Journal of Convex Analysis
Volume 5 (1998), No. 2, 199–220

On Smoothing of Parametric Minimax-Functions and Generalized Max-Functions via Regularization

E. Levitin

Institute of System Analysis, Russian Academy of Science, 117312 Moscow, Russia.

R. Tichatschke

Department of Mathematics, University of Trier, 54286 Trier, Germany.
e-mail: tichat@uni-trier.de

Received October 17, 1996
Revised manuscript received September 29, 1997

The paper deals with a smoothing procedure for parametric minimax-functions, arising from parameter dependend convex-concave games. In this framework generalized max-functions (i.e. maximum value functions of parametric, concave programming problems) are most important special cases of parametric minimax-functions. In general, these functions are non-differentiable and, moreover, non-locally Lipschitzean.

In this paper we suggest to smooth these functions by means of a regularization approach and prove differentiability properties and error estimates of the regularized approximations of the parametric minimax-functions. Uniform convergence (w.r.t. the parameter) of the regularized solutions to the normal solutions is shown. In particular, for generalized max-functions uniform convergence of the regularized solutions to the normal solutions of the corresponding primal and dual parametric optimization problem can be concluded.

This approach gives the possibility to work without directional derivatives for these non-smooth functions and to use the usual differential calculus.

Keywords: parametric minimax problems, semi-infinite optimization, non-differentiable optimization, convex parametric optimization, duality theory, regularization

1991 Mathematics Subject Classification: 90C34, 90C31, 49M45

1. Introduction

There exist a couple of optimization problems

$$f(x) \to \min, \quad \text{s.t. } x \in Q, \tag{1.1}$$

where Q is a closed set in the Euclidean space $\mathbf{X}$ and the function f has the form

$$f(x) := \inf_{y\in\mathcal{Y}} \sup_{z\in\mathcal{Z}} F(x,y,z), \; x \in G, \tag{1.2}$$

with $Q \subset \mathcal{G} \subset \mathbf{X}$; $\mathcal{G}, \mathcal{Y}$ and $\mathcal{Z}$ closed sets in the Euclidean spaces $\mathbf{X}$, $\mathbf{Y}$and $\mathbf{Z}$, respectively. The function $F : \mathbf{X} \times \mathbf{Y} \times \mathbf{Z} \to \mathbb{R}$ is supposed to be continuous. Functions f of type (1.2) we shall call *parametric minimax-functions.*

ISSN 0944-6532 / $ 2.50

In the sequel only such functions are considered, where $\mathcal{Y}$ and $\mathcal{Z}$ are convex sets in $\mathbf{Y}$, $\mathbf{Z}$, respectively and, for each $x \in \mathcal{G}$, $F(x,\cdot,\cdot)$ is convex-concave w.r.t. $\{y,z\}$ on $\mathcal{Y} \times \mathcal{Z}$, i.e. $F(x,\cdot,z)$ is convex w.r.t. y on the set $\mathcal{Y}$ and $F(x,y,\cdot)$ is concave w.r.t. z on $\mathcal{Z}$.

Under these conditions Problem (1.1) generates a parametric convex-concave game with the papameter $x \in \mathcal{G}$. This game can be interpreted as a two-level optimization problem, where in the lower level a (non-cooperative) convex-concave game has to be solved. Indeed, for a ressource-vector $x \in G$, given from the upper level, two players on the lower level ask for a saddle point of the function F on the sets of strategies $\mathcal{Y}$ and $\mathcal{Z}$, respectively. After that, on the upper level the function F has to be minimized on the set of all (parameter dependend) equilibrium states. This kind of problems occur by modeling several hierarchical systems [9].

It is obvious that Problem (1.1) with function f of the form (1.2) is equivalent to the following semi-infinite programming problem (SIP) in the Euclidean space $\mathbb{R}^1 \times \mathbf{X} \times \mathbf{Y}$:

$$t \to \min, \text{ s.t. } t \in \mathbb{R}^1, \; x \in Q, \; y \in \mathcal{Y}, \; F(x,y,z) - t \leq 0 \quad \forall z \in \mathcal{Z}. \tag{1.3}$$

The most important partial case of a function like (1.2), generated by a parametric convex-concave game, is the following kind of max-function:

$$f(x) := \begin{cases} \sup_{z \in D(x)} h(x,z) & , \text{ if } D(x) \neq \emptyset, \\ -\infty & , \text{ if } D(x) = \emptyset, \end{cases} \tag{1.4}$$

with

$$D(x) := \{z \in \mathcal{Z} \subset \mathbf{Z} : h_s(x,z) \leq 0 \ (s \in S_1), \; h_s(x,z) = 0 \ (s \in S_2)\}. \tag{1.5}$$

Indeed, if for the Problem (with fixed x)

$$h(x,z) \to \max, \text{ s.t. } z \in D(x), \tag{1.6}$$

the strong duality theorem (cf., for instance, [10], [17]) is valid (for all $x \in \mathcal{G}$ with $D(x) \neq \emptyset$), then F is the Lagrange function of Problem (1.6):

$$F(x,y,z) := \Pi(x,y,z)^1 \tag{1.7}$$

with $\Pi(x,y,z) = h(x,z) - \sum_{s \in S} y_s h_s(x,z)$, $(S = S_1 \cup S_2)$, where

$$\mathbf{Y} = \mathbb{R}^{|S|}, \; \mathcal{Y} := \{y \in \mathbf{Y} : y_s \geq 0 \ (s \in S_1)\}. \tag{1.8}$$

This is true because of

$$f(x) = \min_{y \in \mathcal{Y}} \varphi(x,y), \text{ where } \varphi(x,y) = \max_{z \in \mathcal{Z}} \Pi(x,y,z). \tag{1.9}$$

If $S = \emptyset$, i.e. the set $D(x) = \mathcal{Z}$ does not depend on x, then f is called a *simple max-function*. In the general case, we shall call functions of type (1.4), (1.5) *generalized max-functions*.

[1]for convenience we change here the sequence of the arguments, usually one has to write $\Pi(x,z,y)$.

In this paper we shall deal only with generalized max-functions f, for which the function Π is concave w.r.t. $z \in \mathcal{Z}$ and the duality relation (1.9) is satisfied for each $x \in \mathcal{G}$. In this case the function $F(x,y,z) := \Pi(x,y,z)$ is convex-concave w.r.t. $\{y,z\}$ on $\mathcal{Y} \times \mathcal{Z}$.

It should be noted that Problem (1.1), (1.4) is equivalent to the following generalized SIP in $\mathbb{R}^1 \times \mathbf{X}$:

$$\begin{aligned} & t \to \min, \\ & \text{s.t. } x \in Q, D(x) \neq \emptyset, \\ & \qquad h(x,z) - t \leq 0 \ \forall z \in D(x). \end{aligned} \tag{1.10}$$

Finally, let us consider the following optimization problem:

$$\begin{aligned} & f_0(x) \to \min, \\ & \text{s.t. } x \in \mathcal{X}, \ f_i(x) \leq 0 \ (i \in I), \ g_j(x) = 0 \ (j \in J), \end{aligned} \tag{1.11}$$

with $\mathcal{X} \subset \mathbf{X}$ a closed convex set, I and J finite sets of indices and, for each $i \in \{0\} \cup I$,

$$f_i(x) = \begin{cases} \sup_{z_i \in D_i(x)} \ h_i(x,z_i), & \text{if } D_i(x) \neq \emptyset, \\ -\infty, & \text{if } D_i(x) = \emptyset, \end{cases} \tag{1.12}$$

$$D_i(x) = \left\{ z_i \in \mathcal{Z}_i \subset \mathbf{Z}_i \ : \ h_{is}(x,z_i) \leq 0 \ (s \in S_{i1}), \ h_{is}(x,z_i) = 0 \ (s \in S_{i2}) \right\}, \tag{1.13}$$

where $h_i, \ h_{i,s} \ (i \in \{0\} \cup I, s \in S_{i1} \cup S_{i2}), \ g_j \ (j \in J)$ are continuously differentiable functions.

It is obvious that Problem (1.11) is equivalent to the following generalized SIP in $\mathbb{R}^1 \times \mathbf{X} \times \prod_{i \in \{0\} \cup I} \mathbf{Z}_i$:

$$\begin{aligned} & t \to \min, \\ & \text{s.t. } x \in \mathcal{X}, \ D_i(x) \neq \emptyset \quad (i \in \{0\} \cup I), \\ & \qquad h_0(x,z_0) - t \leq 0 \qquad \forall z_0 \in D_0(x), \\ & \qquad h_i(x,z_i) \leq 0 \qquad\qquad \forall z_i \in D_i(x) \ (i \in I), \\ & \qquad g_j(x) = 0 \qquad\qquad\quad (j \in J). \end{aligned} \tag{1.14}$$

In particular, Problem (1.14) can be obtained from an *optimization model under uncertainty*:

$$\begin{aligned} & h_0(x,z_0) \to \min, \\ & \text{s.t. } x \in \mathcal{X}, \\ & \qquad h_i(x,z_i) \leq 0 \quad (i \in I), \\ & \qquad g_j(x) = 0 \qquad (j \in J), \end{aligned} \tag{1.15}$$

if the "principle of guaranteed results" (see [9]) is considered, where z_i are the vectors of uncertained parameters belonging to the known sets $D_i(x)$ $(i \in \{0\} \cup I)$, given via (1.13).

There exists a way to reduce the Problems (1.1), (1.2) or (1.1), (1.4), (1.5) to minimization problems with simple max-functions. This way consists in the embedding (or enlargement) of the space $\mathbf{X}$ of sought variables of the initital problem into the space $\mathbf{X} \times \mathbf{Y}$ and it is based on the following obvious formula:

$$\inf_{x\in Q} \inf_{y\in\mathcal{Y}} \sup_{z\in\mathcal{Z}} F(x,y,z) = \inf\left\{\varphi(u) : u = \{x,y\} \in Q \times \mathcal{Y}\right\}, \tag{1.16}$$

with $\varphi(u) = \sup\{F(u,z) : z \in \mathcal{Z}\}$ a simple max-function. Note that we actually have used formula (1.16) in order to apply the equivalence between Problem (1.1), (1.2) and SIP (1.3).

The embedding approach can be used to investigate optimality conditions and stability results (perturbation theory) for the Problems (1.1), (1.2) or (1.1), (1.4), (1.5). In some cases this approach is suitable for constructing numerical methods, solving Problem (1.1) (see [24]).

However, if $\dim \mathbf{X} >> \dim \mathbf{Y}$, this approach leads to an essential enlargement of the dimension of the vector of sought variables. Conversely, if $\dim \mathbf{X} << \dim \mathbf{Y}$, and one has to minimize $\varphi(u) = \sup\{F(u,z) : z \in \mathcal{Z}\}$ on the set $U = Q \times \mathcal{Y}$, and for each $x \in Q$, the determination of a saddle point of the function $F(x,y,z)$ on $\mathcal{Y} \times \mathcal{Z}$ can be performed relatively easy, then this problem is transformed frequently into the minimization of f on Q.

This approach with a succeeding approximate smoothing of the function f is especially natural in case Problem (1.1) has a block-structure, i.e. the spaces $\mathbf{Y}$ and $\mathbf{Z}$ are Euklidean products of the spaces $\mathbf{Y}_k$ and $\mathbf{Z}_k$, respectively ($k \in K, K$ a finite set)

$$\mathcal{Y} := \prod_{k\in K} \mathcal{Y}_k, \;\; \mathcal{Z} := \prod_{k\in K} \mathcal{Z}_k, \;\; (\mathcal{Y}_k \subset \mathbf{Y}_k, \;\; \mathcal{Z}_k \subset \mathbf{Z}_k)$$

and for each $x \in Q, y = \left\{y^k\right\}_{k\in K} \in \mathcal{Y}, \;\; z = \left\{z^k\right\}_{k\in K} \in \mathcal{Z}$ the function F is supposed to be block-separable:

$$F(x,y,z) = \sum_{k\in K} F_k(x,y^k,z^k).$$

Indeed, in this case x is the vector of binding variables and the following decomposition formula is obvious:

$$f(x) = \sum_{k\in K} f_k(x), \;\; \text{with } f_k(x) = \inf_{y^k\in\mathcal{Y}_k} \sup_{z^k\in\mathcal{Z}_k} F_k(x,y^k,z^k).$$

However, in this paper we don't draw attention to this embedding approach.

Optimization problems (1.1), (1.2), or Problems (1.11) with functions f_i of the form (1.12), (1.13), have always a complicated inner structure, in particular, we refer to

- parametric minimax problems (see [3], [4], [9], [10], [21], [22]);
- generalized semi-infinite problems (see [12], [13], [15], [16], [20], [22], [24], [28], [29]).

Analogous optimization problems with parametric minimax-functions or generalized max-functions appear in decomposition methods, when upper level problems are considered (see [1], [2], [18], [21], [22]).

In all these problems differentiability properties of parametric minimax functions or generalized max-functions play an important role.

As known (see [10], [22]), a parametric minimax-function is differentiable at a point $x \in \mathcal{G}$, if the sets

$$\mathcal{Y}^*(x) = \operatorname{Arg\,min}\{\varphi(x,y) : y \in \mathcal{Y}\} \quad \text{and} \quad \mathcal{Z}^*(x) = \operatorname{Arg\,max}\{\psi(x,z) : z \in \mathcal{Z}\},$$

with

$$\varphi(x,y) = \sup\{F(x,y,z) : z \in \mathcal{Z}\} \text{ and } \psi(x,z) = \inf\{F(x,y,z) : y \in \mathcal{Y}\}$$

are non-empty and consist of a single point, respectively. In particular, a generalized max-function, for which Π is concave w.r.t. $z \in \mathcal{Z}$ and the duality relation (1.9) holds, is differentiable at a point $x \in \mathcal{G}$ if the set of optimal solutions of Problem (1.6) and of its dual are non-empty and consist of a single point. However, already in case of linear programs, there may be no uniqueness of the solutions of the primal or dual problem. Even if the function ψ is strongly concave w.r.t. z and, hence, $\mathcal{Z}^*(x)$ consists of the unique point $z^*(x)$, this assumption does not suffice for the existence of a unique optimum for the dual to Problem (1.6). This non-uniqueness appears almost everywhere if

$$\operatorname{card}[S_1(x,z^*(x))] + \operatorname{card}[S_2] > \dim \mathbf{Z},$$

with $S_1(x,z^*(x))$ the set of active indices of constraints (1.5) at the point $\{x,z^*(x)\}$. Therefore, without relatively strong assumptions, it is impossible to ensure a-priorily that the optimal set of Problem (1.6) and of its dual consist of one point only.

In general, parametric minimax-functions and generalized max-functions are non-differentiable functions and, moreover, they are non-locally Lipschitzean. Usually, such a function has only a directional derivative and, in the generic case, this direction is the minimum of a parametric sublinear or the maximum of a parametric superlinear function (see [5], [10], [22]).

Under quite general assumptions the differential expansions and, in particular, the directional derivatives and quasi-subdifferentials for parametric minimax-functions and generalized max-functions, have been considered in a series of publications (see [3] - [7], [10], [20], [22], [23], [25], [26], [28], [29]).

However, the computation of these derivatives is sufficiently complicated and, moreover, as a rule these derivatives are not continuous w.r.t. the variable x. Therefore, to handle with this calculus in numerical algorithms, one deals with several difficulties. This fact makes the theoretical and numerical investigation of optimization problems, involving such functions, much more complicated.

Concerning regularization procedures in convex optimization (also SIP) we refer to [14], [19] [27], [31].

In this paper we suggest to smooth the function f by means of such a regularization procedure for the function F in (1.2) w.r.t. the variables $\{y,z\}$. In particular, in case of generalized max-functions the smoothing procedure is performed by means of regularization of the Lagrange function of Problem (1.6) w.r.t. the variable z as well as to the multiplier y. So we can use the usual differential calculus to investigate regularized approximations of non-differentiable minimax-functions.

This idea is not new. For non-parametrized problems a similar approach was considered in [30]. The same problem can be studied in [11] in connection with the treatment of modified Lagrange functions, too. However, in the papers mentioned, Problem (1.6) and its dual are considered for fixed x only, i.e. for unperturbed problems. Here we are interested in similar differentiability properties depending uniformly on the parameter x.

The paper consists of six sections. After introducing in Section 2 the main assumptions and some preliminary results, in Section 3 the regularization approach is considered and some properties of the regularized primal and dual functions, corresponding to Problem (1.1), are proved. In particular, for $\varepsilon \to 0$ and fixed $x \in G$, convergence of the regularized approximations $f(\varepsilon, x)$ to the original function $f(x)$ is shown. Section 4 deals with differentiability properties of the regularized functions and in Section 5 the convergence of the uniquely defined ε-normal solutions to the normal solutions of the regularized Problem (1.1) and of its dual is investigated. Finally, in Section 6 error estimates for regularized parametric minimax and max-functions and a statement on the exactness of the minimum point of a regularized parametric max-function is proved.

All the results, obtained for parametric minimax-functions, are true also for generalized max-functions if they satify the Assumptions 2.4–2.6 below. However, it should be noted that, if F is the Lagrange function of Problem (1.6) and $\mathcal{Y} = \mathbb{R}_+^{|S_1|} \times \mathbb{R}^{|S_2|}$, then the affinity w.r.t. y of the function F and the simple form of the set $\mathcal{Y}$ permit to specify some properties of the regularized, generalized max-function. In particular, in this case this max-function is a parametrized (with parameter $x \in \mathcal{G}$) modified Lagrange function for the parametric convex programming problem (1.6).

2. General assumptions and preliminary results

Some general assumptions are formulated and discussed for which we suppose that they are satisfied throughout the whole paper.

For $x \in \mathcal{G}, y \in \mathcal{Y}, z \in \mathcal{Z}$ let

$$\varphi(x,y) := \sup\{F(x,y,z) : z \in \mathcal{Z}\}, \ \Phi(x,y) := \operatorname{Arg\,max}\{F(x,y,z) : z \in \mathcal{Z}\}, \tag{2.1}$$

$$\mathcal{Y}^*(x) := \operatorname{Arg\,min}\{\varphi(x,y) : y \in \mathcal{Y}\}, \tag{2.2}$$

$$\psi(x,z) := \inf\{F(x,y,z) : y \in \mathcal{Y}\}, \ \Psi(x,z) := \operatorname{Arg\,min}\{F(x,y,z) : y \in \mathcal{Y}\}, \tag{2.3}$$

$$\mathcal{Z}^*(x) := \operatorname{Arg\,max}\{\psi(x,z) : z \in \mathcal{Z}\}. \tag{2.4}$$

Obviously, $f(x) = \inf\{\varphi(x,y) : y \in \mathcal{Y}\}$.

It should be noted that, if $x \in \mathcal{G}, y \in \mathcal{Y}$ and $\Phi(x,y) \neq \emptyset$, then the function φ as the supremum of a family of continuous functions is lower semi-continuous on the set $\mathcal{G} \times \mathcal{Y}$ at the point $\{x,y\}$. Further, for each $x \in \mathcal{G}, y \in \mathcal{Y}$ it holds that $\varphi(x,y) < +\infty$ and $\Phi(x,y)$ is a non-empty and compact subset in $\mathbf{Z}$ if $\mathcal{Z}$ is compact or F is coercive w.r.t. z. Finally note that, if for any $x \in \mathcal{G}$ there exists a vector $z(x) \in \mathcal{Z}$ such that $F(x,\cdot,z(x))$ is coercive, then $\varphi(x,\cdot)$ is coercive, and so $\mathcal{Y}^*(x)$ is non-empty and compact.

Analogously, if $\Psi(x,z) \neq \emptyset$ for any $x \in \mathcal{G}, z \in \mathcal{Z}$, then ψ is upper semi-continuous on the set $\mathcal{G} \times \mathcal{Z}$ at the point $\{x,z\}$. If F is coercive in y for each fixed $x \in \mathcal{G}, z \in \mathcal{Z}$ or $\mathcal{Y}$ is compact, then for any $x \in \mathcal{G}, z \in \mathcal{Z}$ it holds $\psi(x,z) > -\infty$ and $\Psi(x,z)$ is a non-empty compact subset in $\mathcal{Y}$.

Also analogously, if for any $x \in \mathcal{G}$ there exists a vector $y(x) \in \mathcal{Y}$ such that F is coercive (in the sense $F(x,y(x),\cdot) \to -\infty$), then ψ is coercive (in the sense $\psi(x,\cdot) \to -\infty$) and $\mathcal{Z}^*(x)$ is a non-empty and compact set.

In the sequel we suppose that for the parametric minimax-function (1.1) the following assumptions are satisfied.

Assumption 2.1. $\mathcal{Y}$ and $\mathcal{Z}$ are convex, closed sets in $\mathbf{Y}$ and $\mathbf{Z}$, respectively; F is continuous in $\mathbf{X} \times \mathbf{Y} \times \mathbf{Z}$ and continuously differentiable w.r.t. $x; F$ is convex w.r.t. $y \in \mathcal{Y}$ for all $x \in \mathcal{G}, z \in \mathcal{Z}$ and concave w.r.t. $z \in \mathcal{Z}$ for all $x \in \mathcal{G}, y \in \mathcal{Y}$.

Assumption 2.2. If for $x \in \mathcal{G}, y \in \mathcal{Y}$ $\varphi(x,y) < +\infty$, then $\Phi(x,y)$ is non-empty and $\mathcal{Y}^*(x)$ is non-empty and bounded in $\mathbf{Y}$.

Assumption 2.3. If for $x \in \mathcal{G}, z \in \mathcal{Z}$ $\psi(x,z) > -\infty$, then $\Psi(x,z)$ is non-empty and $\mathcal{Z}^*(x)$ is non-empty and bounded in $\mathbf{Z}$.

Assumption 2.2 is certainly satisfied if, for each $x \in \mathcal{G}, y \in \mathcal{Y}$ such that $\varphi(x,y) < +\infty$, the set $\{z \in \mathcal{Z} : F(x,y,z) \geq \varphi(x,y) - \delta\}$ is bounded in $\mathbf{Z}$ for some $\delta > 0$ (for example, $F(x,y,z_n) \to -\infty$ when $x \in \mathcal{G}, y \in \mathcal{Y}$ and $z_n \in \mathcal{Z}, ||z_n|| \to \infty$) and for any $x \in \mathcal{G}$ there exists $z(x) \in \mathcal{Z}$ such that $F(x,y_n,z(x)) \to +\infty$, when $y_n \in \mathcal{Y}, ||y_n|| \to \infty$.

Analogously, Assumption 2.3 is fulfilled if, for each $x \in \mathcal{G}, z \in \mathcal{Z}$ such that $\psi(x,z) > -\infty$, the set $\{y \in \mathcal{Y} : F(x,y,z) \leq \psi(x,z) + \delta\}$ is bounded in $\mathcal{Y}$ for some $\delta > 0$ (for example, $F(x,y,z_n) \to +\infty$ when $x \in \mathcal{G}, z \in \mathcal{Z}$ and $y_n \in \mathcal{Y}, ||y_n|| \to \infty$) and for any $x \in \mathcal{G}$ there exists $y(x) \in \mathcal{Y}$ such that $F(x,y(x),z_n) \to -\infty$, when $z_n \in \mathcal{Z}, ||z_n|| \to \infty$.

From Assumption 2.1 it follows that, for all $x \in \mathcal{G}$, the function φ is convex w.r.t. $y \in \mathcal{Y}$ and ψ is concave w.r.t. $z \in \mathcal{Z}$. Therefore, for each $x \in \mathcal{G}$, the sets $\mathcal{Y}^*(x)$ and $\mathcal{Z}^*(x)$ are convex. Under the Assumptions 2.1–2.3, for each $x \in \mathcal{G}$, the sets $\mathcal{Y}^*(x)$ and $\mathcal{Z}^*(x)$ are closed, because φ is lower semi-continuous on the set $\{y \in \mathcal{Y} : \varphi(x,y) < +\infty\}$ and ψ is upper semi-continuous on the set $\{z \in \mathcal{Z} : \psi(x,z) > -\infty\}$.

Thus, under the Assumptions 2.1–2.3, for each $x \in \mathcal{G}$, the sets $\mathcal{Y}^*(x), \mathcal{Z}^*(x)$ are convex and compact in $\mathbf{Y}, \mathbf{Z}$, respectively. The Assumptions 2.1–2.3 ensure in addition that the point-to-set maps $\mathcal{Y}^*(x)$ and $\mathcal{Z}^*(x)$ are upper semi-continuous on the set $\mathcal{G}$ (see Theorem 15 in [10]). Therefore, if $x_k \in \mathcal{G}$, $x_k \to x, y_k \in \mathcal{Y}^*(x_k)$, $z_k \in \mathcal{Z}^*(x_k)$, then $\overline{\lim}_{k\to\infty}||y_k|| < +\infty$, $\overline{\lim}_{k\to\infty}||z_k|| < +\infty$.

The Assumptions 2.1–2.3 guarantee also that, for each $x \in \mathcal{G}$, the minimax theorem (see for instance, [10], [17])

$$\min_{y\in\mathcal{Y}} \max_{z\in\mathcal{Z}} F(x,y,z) = \max_{z\in\mathcal{Z}} \min_{y\in\mathcal{Y}} F(x,y,z)$$

is true.

Now, let us formulate analogous assumptions for generalized max-functions of type (1.4), (1.5).

Assumption 2.4. $\mathcal{Z}$ is a convex, closed set in $\mathbf{Z}$; the functions h, h_s $(s \in S)$ are continuously differentiable in $\mathbf{X} \times \mathbf{Z}$; h, $-h_s$ $(s \in S_1)$ are concave w.r.t. $z \in \mathcal{Z}$ for all $x \in \mathcal{G}$; h_s $(s \in S_2)$ are affine w.r.t $z \in \mathcal{Z}$ for all $x \in \mathcal{G}$, i.e., $h_s(x,z) := \langle a_s(x), z\rangle + b_s(x)$, where $a_s(\cdot) : \mathbf{X} \to \mathbf{Z}$ and $b_s(\cdot) : \mathbf{X} \to \mathbb{R}^1$ are continuously differentiable in $\mathbf{X}$.

Assumption 2.4 ensures that Problem (1.6) is a convex optimization problem for each $x \in \mathcal{G}$.

Assumption 2.5. If $\sup\{\Pi(x,y,z) : z \in \mathcal{Z}\} < +\infty$ for $x \in \mathcal{G}, y \in \mathcal{Y}$, then the set $\operatorname{Arg\,max}\{\Pi(x,y,z) : z \in \mathcal{Z}\}$ is non-empty and $\operatorname{Arg\,min}\{\sup_{z\in\mathcal{Z}} \Pi(x,y,z) : y \in \mathcal{Y}\}$ is non-empty and bounded in $\mathbf{Y}$.

Assumption 2.6. If $D(x) \neq \emptyset$ for $x \in \mathcal{G}$, then the set $\operatorname{Arg\,max}\{h(x,z) : z \in D(x)\}$ is non-empty.

As known from the theory of convex analysis, Assumption 2.5 is satisfied if Π is coercive (in the sense $\Pi(x,y,z_n) \to -\infty$ for each $x \in \mathcal{G}, y \in \mathcal{Y}, \{z_n\} \in \mathbf{Z}$, with $z_n \in \mathcal{Z}, ||z_n|| \to \infty$) and if the following regularity assumption is fulfilled for the set $D(x)$:

(R) For each $x \in \mathcal{G}$ the zero vector of the space $\mathbb{R}^{|S_2|}$ is an interior point of the set

$$\left\{\ell \in \mathbb{R}^{|S_2|} : \ell_s = \langle a_s(x), t\rangle + b_s(x),\ (s \in S_2),\ z \in \mathcal{Z}\right\}$$

and there exists a point $\tilde{z}(x) \in D(x)$ such that $h_s(x, \tilde{z}(x)) < 0$ $(s \in S_1)$.

Note that, in case $S_2 = \emptyset$ (i.e. Problem (1.6) is given without equality constraints in $D(x)$), regularity of the set $D(x)$ (for a given $x \in \mathcal{G}$) is equivalent to the modified Slater condition.

Regularity of $D(x)$ ensures that constants $c_1(x) > 0$ and $c_2(x) > 0$ exist such that

$$\varphi(x,y) \geq c_1(x) + c_2(x)\left(\sum_{s\in S_1} y_s + \sum_{s\in S_2} |y_s|\right) \quad \forall x \in \mathcal{G}, \forall y \in \mathcal{Y} \tag{2.5}$$

(cf. [17], Chapt. 4, Theor. 6.5). From (2.5) it follows that for fixed $x \in \mathcal{G}$

$$\varphi(x,y) \to +\infty \text{ for } y \in \mathcal{Y},\quad \left(\sum_{s\in S_1} y_s + \sum_{s\in S_2} |y_s|\right) \to +\infty,$$

hence, Assumption 2.6 is fulfilled.

Assumption 2.6 is also fulfilled, if $\mathcal{Z}$ is compact or h is coercive on $D(x)$ (in the sense $h(x,z_n) \to -\infty$ for each $x \in \mathcal{G}$ and any sequence $\{z_n\} \in \mathbf{Z}$ such that $z_n \in D(x), ||z_n|| \to \infty$).

Proposition 2.7. *Suppose that the generalized max-function* (1.4), (1.5) *satisfies the Assumptions* 2.4–2.6. *Let* $\mathbf{Y} = \mathbb{R}^{|S|}, \mathcal{Y} = \{y \in \mathbf{Y} : y_s \geq 0\ \forall s \in S_1\}$, $F(x,y,z) = h(x,z) - \sum_{s\in S} y_s h_s(x,z)$ *for* $x \in \mathbf{X}, y \in \mathbf{Y}, z \in \mathbf{Z}$. *Then* F *satisfies the Assumptions* 2.1–2.3.

Proof. Let the functions φ, ψ and the sets $\Phi(x,y), \Psi(x,z), \mathcal{Y}^*(x)$, and $\mathcal{Z}^*(x)$ be defined by (2.1) - (2.4). Then it is obvious that φ is the dual function of Problem (1.6), $\Phi(x,y)$

is the set of maximum points of the Lagrange function $\Pi(x, y, z)$ on the set $\mathcal{Z}$, and $\mathcal{Y}^*(x)$ is the set of optimal solutions for the dual problem to Problem (1.6). Moreover, for each $x \in \mathcal{G}, z \in \mathcal{Z}$, due to the affinity of $\Pi(x, \cdot, z)$ and the form of $\mathcal{Y}$ in (1.8), we have

$$\psi(x, z) = \begin{cases} h(x, z) & , \text{ if } z \in D(x), \\ -\infty & , \text{ if } z \notin D(x), \end{cases}$$

$$\Psi(x, z) = \begin{cases} \{y \in \mathcal{Y} : y_s = 0 \ (s \in S_1), \ h_s(x, z) < 0\} & , \text{ if } z \in D(x), \\ \emptyset & , \text{ if } z \notin D(x). \end{cases}$$

Therefore, $\mathcal{Z}^*(x) = \operatorname{Arg\,max}\{h(x, z) : z \in D(x)\}$, if $D(x) \neq \emptyset$. Now, the statement follows from these facts. □

In the sequel we suppose that for the generalized max-function (1.4), (1.5) the Assumptions 2.4–2.6 are fulfilled. Therefore, all statements which can be proved for parametric minimax-functions under the Assumptions 2.1–2.3 are also valid for generalized max-functions.

A function ϑ in the Euclidean space $\mathbf{T}$ is called *strongly convex on the convex set* $\Theta \subset \mathbf{T}$ *with the constant* κ ($\kappa > 0$), if

$$\vartheta\left(\frac{1}{2}(t' + t'')\right) \leq \frac{1}{2}\vartheta(t') + \frac{1}{2}\vartheta(t'') - \kappa||t' - t''||^2 \quad \forall t', t'' \in \Theta.$$

Function ϑ is called *strongly concave on the convex set* Θ, if $-\vartheta$ is strongly convex on Θ.

The following two statements are essential in order to investigate the smoothing procedure, suggested below.

Theorem 2.8. *Let $\mathcal{U}$ be a closed set in the Euclidean space $\mathbf{U}$, $\mathcal{V}$ be a closed convex set in the Euclidean space $\mathbf{V}$. Assume that ϑ is a continuous function on $\mathbf{U}\times\mathbf{V}$, continuously differentiable w.r.t. u and, for each $u \in \mathcal{U}$, strongly convex w.r.t. v on the set $\mathcal{V}$ with the constant κ independent of $u \in \mathcal{U}$. Then the following statements are true:*

(i) *for each $u \in \mathcal{U}$ the minimum w.r.t. v of the function ϑ on the set $\mathcal{V}$ is attained in the unique point $v(u)$, and*

$$\vartheta(u, v) - \vartheta(u, v(u)) \geq 4\kappa||v - v(u)||^2 \quad \forall v \in \mathcal{V}; \tag{2.6}$$

(ii) *the operator $v(u)$ is continuous on $\mathcal{U}$;*

(iii) *if, additionally to the assumptions made above, for each $u \in \mathcal{U}$ there exist $\rho = \rho(u) > 0$ and $L = L(u) > 0$ such that on the set $\{(u', v') \in \mathcal{U} \times \mathcal{V} : ||u' - u|| \leq \rho, ||v' - v(u)|| \leq \rho\}$ the gradient $\nabla_u \vartheta(u, v)$ is Lipschitz with the constant L, then for each $u \in \mathcal{U}$ the operator $v(\cdot)$ is Lipschitz on the set $\{u' \in \mathcal{U} : ||u' - u|| \leq \rho\}$ with the constant $(4\kappa)^{-1}L$.*

Proof. Existence and uniqueness of the point $v(u)$ follows from Theorem 1.15, Chapter 4 in [22] and estimate (2.6) is a consequence of Theorem 2.11, Chapter 4 in [22]. Statement (ii) follows from (2.6) and the continuity of the function $\vartheta(u, v(u))$ on $\mathcal{U}$. Finally, the Lipschitz property of the operator $v(\cdot)$ with constant $(4\kappa)^{-1}L$ is due to Theorem 6.1 in [23]. □

Recall that in the Euclidean space the squared norm is a strongly convex function on the whole space with the constant $\kappa = \frac{1}{4}$. Thus, from Theorem 2.8 (i) it follows that, for each $x \in \mathcal{G}$, the function $||z||^2$ has a unique minimizer on the closed set $\mathcal{Z}^*(x)$, and $||y||^2$ has a unique minimizer on the closed set $\mathcal{Y}^*(x)$.

The vectors

$$\hat{z}(x) := \arg\min\left\{||z||^2 : z \in \mathcal{Z}^*(x)\right\} \text{ and } \hat{y}(x) := \arg\min\left\{||y||^2 : y \in \mathcal{Y}^*(x)\right\} \tag{2.7}$$

are called *normal points* of the sets $\mathcal{Z}^*(x)$ and $\mathcal{Y}^*(x)$, respectively.

From (2.6) with $\kappa = \frac{1}{4}$ it follows

$$||z - \hat{z}(x)||^2 \leq ||z||^2 - ||\hat{z}(x)||^2 \quad \forall z \in \mathcal{Z}^*(x) \tag{2.8}$$

and

$$||y - \hat{y}(x)||^2 \leq ||y||^2 - ||\hat{y}(x)||^2 \quad \forall y \in \mathcal{Y}^*(x). \tag{2.9}$$

Let $\mathcal{U}$ be a closed set in the Euclidean space $\mathbf{U}, u_0 \in \mathcal{U}, m$ be a continuous function on $\mathcal{U}$. The function m is called *$\mathcal{U}$-differentiable at the point u_0* if there exists a vector $\ell \in \mathbf{U}$ such that

$$m\left(u_n\right) = m\left(u_0\right) + \langle \ell, u_n - u_0 \rangle + o\left(||u_n - u_0||\right) \ \forall \left\{u_n\right\} \in \mathcal{U}, u_n \to u_0.$$

The vector ℓ is called *$\mathcal{U}$-gradient of m in the point u_0* and it is denoted by $\nabla^{\mathcal{U}} m(u_0)$.

Theorem 2.9. *Assume that the hypotheses of* Theorem 2.8 *are satisfied. Then the function*

$$m(u) := inf\{\vartheta(u,v) : \ v \in \mathcal{V}\}$$

is $\mathcal{U}$-differentiable in each point $u \in \mathcal{U}$ and its gradient $\nabla^{\mathcal{U}} m(u) = \nabla_u \vartheta(u, v(u))$ is continuous on the set U. If, moreover, for each $u \in \mathcal{U}$, constants $\rho = \rho(u) > 0$ and $L = L(u) > 0$ exist such that on the set $\{(u', v') \in \mathcal{U} \times \mathcal{V} : \|u' - u\| \leq \rho, \ \|v' - v(u)\| \leq \rho\}$ the gradient $\nabla_u \vartheta(u, v(u))$ is Lipschitz with the constant L, then for each $u \in \mathcal{U}$ the vector-function $\nabla m(u)$ is Lipschitz on the set $\{u' \in \mathcal{U} : \|u' - u\| \leq \rho\}$.

Proof. On the one hand, if $u_k, u \in \mathcal{U}$ and $u_k \to u$, then

$$\begin{aligned} m(u_k) \leq \vartheta\left(u_k, v(u)\right) &= \vartheta(u, v(u)) + \langle \nabla_u \vartheta(u, v(u)), u_k - u \rangle + o\left(||u_k - u||\right) \\ &= m(u) + \langle \nabla_u \vartheta\left(u, v(u)\right), u_k - u \rangle + o\left(||u_k - u||\right). \end{aligned}$$

On the other hand, due to the continuous differentiability of ϑ w.r.t. u, and $v(u_k) \to v(u)$ (which follows from Assumption (ii) in Theorem 2.8), we obtain

$$\begin{aligned} m(u_k) &= \vartheta\left(u_k, v(u_k)\right) = \vartheta\left(u, v(u_k)\right) + \langle \nabla_u \vartheta\left(u, v(u_k)\right), u_k - u \rangle \\ &+ \int_0^1 \langle \nabla_u \vartheta\left(u + t(u_k - u), v(u_k)\right) - \nabla_u \vartheta\left(u, v(u_k)\right), u_k - u \rangle dt \\ &\geq \inf\left\{\vartheta(u,v) : v \in \mathcal{V}\right\} + \langle \nabla_u \vartheta\left(u, v(u)\right), u_k - u \rangle + o\left(||u_k - u||\right). \end{aligned}$$

Hence, $\nabla_u \vartheta(u, v(u))$ is the gradient of m in the point u.

Finally, under the assumptions made, Theorem 2.8 permits to conclude that the operator $\nabla m(u)$ is Lipschitz on the set $\{u' \in \mathcal{U} : ||u' - u|| \leq \rho\}$. □

3. Regularization of parametric minimax-functions and generalized max-functions

For arbitrarily fixed $\varepsilon > 0$ we consider the following regularization of the parametric minimax-function

$$f(\varepsilon, x) := \min_{y \in \mathcal{Y}} \max_{z \in \mathcal{Z}} F(\varepsilon, x, y, z), \tag{3.1}$$

with

$$F(\varepsilon, x, y, z) := F(x, y, z) + \varepsilon ||y||^2 - \varepsilon ||z||^2 \tag{3.2}$$

(obviously, $f(\varepsilon, x)_{|\varepsilon=+0} = f(x)$).

In this way $F(\varepsilon, x, y, z)$ becomes strongly convex w.r.t. y and strongly concave w.r.t. z for each $\varepsilon > 0$, and we have the possibility to construct the following smooth functions:

$$\varphi(\varepsilon, x, y) := \max_{z \in \mathcal{Z}} F(\varepsilon, x, y, z), \quad \psi(\varepsilon, x, z) := \min_{y \in \mathcal{Y}} F(\varepsilon, x, y, z). \tag{3.3}$$

It is clear that

$$\varphi(\varepsilon, x, y) = \varepsilon ||y||^2 + \max\{F(x, y, z) - \varepsilon ||z||^2 : z \in \mathcal{Z}\}, \tag{3.4}$$

$$\psi(\varepsilon, x, z) = -\varepsilon ||z||^2 + \min\{F(x, y, z) + \varepsilon ||y||^2 : y \in \mathcal{Y}\}, \tag{3.5}$$

and $\varphi(\varepsilon, x, y)_{|\varepsilon=+0} = \varphi(x, y), \psi(\varepsilon, x, z)_{|\varepsilon=+0} = \psi(x, z)$ $\forall x \in \mathcal{G}, y \in \mathcal{Y}, z \in \mathcal{Z}$.

Due to the definition of the functions φ, ψ and the minimax theorem it holds for arbitrary $\varepsilon > 0$

$$f(\varepsilon, x) = \min\{\varphi(\varepsilon, x, y) : y \in \mathcal{Y}\} = \max\{\psi(\varepsilon, x, z) : z \in \mathcal{Z}\}. \tag{3.6}$$

From Theorem 2.8 it follows that the sets

$$\Phi(\varepsilon, x, y) := \operatorname{Arg\,min}\{F(\varepsilon, x, y, z) : z \in \mathcal{Z}\}$$

and

$$\Psi(\varepsilon, x, z) := \operatorname{Arg\,min}\{F(\varepsilon, x, y, z) : y \in \mathcal{Y}\}$$

consist of the single points $z(\varepsilon, x, y)$ and $y(\varepsilon, x, z)$, respectively.

Obviously, for generalized max-function (1.4) the regularization (3.1), (3.2) takes the form:

$$f(\varepsilon, x) = \min_{y \in \mathcal{Y}} \max_{z \in \mathcal{Z}} \left\{ h(x, z) - \sum_{s \in S} y_s h_s(x, z) + \varepsilon \sum_{s \in S} y_s^2 - \varepsilon ||z||^2 \right\}. \tag{3.7}$$

It should be noted that the values of ψ can be computed explicitely. Indeed, the equation

$$\min \left\{ \sum_{s \in S} \left(\varepsilon y_s^2 - y_s h_s(x, z) \right) : y_s \geq 0 \ (s \in S_1), \ y_s \in \mathbb{R}^1 \ (s \in S_2) \right\}$$

$$= \sum_{s\in S_1} \min_{y_s\geq 0} \{\varepsilon y_s^2 - y_s h_s(x,z)\} + \sum_{s\in S_2} \min_{y_s\in\mathbb{R}^1} \{\varepsilon y_s^2 - y_s h_s(x,z)\}$$

$$= -(4\varepsilon)^{-1} \left(\sum_{s\in S_1} [h_s^+(x,z)]^2 + \sum_{s\in S_2} h_s^2(x,z) \right),$$

with $h_s^+(x,z) = \max\{h_s(x,z),\ 0\}$, leads to

$$\psi(\varepsilon,x,z) = -\varepsilon||z||^2 + h(x,z) - (4\varepsilon)^{-1} \left(\sum_{s\in S_1} \left[h_s^+(x,z)\right]^2 + \sum_{s\in S_2} h_s^2(x,z) \right). \tag{3.8}$$

Therefore, for arbitrary $\varepsilon > 0$, the function ψ corresponds to the penalized Problem (1.6) with a regularized quadratic penalty and the penalty coefficient $(4\varepsilon)^{-1}$.

Now, let us return to the general case of parametric minimax-functions. We formulate some simple properties of the perturbed functions φ and ψ.

Theorem 3.1. *The following statements are true:*

(i) *For arbitrary $\varepsilon > 0, x \in \mathcal{G}, y \in \mathcal{Y}$ the function $F(x,y,z) - \varepsilon||z||^2$ is strongly concave w.r.t. z on the set $\mathcal{Z}$ (with the coefficient $\kappa(\varepsilon) = \frac{1}{4}\varepsilon$) and attains its maximum w.r.t. $z \in \mathcal{Z}$ in the unique point $z(\varepsilon,x,y)$. Moreover, for each fixed $\varepsilon > 0$, the vector-function $z(\varepsilon,x,y)$ is continuous w.r.t. $\{x,y\}$ on the set $\mathcal{G}\times\mathcal{Y}$.*

(ii) *For fixed $\varepsilon > 0, x \in \mathcal{G}$, the function φ is strongly convex w.r.t. y on the set $\mathcal{Y}$ with the coefficient $\frac{1}{4}\varepsilon$, independent of $x \in \mathbf{X}$.*
φ attains its minimum w.r.t. $y \in \mathcal{Y}$ in the unique point $y(\varepsilon,x)$. Moreover, for fixed $\varepsilon > 0$, the vector-function $y(\varepsilon,x)$ is continuous w.r.t. x on the set $\mathcal{G}$.

(iii) *For each $\varepsilon > 0, x \in \mathcal{G}$, the function ψ is strongly concave w.r.t. z on the set $\mathcal{Z}$ with the coefficient $\frac{1}{4}\varepsilon$, independent of $x \in \mathbf{X}$.*
ψ attains its minimum w.r.t. $z \in \mathcal{Z}$ in the unique point $z(\varepsilon,x)$. For fixed $\varepsilon > 0$, the vector-function $z(\varepsilon,x)$ is continuous w.r.t. x on the set $\mathcal{G}$.

Proof. (i) According to the concavity of the function F w.r.t. $z \in \mathcal{Z}$, for each $x \in \mathcal{G}, y \in \mathcal{Y}$, the function $F(x,y,z) - \varepsilon||z||^2$ is strongly concave w.r.t. $z \in \mathcal{Z}$ with the coefficient $\frac{1}{4}\varepsilon$. Now, for arbitrary fixed $\varepsilon > 0$, put $\mathbf{U} := \mathbf{X}\times\mathbf{Y}$, $\mathcal{U} := \mathcal{X}\times\mathcal{Y}$, $\mathbf{V} := \mathbf{Z}$, $\mathcal{V} := \mathcal{Z}$, and for arbitrary $u = (x,y) \in \mathcal{U}$, $v := z \in \mathcal{Z}$ set $\vartheta(u,v) := -F(x,y,z) + \varepsilon||z||^2$. Thus, we obtain

$$\min\{\vartheta(u,v) : v \in \mathcal{V}\} = -\max\{F(\varepsilon,x,y,z) : z \in \mathcal{Z}\}$$

and statement (i) immediately follows from Theorem 2.8 (i) and (ii).

(ii) The function $\max\{F(x,y,z) - \varepsilon||z||^2 : z \in \mathcal{Z}\}$ is convex in y, because it is the maximum (w.r.t. $z \in \mathcal{Z}$) of a family of affine functions in y. Therefore, for arbitrarily fixed $\varepsilon > 0, x \in \mathcal{G}$, the function φ is stronly convex w.r.t. y (on the set $\mathcal{Y}$) with the coefficient $\frac{1}{4}\varepsilon$, independent of $x \in \mathbf{X}$. Hence, statement (ii) immediately follows from the first two statements in Theorem 2.8, too.

(iii) Because the function $\min\{F(x,y,z) + \varepsilon||y||^2 : y \in \mathcal{Y}\}$ is concave in z, as minimum (w.r.t. $y \in \mathcal{Y}$) of a family of concave functions in z, we conclude that for arbitrarily fixed $\varepsilon > 0, x \in \mathcal{G}$, the function ψ is strongly concave w.r.t. z (on the set $\mathcal{Z}$) with the coefficient $\frac{1}{4}\varepsilon$, independent of $x \in \mathbf{X}$. Hence, statement (iii) immediately follows from Theorem 2.8, too. □

In Section 4 we will prove that for arbitrarily fixed $x \in \mathcal{G}$ the estimate

$$|f(\varepsilon, x) - f(x)| \leq O(\varepsilon)$$

is true and that the function $f(\varepsilon, x)$ is differentiable w.r.t. x on $\mathcal{G}$ for arbitrarily fixed $\varepsilon > 0$. Thus, the basic advantage of the regularization consists in a smooth approximation of a, in general, non-smooth function f.

Theorem 3.2. *For arbitrary $\varepsilon > 0, x \in \mathcal{G}$ the equality*

$$\min_{y \in \mathcal{Y}} \varphi(\varepsilon, x, y) = \max_{z \in \mathcal{Z}} \psi(\varepsilon, x, z) \tag{3.9}$$

holds true.

Proof. Due to Theorem 3.1 (ii) for each fixed $\varepsilon > 0, x \in \mathcal{G}$ the function φ attains its minimum w.r.t. y on the set $\mathcal{Y}$ in the unique point $y(\varepsilon, x)$. Hence, on account of the relations (cf. (3.3))

$$\min_{y \in \mathcal{Y}} \varphi(\varepsilon, x, y) = \min_{y \in \mathcal{Y}} \max_{z \in \mathcal{Z}} \; F(\varepsilon, x, y, z), \quad \max_{z \in \mathcal{Z}} \psi(\varepsilon, x, z) = \max_{z \in \mathcal{Z}} \min_{y \in \mathcal{Y}} F(\varepsilon, x, y, z),$$

equality (3.9) follows from the minimax-theorem (cf., for instance [10], [17]). □

Remark 3.3. From Theorem 3.2 and formula (3.8) it follows for a generalized max-function f that

$$f(\varepsilon, x) = \max_{z \in \mathcal{Z}} \left\{ -\varepsilon ||z||^2 + h(x, z) - (4\varepsilon)^{-1} \left(\sum_{s \in S_1} \left[h_s^+(x, z) \right]^2 + \sum_{s \in S_2} h_s^2(x, z) \right) \right\}, \tag{3.10}$$

i.e., $f(\varepsilon, x)$ is a max-function on the set $\mathcal{Z}$. Observe that here $\mathcal{Z}$ does not depend on x and, hence, $f(\varepsilon, x)$ is a simple max-function.

Theorem 3.4. *Let $\varepsilon \to +0, x \in \mathcal{G}, x \to x_0$. Then the relations*

$$\begin{aligned} \operatorname{dist}(z(\varepsilon, x), \mathcal{Z}^*(x_0)) &:= \min \{||z(\varepsilon, x) - z'|| : z' \in \mathcal{Z}^*(x_0)\} \to 0, \\ \operatorname{dist}(y(\varepsilon, x), \mathcal{Y}^*(x_0)) &:= \min \{||y(\varepsilon, x) - y'|| : y' \in \mathcal{Y}^*(x_0)\} \to 0 \end{aligned} \tag{3.11}$$

are true.

Proof. Let $w = (\varepsilon, x)$ be a vector of the closed set $\mathcal{W} = \mathbb{R}^1 \times \mathcal{G}$ in the Euclidean space $\mathbf{W} = \mathbb{R}^1 \times \mathbf{X}$. The function $F(\varepsilon, x, y, z) = F(w, y, z)$ is continuous on the Cartesian product $\mathbf{W} \times \mathbf{Y} \times \mathbf{Z}$. Moreover, on the set $\mathcal{W} \times \mathcal{Y} \times \mathcal{Z}$ this function is concave w.r.t. z (strongly concave w.r.t. z if $\varepsilon > 0$) and convex w.r.t. y (strongly convex w.r.t. y if $\varepsilon > 0$). Now we consider for arbitrary $w = (\varepsilon, x) \in \mathcal{W}$ the sets

$$\mathcal{Y}^*(w) := \operatorname{Arg\,min}\{\varphi(w, y) : y \in \mathcal{Y}\}, \;\; \mathcal{Z}^*(w) := \operatorname{Arg\,max}\{\psi(w, z) : z \in \mathcal{Z}\}.$$

Let $w_0 = (0, x_0)$. Obviously, $w_0 \in \mathcal{W}$, due to the closeness of $\mathcal{G}$. According to the Assumptions 2.2 and 2.3, the sets $\mathcal{Y}^*(w_0)$ and $\mathcal{Z}^*(w_0)$ are non-empty and bounded. Therefore, due to Theorem 15 in [10], the sets $\mathcal{Y}^*(w)$ and $\mathcal{Z}^*(w)$ are also non-empty if $\varepsilon > 0, x \in \mathcal{G}$ and the values ε and $||x - y_0||$ are chosen sufficiently small. Moreover, the point-to-set mappings $\mathcal{Y}^*(\cdot)$ and $\mathcal{Z}^*(\cdot)$ are upper semi-continuous on the set $\mathcal{W}$ in the point w_0. Now, using Theorem 3.1, with $\varepsilon > 0, x \in \mathcal{G}, w = (\varepsilon, x)$, one can conclude that $\mathcal{Z}^*(w)$ consists of the unique point $z(\varepsilon, x)$ and $\mathcal{Y}^*(w)$ of the unique point $y(\varepsilon, x)$, respectively. Using the upper semi-continuity of the point-to-set-mappings $\mathcal{Y}^*(\cdot)$ and $\mathcal{Z}^*(\cdot)$ in the point w_0, the statement follows immediately. □

Corollary 3.5. *For arbitrary fixed $x \in \mathcal{G}$ and $\varepsilon \to +0$ it follows that*

$$\operatorname{dist}(z(\varepsilon, x), \mathcal{Z}^*(x)) \to 0, \quad \operatorname{dist}(y(\varepsilon, x), \mathcal{Y}^*(x)) \to 0.$$

4. Differentiability of the smoothened function $f(\varepsilon, x)$ and estimates for $|f(\varepsilon, x) - f(x)|$

Recall that $\mathbf{U} = \mathbf{X} \times \mathbf{Y}$, $\mathcal{U} = \mathcal{G} \times \mathcal{Y}$.

Proposition 4.1. *For arbitrary $\varepsilon > 0$ the function φ in* (3.4) *is continuously differentiable in each point $u = (x, y) \in \mathcal{U}$ and its $\mathcal{U}$-gradient $\nabla_u^{\mathcal{U}} \varphi(\varepsilon, u)$ is continuous on the set $\mathcal{U}$ and has the form*

$$\nabla_u^{\mathcal{U}} \varphi(\varepsilon, u) = \nabla_u F(\varepsilon, x, y, z(\varepsilon, u)).$$

The proof follows immediately from Theorem 2.9 if we put, for fixed $\varepsilon > 0$,

$$\mathbf{V} := \mathbf{Z},\ \mathcal{V} := \mathcal{Z},\ m(u) := \min\{-F(\varepsilon, x, y, z) : z \in \mathcal{Z}\}$$

and take into account that $m(u) = -\max\{F(\varepsilon, x, y, z) : z \in \mathcal{Z}\} = -\varphi(\varepsilon, u)$.

Remark 4.2. If the gradient $\nabla_{\{x,y\}} F(x, y, z)$ is locally Lipschitz w.r.t. $\{x, y\}$ on the whole space $\mathbf{X} \times \mathbf{Y} \times \mathbf{Z}$, then Theorem 2.9 ensures that, for fixed $\varepsilon > 0$, the $\mathcal{U}$-gradient $\nabla_u^{\mathcal{U}} \varphi(\varepsilon, u)$ is locally Lipschitz on the set $\mathcal{U}$.

Proposition 4.3. *For arbitrary $\varepsilon > 0$ the function ψ in* (3.5) *is continuously differentiable in each point $(x, z) \in \mathbf{X} \times \mathbf{Z}$. If f is a generalized max-function and $f(\varepsilon, x)$ has the form* (3.7), *then the gradient of ψ w.r.t. $\{x, z\}$ has the form*

$$\nabla_{\{x,z\}} \psi(\varepsilon, x, z) = \{\nabla_x \psi(\varepsilon, x, z), \nabla_z \psi(\varepsilon, x, z)\}$$

with

$$\nabla_x \psi(\varepsilon, x, z) = \nabla_x h(x, z) - \varepsilon^{-1} \left(\sum_{s \in S_1} h_s^+(x, z) \nabla_x h_s(x, z) + \sum_{s \in S_2} h_s(x, z) \nabla_x h_s(x, z) \right),$$

$$\begin{aligned} \nabla_z \psi(\varepsilon, x, z) &= -2\varepsilon z + \nabla_z h(x, z) \\ &\quad -\varepsilon^{-1} \left(\sum_{s \in S_1} h_s^+(x, z) \nabla_z h_s(x, z) + \sum_{s \in S_2} h_s(x, z) \nabla_z h_s(x, z) \right). \end{aligned}$$

The proof follows immediately from formula (3.8).

Theorem 4.4. *For arbitrary $\varepsilon > 0$ the function f in (3.1) is continuously differentiable w.r.t. x on the set $\mathcal{G}$ and its gradient w.r.t. x has the form*

$$\nabla_x f(\varepsilon, x) = \nabla_x \varphi\left(\varepsilon, x, y(\varepsilon, x)\right) = \nabla_x F\left(\varepsilon, x, y(\varepsilon, x), z(\varepsilon, x, y(\varepsilon, x))\right). \tag{4.1}$$

Proof. With $\vartheta(u, v) = -F(\varepsilon, x, y, z) = -F(\varepsilon, u, v)$ we obtain from Theorem 2.9 that there exists the gradient

$$\nabla_u^{\mathcal{U}} \varphi(\varepsilon, u) = \nabla_u F(\varepsilon, u, z(\varepsilon, u)).$$

In particular,

$$\nabla_x^{\mathcal{U}} \varphi(\varepsilon, x, y) = \nabla_x F(\varepsilon, x, y, z(\varepsilon, x, y)) \tag{4.2}$$

holds true. Again, applying Theorem 2.9 to the function $\vartheta(u, v) = \varphi(\varepsilon, x, y)$ with $u = x,\ v = y,\ \mathcal{V} = \mathcal{Y}$, and $m(u) = f(\varepsilon, x)$, we get

$$\nabla_x f(\varepsilon, x) = \nabla_x \varphi(\varepsilon, x, y(\varepsilon, x)). \tag{4.3}$$

Thus, the inequalities (4.2) and (4.3) lead to (4.1).

Continuity of $\nabla_x f(\varepsilon, x)$ w.r.t. x on the set $\mathcal{G}$ follows from formula (4.1). Continuity of $\nabla_x^{\mathcal{U}} \varphi(\varepsilon, x, y)$ w.r.t. x on the set $\mathcal{U} = \mathcal{G} \times \mathcal{Y}$ can be concluded from the continuity of the operator $y(\varepsilon, x)$ on the set $\mathcal{G}$ and also from the continuity of the operator $z(\varepsilon, x, y)$ on the set $\mathcal{U}$. $\square$

Now, in case of a generalized max-function we obtain also another formula for $\nabla_x f(\varepsilon, x)$.

Theorem 4.5. *Let f be a generalized max-function and $f(\varepsilon, x)$ be of the form (3.7). Then, for arbitrary $\varepsilon > 0, x \in \mathcal{G}$, it holds*

$$\begin{aligned}\nabla_x f(\varepsilon, x) &= \nabla_x h(x, z(\varepsilon, x)) \\ &\quad -\varepsilon^{-1} \left(\sum_{s \in S_1} h_s^+(x, z(\varepsilon, x)) \nabla_x h_s(x, z(\varepsilon, x)) + \sum_{s \in S_2} h_s(x, z(\varepsilon, x)) \nabla_x h_s(x, z(\varepsilon, x)) \right).\end{aligned}$$

The proof follows from the expression of f in (3.10) and from Proposition (4.2).

Under the continuity, closedness and convexity assumptions and $inf_{y \in \mathcal{Y}} \varphi(x, y) = sup_{z \in \mathcal{Z}}$ $\psi(x, z)$ $(= f(x))$, with possible no existence of a saddle point for $F(x, \cdot, \cdot)$, we have: $\lim_{\varepsilon \to 0} f(\varepsilon, x) = f(x)$. This is more or less standard. In fact, for each $x \in \mathcal{G}, F(\epsilon, x, \cdot, \cdot)$ has a unique saddle-point on $\mathcal{Y} \times \mathcal{Z}$ and $f(\varepsilon, x)$ is equal to the saddle value. Existence results from convex analysis (see [8], Chap. VI, Prop. 2.2), and uniqueness results from the strict convexity-concavity. Writing the double inequality of saddle-point we deduce easily that

$$sup_{z \in \mathcal{Z}} \psi(x, z) \leq \liminf_{\varepsilon \to 0} f(\varepsilon, x) \leq \limsup_{\varepsilon \to 0} f(\varepsilon, x) \leq \inf_{y \in \mathcal{Y}} \varphi(x, y).$$

Now, let $\epsilon > 0$ be given and we look for estimates $|f(\varepsilon, x) - f(x)|$.

Theorem 4.6. *For arbitrary* $\varepsilon > 0$, $x \in \mathcal{G}$, $y^*(x) \in \mathcal{Y}^*(x)$, $z^*(x) \in \mathcal{Z}^*(x)$ *it holds*

$$f(\varepsilon, x) \le f(x) - \varepsilon||z(\varepsilon, x)||^2 + \varepsilon||y^*(x)||^2, \tag{4.4}$$

$$f(\varepsilon, x) \ge f(x) - \varepsilon||z^*(x)||^2 + \varepsilon||y(\varepsilon, x)||^2. \tag{4.5}$$

Proof. Inequality (4.4) follows from (3.6) and the relations

$$\begin{aligned} f(\varepsilon, x) &= \psi(\varepsilon, x, z(\varepsilon, x)) \le F(\varepsilon, x, y^*(x), z(\varepsilon, x)) \\ &= F(x, y^*(x), z(\varepsilon, x)) - \varepsilon||z(\varepsilon, x)||^2 + \varepsilon||y^*(x)||^2 \\ &\le \max_{z \in \mathcal{Z}} F(x, y^*(x), z) - \varepsilon||z(\varepsilon, x)||^2 + \varepsilon||y^*(x)||^2. \end{aligned}$$

Due to $y^*(x) \in \mathcal{Y}^*(x)$, we get from (3.3)

$$\max_{z \in \mathcal{Z}} F(x, y^*(x), z) = \varphi(x, y^*(x)) = f(x).$$

Analogously, inequality (4.5) follows from (3.6) and the relations

$$\begin{aligned} f(\varepsilon, x) &= \varphi(\varepsilon, x, y(\varepsilon, x)) \ge F(\varepsilon, x, y(\varepsilon, x), z^*(x)) \\ &= F(x, y(\varepsilon, x), z^*(x)) - \varepsilon||z^*(x)||^2 + \varepsilon||y(\varepsilon, x)||^2 \\ &\ge \min_{y \in \mathcal{Y}} F(x, y, z^*(x)) - \varepsilon||z^*(x)||^2 + \varepsilon||y(\varepsilon, x)||^2. \end{aligned}$$

With regard to $z^*(x) \in \mathcal{Z}^*(x)$,

$$\min_{y \in \mathcal{Y}} F(x, y, z^*(x)) = \psi(x, z^*(x)) = f(x).$$

□

Recall that, according to (2.7), $\hat{z}(x)$, $\hat{y}(x)$ denote the norm-minimal points of the sets $\mathcal{Z}^*(x)$, $\mathcal{Y}^*(x)$, respectively.

Corollary 4.7. *For arbitrary* $\varepsilon > 0$, $x \in \mathcal{G}$ *the following two-sided estimate is true:*

$$f(x) - \varepsilon||\hat{z}(x)||^2 \le f(\varepsilon, x) \le f(x) + \varepsilon||\hat{y}(x)||^2. \tag{4.6}$$

To prove this result it is sufficient to choose in the inequalities (4.4), (4.5):

$$\begin{aligned} y^*(x) = \hat{y}(x) = \arg\min \left\{ ||y||^2 : y \in \mathcal{Y}^*(x) \right\}, \\ z^*(x) = \hat{z}(x) = \arg\min \left\{ ||z||^2 : z \in \mathcal{Z}^*(x) \right\}. \end{aligned}$$

Corollary 4.8. *For arbitrary* $\varepsilon > 0, x \in \mathcal{G}$ *the following estimate is true:*

$$||z(\varepsilon, x)||^2 + ||y(\varepsilon, x)||^2 \le ||\hat{z}(x)||^2 + ||\hat{y}(x)||^2. \tag{4.7}$$

5. Convergence of the vector functions $z(\varepsilon, x)$ and $y(\varepsilon, x)$ for $\varepsilon \to +0$

First we consider, for fixed $x \in \mathcal{G}$, the behavior of the points $z(\varepsilon, x),\ y(\varepsilon, x)$ for $\varepsilon \to +0$.

Theorem 5.1. *For fixed $x \in \mathcal{G}$, $\varepsilon \to 0$ it holds*

$$z(\varepsilon, x) \to \hat{z}(x),\ \ y(\varepsilon, x) \to \hat{y}(x),$$

where, according to (2.7), $\hat{z}(x)$, $\hat{y}(x)$ are the normal points of the convex, compact sets $\mathcal{Z}^*(x)$ and $\mathcal{Y}^*(x)$, respectively.

Proof. In view of Corollary 3.5 for $\varepsilon \to +0$ the relations

$$\operatorname{dist}(z(\varepsilon, x), \mathcal{Z}^*(x)) \to 0,\ \operatorname{dist}(y(\varepsilon, x), \mathcal{Y}^*(x)) \to 0$$

are true. Let $\tilde{z}(\varepsilon, x),\ \tilde{y}(\varepsilon, x)$ are the projections of the points $z(\varepsilon, x),\ y(\varepsilon, x)$ onto the sets $\mathcal{Z}^*(x)$ and $\mathcal{Y}^*(x)$, respectively. Then, we have

$$||z(\varepsilon, x) - \tilde{z}(\varepsilon, x)|| \to 0, \quad ||y(\varepsilon, x) - \tilde{y}(\varepsilon, x)|| \to 0 \text{ for } \varepsilon \to +0. \tag{5.1}$$

Hence,

$$||z(\varepsilon, x)||^2 + ||y(\varepsilon, x)||^2 = ||\tilde{z}(\varepsilon, x)||^2 + ||\tilde{y}(\varepsilon, x)||^2 + \xi(\varepsilon, x), \tag{5.2}$$

with $\xi(\varepsilon, x) \to 0$ for $\varepsilon \to +0$ and fixed $x \in \mathcal{G}$. With regard to (5.2) and (4.5), we obtain for each $\varepsilon \to +0,\ x \in \mathcal{G}$ that

$$||\tilde{z}(\varepsilon, x)||^2 + ||\tilde{y}(\varepsilon, x)||^2 + \xi(\varepsilon, x) \le ||\hat{z}(x)||^2 + ||\hat{y}(x)||^2. \tag{5.3}$$

Summing up the inequalities (2.8) and (2.9) with $z = \tilde{z}(\varepsilon, x),\ y = \tilde{y}(\varepsilon, x)$, we get

$$||\tilde{z}(\varepsilon, x)||^2 + ||\tilde{y}(\varepsilon, x)||^2 \ge ||\hat{z}(x)||^2 + ||\hat{y}(x)||^2 + ||\tilde{z}(\varepsilon, x) - \hat{z}(x)||^2 + ||\tilde{y}(\varepsilon, x) - \hat{y}(x)||^2. \tag{5.4}$$

For fixed $x \in \mathcal{G}$, (5.3) and (5.4) lead to

$$||\tilde{z}(\varepsilon, x) - \hat{z}(x)||^2 + ||\tilde{y}(\varepsilon, x) - \hat{y}(x)||^2 \le -\xi(\varepsilon, x) \to 0 \text{ for } \varepsilon \to +0. \tag{5.5}$$

Now, to finish the proof, we have to use the relations (5.1) and (5.5). □

In order to solve optimization problems with functions of the type (1.1) by means of a smoothing procedure, sometimes it is important to know under which assumptions the following convergence happens:

$$z(\varepsilon_k, x_k) \to \hat{z}(x), \quad y(\varepsilon_k, x_k) \to \hat{y}(x) \text{ for } \varepsilon_k \to +0, x_k \in \mathcal{G}, x_k \to x. \tag{5.6}$$

If that would be true, then uniform convergence of

$$z(\varepsilon, x) \to \hat{z}(x), \quad y(\varepsilon, x) \to \hat{y}(x) \text{ for } \varepsilon \to 0 \text{ on each compact set } \mathcal{G}$$

would be an important conclusion of (5.6). Under additional assumptions on the point-to-set-mappings $\mathcal{Z}^*(x)$ and and $\mathcal{Y}^*(x)$, relation (5.6) turns out to be true. To prove that, we have to strengthen the statements in the Theorems 3.4 and 5.1.

Let $x \in \mathcal{G}$. The normal points $\hat{z}(x)$, $\hat{y}(x)$ are called *stable points* of the point-to-set-mappings $\mathcal{Z}^*(x),\ \mathcal{Y}^*(x)$, respectively, if for each sequence $\{x_k\} \in \mathcal{G},\ x_k \to x$, there exist sequences $\{z_k\}$ and $\{y_k\}$ such that

$$z_k \in \mathcal{Z}^*(x_k),\ z_k \to \hat{z}(x) \ \text{ and } \ y_k \in \mathcal{Y}^*(x_k),\ y_k \to \hat{y}(x).$$

Theorem 5.2. *Assume that $x \in \mathcal{G}$ and that the points $\hat{z}(x)$, $\hat{y}(x)$ are stable points of the sets $\mathcal{Z}^*(x)$, $\mathcal{Y}^*(x)$, respectively, w.r.t. the sequence $\{x_k\} \in \mathcal{G}, x_k \to x$. Then, if $x_k \in \mathcal{G}$, $x_k \to x$, it holds:*

$$z(\varepsilon_k, x_k) \to \hat{z}(x),\ y(\varepsilon_k, x_k) \to \hat{y}(x) \text{ for } \varepsilon_k \to +0.$$

Proof. The statement can be proved by means of the same scheme as in the proof of Theorem 5.1. Due to Theorem 3.4, we have

$$\operatorname{dist}(z(\varepsilon_k, x_k), \mathcal{Z}^*(x)) \to 0, \quad \operatorname{dist}(y(\varepsilon_k, x_k), \mathcal{Y}^*(x)) \to 0.$$

If $\tilde{z}(\varepsilon_k, x_k)$ and $\tilde{y}(\varepsilon_k, x_k)$ are projections of $z(\varepsilon_k, x_k)$ and $y(\varepsilon_k, x_k)$ onto the convex, compact sets $\mathcal{Z}^*(x)$, $\mathcal{Y}^*(x)$, respectively, then

$$||z(\varepsilon_k, x_k) - \tilde{z}(\varepsilon_k, x_k)|| \to 0, \quad ||y(\varepsilon_k, x_k) - \tilde{y}(\varepsilon_k, x_k)|| \to 0. \tag{5.7}$$

Hence,

$$||z(\varepsilon_k, x_k)||^2 + ||y(\varepsilon_k, x_k)||^2 = ||\tilde{z}(\varepsilon_k, x_k)||^2 + ||\tilde{y}(\varepsilon_k, x_k)||^2 + \xi_k, \text{ with } \xi_k \to 0. \tag{5.8}$$

From relation (4.5) we obtain that

$$||z(\varepsilon_k, x_k)||^2 + ||y(\varepsilon_k, x_k)||^2 \le ||\hat{z}(x_k)||^2 + \hat{y}(x_k)||^2. \tag{5.9}$$

Due to the assumption made there exist sequences $\{z_k\}, \{y_k\}$ such that

$$z_k \in \mathcal{Z}^*(x_k),\ z_k \to \hat{z}(x) \text{ and } y_k \in \mathcal{Y}^*(x_k),\ y_k \to \hat{y}(x).$$

In view of

$$||\hat{z}(x_k)||^2 = \min\{||z||^2 : z \in Z^*(x_k)\} \le ||z_k||^2 \to ||\hat{z}(x)||^2$$

and

$$||\hat{y}(x_k)||^2 = \min\left\{||y||^2 : y \in \mathcal{Y}^*(x_k)\right\} \le ||y_k||^2 \to ||\hat{y}(x)||^2,$$

one can conclude that

$$||\hat{z}(x_k)||^2 + ||\hat{y}(x_k)||^2 \le ||\hat{z}(x)||^2 + ||\hat{y}(x)||^2 + \eta_k, \text{ with } \eta_k \to 0. \tag{5.10}$$

The relations (5.8)-(5.10) lead to

$$||\tilde{z}(\varepsilon_k, x_k)||^2 + ||\tilde{y}(\varepsilon_k, x_k)||^2 \le ||\hat{z}(x)||^2 + ||\hat{y}(x)||^2 + (\eta_k - \xi_k). \tag{5.11}$$

Summing up the inequalities (2.8), (2.9) with $z := \tilde{z}(\varepsilon_k, x_k) \in \mathcal{Z}^*(x)$, $y := \tilde{y}(\varepsilon_k, x_k) \in \mathcal{Y}^*(x)$, we get

$$\begin{aligned} &||\tilde{z}(\varepsilon_k, x_k)||^2 + ||\tilde{y}(\varepsilon_k, x_k)||^2 \\ &\quad \ge ||\hat{z}(x)||^2 + ||\hat{y}(x)||^2 + ||\tilde{z}(\varepsilon_k, x_k) - \hat{z}(x)||^2 + ||\tilde{y}(\varepsilon_k, x_k) - \hat{y}(x)||^2. \end{aligned} \tag{5.12}$$

Now, the inequalities (5.11), (5.12) ensure that

$$||\tilde{z}(\varepsilon_k, x_k) - \hat{z}(x)||^2 + ||\tilde{y}(\varepsilon_k, x_k) - \hat{y}(x)||^2 \le (\eta_k - \xi_k) \to 0 \tag{5.13}$$

and the statement follows from (5.7), (5.13). □

Corollary 5.3. *Let $\mathcal{H} \subset \mathcal{G}$ be a compact set. Assume that for each $x \in \mathcal{H}$, according to the sequence $\{x_k\} \in \mathcal{H}$, $x_k \to x$, the vectors $\hat{z}(x)$, $\hat{y}(x)$ are stable points of the point-to-set-mappings $\mathcal{Z}^*(\cdot)$ and $\mathcal{Y}^*(\cdot)$, respectively. Then, for $\varepsilon \to +0$, uniform convergence (w.r.t. $x \in \mathcal{H}$) of*

$$z(\varepsilon, x) \to \hat{z}(x),\ y(\varepsilon, x) \to \hat{y}(x)$$

is true.

6. Estimates of the extremal values of parametric minimax-functions and their regularizations

Proposition 6.1. *Let $\mathcal{H} \subset \mathcal{G}$ be a compact set. Then the relations*

$$\hat{p}(\mathcal{H}) := \sup\left\{||\hat{z}(x)||^2 : x \in \mathcal{H}\right\} < +\infty, \quad \hat{q}(\mathcal{H}) := \sup\left\{||\hat{y}(x)||^2 : x \in \mathcal{H}\right\} < +\infty$$

are true.

Proof. From the upper semi-continuity of the point-to-set-mappings $\mathcal{Z}^*(\cdot)$ and $\mathcal{Y}^*(\cdot)$ on $\mathcal{G}$ it follows that the functions

$$p(x) := \max\{||z||^2 : z \in \mathcal{Z}^*(x)\}, \quad q(x) := \max\{||y||^2 : y \in \mathcal{Y}^*(x)\}$$

are upper semi-continuous on the set $\mathcal{G}$. Indeed, let $x_k,\ x \in \mathcal{G},\ x_k \to x$ and

$$z_k = \arg\max\left\{||z||^2 : z \in \mathcal{Z}^*(x_k)\right\}, \quad y_k = \arg\max\left\{||y||^2 : y \in \mathcal{Y}^*(x_k)\right\}.$$

In view of the compactness of the sets $\mathcal{Z}^*(\cdot)$ and $\mathcal{Y}^*(\cdot)$ in the spaces $\mathbf{Z}$ and $\mathbf{Y}$, respectively, these points $z_k,\ y_k$ exist. Due to the upper semi-continuity of the maps $\mathcal{Z}^*(\cdot)$ and $\mathcal{Y}^*(\cdot)$ on $\mathcal{G}$, for arbitrary $\rho > 0$, starting with some number $k = k(\rho)$, the points $z_k,\ y_k$ are included in some ρ-neighborhood of the compact sets $\mathcal{Z}^*(x)$ and $\mathcal{Y}^*(x)$, respectively. Hence, taking into account that $\rho > 0$ can be chosen arbitrarily small, we obtain

$$\overline{\lim}_{k\to\infty} p\,(x_k) = \overline{\lim}_{k\to\infty}||z_k||^2 \le p(x), \quad \overline{\lim}_{k\to\infty} q\,(x_k) = \overline{\lim}_{k\to\infty}||y_k||^2 \le q(x).$$

From the inequalities

$$||\hat{z}(x)||^2 \le p(x), \quad ||\hat{y}(x)||^2 \le q(x)$$

and the boundedness from above of an upper semi-continuous function on an arbitrary compact set it follows that

$$\hat{p}(\mathcal{H}) < +\infty, \quad \hat{q}(\mathcal{H}) < +\infty.$$

□

Proposition 6.2. *Let $\mathcal{H} \subset \mathcal{G}$ be a compact set. Then the two-sided estimate*

$$f(x) - \hat{p}(\mathcal{H})\varepsilon \le f(\varepsilon, x) \le f(x) + \hat{q}(\mathcal{H})\varepsilon \quad \forall x \in \mathcal{H}, \varepsilon > 0$$

is true.

The proof immediately follows from Corollary 4.7 and Proposition 6.1.

Corollary 6.3. *For arbitrary compact set $\mathcal{H} \subset \mathcal{G}$ and arbitrary $\varepsilon > 0$ the following estimate holds:*

$$\sup_{x\in\mathcal{H}} |f(\varepsilon, x) - f(x)| \le \varepsilon \max\{\hat{p}(\mathcal{H}), \hat{q}(\mathcal{H})\},$$

i.e., on arbitrary compact set $\mathcal{H} \subset \mathcal{G}$, for $\varepsilon \to +0$, the estimate

$$|f(\varepsilon, x) - f(x)| \le O(\varepsilon)$$

holds uniformly (w.r.t. $x \in \mathcal{H}$).

Now, we consider Problem (1.1). Denote

$$f_* = \inf \{f(x) : x \in Q\}$$

and assume that $f_* > -\infty$.

In the sequel we give a two-sided estimate for the exactness of the approximate solution of Problem (1.1), calculated by means of the corresponding regularization of the function f.

Theorem 6.4. *Assume that for some $\bar{x} \in Q$ the set $\mathcal{H} = \{x \in Q : f(x) \leq f(\bar{x})\}$ is bounded in $\mathbf{X}$. Then the estimates*

$$\min \{f(\varepsilon, x) : x \in \mathcal{H}\} - \hat{q}(\mathcal{H})\varepsilon \leq f_* \leq \min \{f(\varepsilon, x) : x \in \mathcal{H}\} + \hat{p}(\mathcal{H})\varepsilon \tag{6.1}$$

are true.

Proof. Note that the closeness of $\mathcal{H}$ follows from the closeness of Q and the continuity of f on $\mathcal{G}$, hence, $\mathcal{H}$ is compact. Obviously, $f_* = \min \{f(x) : x \in \mathcal{H}\}$. Due to Proposition 6.2, we obtain that

$$f_* - \hat{p}(\mathcal{H})\varepsilon \leq \min \{f(\varepsilon, x) : x \in \mathcal{H}\} \leq f_* + \hat{q}(\mathcal{H})\varepsilon,$$

proving inequality (6.1). □

Now, let $\bar{x} \in Q$, $\mathcal{H} = \{x \in Q : f(x) \leq f(\bar{x})\}$, $\varepsilon > 0$, $\delta \geq 0$. Finally, we consider the regularized problem

$$f_*(\varepsilon) = \min \{f(\varepsilon, x) : x \in \mathcal{H}\} \tag{6.2}$$

and denote by

$$\mathcal{M}_\delta(\varepsilon) := \{x \in \mathcal{H} : f(\varepsilon, x) \leq f_*(\varepsilon) + \delta\}$$

the δ-optimal set of the regularized function f.

If there is known some point of $\mathcal{M}_\delta(\varepsilon)$, then the right-hand side of (6.1) can be improved.

Theorem 6.5. *For some $\bar{x} \in Q$ let the set $\mathcal{H} = \{x \in Q : f(x) \leq f(\bar{x})\}$ be bounded in $\mathbf{X}$ and for some $\varepsilon > 0$, $\delta > 0$ let $x_\delta(\varepsilon) \in \mathcal{M}_\delta(\varepsilon)$. Then*

$$f(\varepsilon, x_\delta(\varepsilon)) - \delta - \hat{q}(\mathcal{H})\varepsilon \leq f_* \leq f\,(\varepsilon, x_\delta(\varepsilon)) + \varepsilon||\hat{z}(x_\delta(\varepsilon))||^2 - \varepsilon||y(\varepsilon, x_\delta(\varepsilon))||^2 \tag{6.3}$$

is true.

Proof. The left-hand side of estimate (6.3) immediately follows from the left-hand side of inequality (6.1). From (4.3) and Theorem 5.1 we get

$$\begin{aligned} f(\varepsilon, x_\delta(\varepsilon)) &\geq f(x_\delta(\varepsilon)) - \varepsilon||\hat{z}(x_\delta(\varepsilon))||^2 + \varepsilon||y(\varepsilon, x_\delta(\varepsilon))||^2 \\ &\geq \min\{f(x) : x \in \mathcal{H}\} - \varepsilon||\hat{z}(x_\delta(\varepsilon))||^2 + \varepsilon||y(\varepsilon, x_\delta(\varepsilon))||^2 \\ &= f_* - \varepsilon||\hat{z}(x_\delta(\varepsilon))||^2 + \varepsilon||y(\varepsilon, x_\delta(\varepsilon))||^2. \end{aligned}$$

□

Acknowledgements. We thank the referee for valuable comments and suggestions.

References

[1] K. Beer: Solution of Large Scale Linear Optimization Problems, Deutscher Verlag der Wissenschaften, Berlin, 1977.

[2] K. Beer: Decomposition approach in optimization, in: K. H. Elster (ed.): Modern Mathematical Methods of Optimization, Akademie Verlag, Berlin (1993) 116–131.

[3] B. S. Darkhovsky, E. S. Levitin: On differential properties of the value function of a finite dimensional parametric min-max-problem, Kibernetika, Kiev, Part. I: 4 (1976) 58–69, Part II: 5 (1976) 115–123. (in Russian)

[4] V. F. Demyanov: Minimax: Directional Differentiation, University of Leningrad, Leningrad, 1975. (in Russian)

[5] V. F. Demyanov, A. M. Rubinov: Quasidifferential Calculus, Optimization Software, Inc. Publ. Division, New York, 1986.

[6] V. F. Demyanov, I. S. Zabrodin: Directional differentiability of a continuous maximum function of quasidifferentiable functions, Math. Prog. Study 29 (1986) 108–117.

[7] V. F. Demyanov, D. Pallaschke (eds.): Nondifferential Optimization Methods and Applications, Springer-Verlag, Berlin et al., 1987.

[8] I. Ekeland, R. Temam: Convex Analysis and Variational Problems, North-Holland, Elsevier, 1976.

[9] Y. B. Germeyer: Games with Non-antagonistic Interests, Nauka, Moscow, 1976. (in Russian)

[10] E. G. Gol'stein: Konvexe Optimierung, WTG, Bd. 122, Akademie Verlag, Berlin, 1973.

[11] E. G. Gol'stein, N. V. Tretyakov: Modified Lagrange Functions, J. Wiley & Sons, New York, 1995.

[12] R. Hettich, G. Still: Second order optimality conditions for generalized semi-infinite programming problems, Optimization 34 (1995) 195–211.

[13] H. Th. Jongen, J. Rueckmann, O. Stein: Generalized semi-infinite optimization: A first order optimality condition and examples, No 95-24, Dept. of Mathematics, University of Trier, Trier, 1995. (Preprint)

[14] A. Kaplan, R. Tichatschke: Stable Methods for Ill-posed Variational Problems, Akademie Verlag, Berlin, 1994.

[15] A. Kaplan, R. Tichatschke: On a class of terminal variational problems, Proceedings of the PARAOPT-Conference, Enschede, June 1995, in: Parametric Optimization and Related Topics IV, (eds. J. Guddat, H. Th. Jongen, F. Nozicka, G. Still and F. Twilt), Verlag Peter Lang, Frankfurt/M. (1996) 185–199.

[16] A. Kaplan, R. Tichatschke: On the numerical treatment of a class of semi-infinite terminal problems, Optimization 41 (1997) 1–36.

[17] S. Karlin: Mathematical Methods and Theory in Games, Programming and Economics, Pergamon Press, London - Paris, 1959.

[18] L. S. Lasdon: Optimization Theory for Large Systems, Macmillan Comp., London - New York, 1975.

[19] E. S. Levitin, B. T. Polyak: Convergence of minimizing sequences in conditional extremum problems, Soviet Math. Doklady 7 (1966) 764–767.

[20] E. S. Levitin: Second-order conditions in nonsmooth mathematical programming problems with applications in minimax problems with binding variables, Soviet Math. Doklady 20 (1979) 59–63.

[21] E. S. Levitin: Decomposition choice of the decent direction in a block programming problem with binding variables and constraints, in: Models and Methods of Optimization, (ed. E.R. Smolyakov) issue 7, pp. 43-54, All-Union Institute for System Research (VNIISI); Moscow 1990. (in Russian)

[22] E. S. Levitin: Perturbation Theory in Mathematical Programming and its Applications, J. Wiley & Sons, Chichester et al., 1994.

[23] E. S. Levitin: Differential properties of parametric minimum functions and extremal maps, Proceedings of the PARAOPT-Conference, Enschede, June 1995, in: Parametric Optimization and Related Topics IV, (eds. J. Guddat, H. Th. Jongen, F. Nozicka, G. Still and F. Twilt), Verlag Peter Lang, Frankfurt/M. (1996) 213–226.

[24] E. S. Levitin, R. Tichatschke: A branch-and-bound approach for solving a class of generalized semi-infinite programming problems, Journal of Global Optimization 13 (1998) 3.

[25] B. Luderer: The quasidifferential of an optimal value function in non-smooth programming, Optimization 20 (1989) 597–613.

[26] B. Luderer: Directional derivative estimates for the optimal value function of the quasidifferentiable programming problem, Math. Programming 51 (1991) 333–348.

[27] B. T. Polyak: Introduction to Optimization, Optimization Software, Inc. Publ. Division, New York, 1987.

[28] A. Shapiro: Second-order derivatives of extremal value functions and optimality conditions for semi-infinite programs, Mathematics of Operation Research 10 (1985) 207–219.

[29] A. Shapiro: Sensitivity analysis of parametrized programs via generalized equations, SIAM J. Control and Optimization 32 (1994) 553–571.

[30] V. D. Skarin: Regularization of minimax problems appearing in convex programming, Journal Comp. Maths. Math. Phys. 17 (1977). (in Russian)

[31] A. N. Tikhonov, V. J. Arsenin: Methods for Solving Ill-posed Problems, J. Wiley & Sons, New York, 1977.