathbb{R}^2_+$ with the following properties:

$$q(1,0) = 0, \qquad \lim_{w \to +\infty} q(1,w) = +\infty. \tag{63}$$

Clearly, the first equality in (63) holds if and only if (61) is valid for the function φ; the second equality in (63) holds if and only if (61) is valid for φ. Assume that ψ is either a proper strictly exterior or proper strictly common convolution function. Then (see (43) and (44)) $X_0 = \{x \in X : \psi(g(x)) \leq 0\}$. Let

$$h(u,v;\omega) := \varphi(u,v;\omega) = q(u_+, \omega\psi_+(v)), \qquad (u,v;\omega) \in \mathbb{R}^{1+m} \times \mathbb{R}_+. \tag{64}$$

Let $Y \in \mathcal{Y}$. We assume again that

$$\inf_{x \in Y} f(x) > 0,$$

however, we do not require now some properties like uniform continuity of f on Y or upper semicontinuity of the mapping G. Instead we assume that the following assumption is valid, which shows that the single constraint $\psi(g(x)) \leq 0$ is essential.

Assumption 8.1. *There exists a sequence $x_k \in Y$ such that $\psi(g(x_k)) > 0$, $\psi(g(x_k)) \to 0$ and $f(x_k) \to \rho := \min_{x \in X_0} f(x)$.*

We also consider the perturbation function β of the problem $P(f, \psi(g))$. By definition

$$\beta(f, \psi(g); w) = \inf\{f(x) : x \in Y,\ \psi(g(x)) \leq w\}.$$

The following result holds (see [20] and also [18]).

Theorem 8.1. *Let Assumption* 8.1 *holds and let h be a separation function defined by* (64), *where q is IPH function with properties* (63). *Then the zero duality gap property holds if and only if the perturbation function β is lower semicontinuous at the origin.*

The following result demonstrates that properties (63) are essential for the validity of the zero duality gap property.

Theorem 8.2. *Let $m = 1$ and $\psi(g) = g$. Let q be a continuous IPH function defined on $\mathbb{R}^2_+$ and the zero duality gap property holds (with respect to h defined by* (64)*) for all problems $P(f,g)$ such that Assumption* 8.1 *is valid and the perturbation function β is lower semicontinuous. Then* (63) *holds.*

We now discuss the existence of $\omega \in \Omega = \mathbb{R}_+$ such that the problem

$$\min_{x \in Y} h(f(x), g(x); \omega) \to \max \quad \text{subject to} \ \ \omega \in \Omega \tag{65}$$

has a solution. A solution of this problem is called the exact penalty parameter. The following results hold (see [20] and also [18]):

Theorem 8.3. *Let $P(f,g)$ be a problem such that Assumption* 8.1 *holds and the perturbation function $\beta(f, \psi(g); \cdot)$ is lower semicontinuous. Then there exists a continuous IPH function q with properties* (63) *such that the problem* (65) *has a solution (here h defined by* (64)*).*

Theorem 8.4. *Let $m = 1$ and $\psi(g) = g$. Let Y be a non-discrete set in the following sense: there exists a function defined on Y and mapping onto $\mathbb{R}$. Then for each continuous IPH function q with properties* (63) *there exists a problem $P(f,g)$ such that Assumption* 8.1 *holds, the perturbation function $\beta(f,g;\cdot)$ is lower semicontinuous and the supremum $\sup\limits_{\omega \in \Omega} h(u, v; \omega)$ does not attain.*

Remark 8.1. *It follows from Theorem* 8.1 *that the validity of zero duality gap property does not depend on the choice of a continuous IPH function with properties* (63)*. At the same time, Theorem* 8.3 *demonstrates that the existence of $\omega' \in \Omega$ such that*

$$\min_{x \in Y} h(f(x), g(x); \omega') = \min_{x \in X_0} f(x) \tag{66}$$

does depend on this choice. Theorem 8.4 *shows that we can not find a continuous IPH function such that* (66) *holds for all problems $P(f,g)$ with the lower semicontinuous at the origin perturbation function.*

Acknowledgements

This research has been supported by Australian Research Council (Grant No. A49906152) and by Russian Foundation for Basic Research (Grants No. 99-01-01186 and No. 00-15-96080).

The authors are very thankful to Prof. F. Giannessi and anonymous referees for valuable comments.

REFERENCES

[1] M.Yu. Andramonov, (2000). An approach to constructing generalized penalty function. In: A.M. Rubinov and B.M. Glover (Eds.), *Optimization and related topics*, 3–14. Kluwer Academic Publishers.

[2] J.-P. Aubin and I. Ekeland (1984). *Applied Nonlinear Analysis*, John Wiley and Sons, New York.

[3] J. Burke (1991). An exact penalization viewpoint of constrained optimization, *SIAM J. Control and Optimization*, **29**, 968–998.

[4] V.F. Demyanov, G. Di Pillo and F. Facchinei (1998). Exact penalization via Dini and Hadamard conditional derivatives, *Optimization Methods and Software*, **9**, 19–36.

[5] L.C. Dixon and G.P. Szego (1975, 1978). *Towards Global Optimization* (part I, part II), North Holland, Amsterdam.

[6] I.I. Eremin (1967). The "penalty" method in convex programming. *Dokl. Akad. Nauk SSSR*, **173**, 748–751 (in Russian).

[7] Y.G. Evtushenko and V.G. Zhadan (1990). Exact auxiliary functions in optimization problems, *U.S.S.R. Comput. Maths. Math. Phys.*, **30**, 31–42.

[8] Y.G. Evtushenko and V.G. Zhadan (1990). Exact auxiliary functions, *Informatica*, **1**, 40–58.

[9] Y.G. Evtushenko and V.G. Zhadan (1992). Exact auxiliary functions in non-convex optimization. In: W. Oettli and D. Pallaschke (Eds.), *Advances in Optimization*, (Lecture Notes in Economics and Mathematical Systems, 382), 217–226. Springer-Verlag, Berlin.

[10] Y.G. Evtushenko, A.M. Rubinov and V.G. Zhadan (2001). General Lagrange-type functions in constrained global optimization. Part II: Exact Auxiliary functions, *Optimization Methods and Software*, **16**, 231–256.

[11] F. Giannessi (1984). Theorems of alternative and optimality conditions, *J. Optimization Theory and Applications*, **42**, 331–365.

[12] F. Giannessi (1994). General optimality conditions via a separation scheme. In: E. Spedicato (Ed.), *Algorithms for Continuous Optimization*, 1–23. Kluwer Academic Publishers.

[13] C.J. Goh and X.Q. Yang (2001). A Nonlinear Lagrangian theory for non-convex optimization, *Journal of Optimization Theory and Applications*, **109**, 99–121.

[14] S.P. Han and O.L. Mangasarian (1979). Exact penalty functions in nonlinear programming, *Mathematical Programming*, **17**, 251–269.

[15] G. Di Pillo and L. Grippo (1979). A new class of augmented Lagrangian in nonlinear programming, *SIAM J. Control Optim.*, **17**, 618–628.

[16] G. Di Pillo and L. Grippo (1989). Exact penalty functions in constrained optimization, *SIAM J. Control Optim.*, **27**, 1333–1360.

[17] R.T. Rockafellar and R.J-B. Wets (1998). *Variational Analysis.* Springer-Verlag, Berlin-Heidelberg.

[18] A.M. Rubinov (2000). *Abstract Convexity and Global Optimization.* Kluwer Academic Publishers.

[19] A.M. Rubinov, B.M. Glover and X.Q. Yang (1999). Extended Lagrange and penalty functions in continuous optimization, *Optimization*, **46**, 327–351.

[20] A.M. Rubinov, B.M. Glover and X.Q. Yang (2000). Decreasing functions with application to penalization. *SIAM Journal Optimization*, **10**, 289–313.

[21] A.M. Rubinov and A. Uderzo. On global optimality conditions via separation functions. *Journal of Optimization Theory and Applications*, (to appear).

[22] X.Q. Yang and K.L. Teo, Nonlinear Lagrangian functions and applications to semi-infinite programs, optimization and numerical algebra. *Annals of Operations Research*, (to appear).

[23] W. Zangwill (1967). Nonlinear programming via penalty function. *Management Sciences*, **13**, 344–358.