

## STABILITY OF EXACT PENALTY FOR CLASSES OF CONSTRAINED MINIMIZATION PROBLEMS IN BANACH SPACES

Alexander J. Zaslavski

Dedicated to Professor Wataru Takahashi on the occasion of his 65th birthday

**Abstract.** In this paper we use the penalty approach in order to study two constrained minimization problems in Banach spaces. A penalty function is said to have the generalized exact penalty property if there is a penalty coefficient for which approximate solutions of the unconstrained penalized problem are close enough to approximate solutions of the corresponding constrained problem. In this paper we show that the generalized exact penalty property holds and is stable under perturbations of objective functions, constraint functions and the right-hand side of constraints.

### 1. INTRODUCTION

Penalty methods are an important and useful tool in constrained optimization. See, for example, [2-5, 7-10, 12, 13, 15, 17-19] and the references mentioned there.

The notion of exact penalization was introduced by Eremin [7] and Zangwill [15] for use in the development of algorithms for nonlinear constrained optimization. Since that time exact penalty functions have continued to play a key role in the theory of mathematical programming. For discussions and various applications of exact penalization to various constrained optimization problems see [2, 3, 5, 9, 11].

In this paper we use the penalty approach in order to study constrained minimization problems with locally Lipschitzian constraints in Banach spaces. A penalty function is said to have the exact penalty property [2, 3, 5, 9] if there is a penalty coefficient for which a solution of an unconstrained penalized problem is a solution of the corresponding constrained problem. We study two constrained nonconvex minimization problems with Lipschitzian (on bounded sets) objective functions. The first problem is an equality-constrained problem in a Banach space with a locally

---

Received April 11, 2008.

2000 *Mathematics Subject Classification*: 49M30, 90C26, 90C30.

*Key words and phrases*: Clarke's generalized gradient, Ekeland's variational principle, Minimization problem, Penalty function.

Lipschitzian constraint function and the second problem is an inequality-constrained problem in a Banach space with a locally Lipschitzian constraint function.

In [17] we considered these two problems and established a very simple sufficient condition for the exact penalty property. In particular, the problem  $f(x) \rightarrow \min$  subject to  $g(x) = c$  possesses the exact penalty if the real number  $c$  is not a critical value of the function  $g$ . In other words the set  $g^{-1}(c)$  does not contain a critical point of the function  $g$ . Usually the exact penalty property is related to calmness of the perturbed constraint function. In [17] and here we use the assumption of the different nature which is not difficult to verify.

Note that in [17] we used the following notion of a critical point of a Lipschitzian function introduced in [16]:

A point  $z$  is a critical point of the function  $g$  if  $0 \in \partial g(z)$  where  $\partial g(z)$  is Clarke's generalized gradient of  $g$  at  $z$  [4].

More precisely, in [17] we considered the following constrained minimization problems:

$$(P_{c,e}^{(f,g)}) \quad f(x) \rightarrow \min \text{ subject to } x \in g^{-1}(c)$$

and

$$(P_{c,i}^{(f,g)}) \quad f(x) \rightarrow \min \text{ subject to } x \in g^{-1}((-\infty, c]),$$

where  $g$  is a locally Lipschitzian function defined on a Banach space  $X$ ,  $f : X \rightarrow R^1$  is a function which is Lipschitzian on all bounded subsets of  $X$  and which satisfies a growth condition common in the literature and  $c$  is a real number such that the set  $g^{-1}(c)$  is nonempty.

We associate with these two problems the corresponding families of unconstrained minimization problems

$$(P_{\lambda,c,e}^{(f,g)}) \quad f(x) + \lambda |g(x) - c| \rightarrow \min, \quad x \in X$$

and

$$(P_{\lambda,c,i}^{(f,g)}) \quad f(x) + \lambda \max\{g(x) - c, 0\} \rightarrow \min, \quad x \in X,$$

where  $\lambda > 0$ .

The main result of [17, Theorem 1.1] implies that if the space  $X$  is finite-dimensional,  $c$  is not a critical value of  $g$  and if  $\lambda$  is sufficiently large, then any solution of problem  $(P_{\lambda,c,e}^{(f,g)})$  is a solution of problem  $(P_{c,e}^{(f,g)})$  and any solution of problem  $(P_{\lambda,c,i}^{(f,g)})$  is a solution of problem  $(P_{c,i}^{(f,g)})$ . Thus if the space  $X$  is finite-dimensional, then problems  $(P_{c,e}^{(f,g)})$  and  $(P_{c,i}^{(f,g)})$  possess the classical exact penalty property.

If the space  $X$  is infinite-dimensional, then the existence of solutions of problems  $(P_{\lambda,c,e}^{(f,g)})$  and  $(P_{\lambda,c,i}^{(f,g)})$  is not guaranteed and in this situation we consider approximate solutions which always exist.

Consider a minimization problem  $h(z) \rightarrow \min, z \in X$  where  $h : X \rightarrow \mathbb{R}^1$  is a bounded from below function. We say that  $x \in X$  is a  $\delta$ -approximate solution of the problem  $h(z) \rightarrow \min, z \in X$ , where  $\delta > 0$ , if  $h(x) \leq \inf\{h(z) : z \in X\} + \delta$ .

Since in [17-19] and in this paper we consider minimization problems in a general Banach space the existence of their solutions is not guaranteed. Therefore we are interested in approximate solutions of the unconstrained penalized problems and in approximate solutions of the corresponding constrained problems. Theorem 1.1 of [17] establishes the existence of a penalty coefficient for which approximate solutions of the unconstrained penalized problem are close enough to approximate solutions of the corresponding constrained problem. This is a novel approach in the penalty type methods.

More precisely, if  $c$  is not a critical value of  $g$ , then Theorem 1.1 of [17] establishes the following property which will be called here as the generalized exact penalty property:

For each  $\epsilon > 0$  there exists  $\delta(\epsilon) > 0$ , which depends only on  $\epsilon$ , such that if  $\lambda \geq \bar{\lambda}$  and  $x$  is a  $\delta$ -approximate solution of  $(P_{\lambda,c,e}^{(f,g)})$  ( $(P_{\lambda,c,i}^{(f,g)})$  respectively), then there exists a  $(\bar{\lambda}\epsilon)$ -approximate solution of  $(P_{c,e}^{(f,g)})$  ( $(P_{c,i}^{(f,g)})$  respectively) such that  $\|y - x\| \leq \epsilon$ .

Here  $\bar{\lambda}$  is a positive constant which does not depend on  $\epsilon$ .

It is not difficult to see that the generalized exact penalty property implies that any exact solution of the unconstrained penalized problem whose penalty coefficient is larger than  $\bar{\lambda}$ , is an exact solution of the corresponding constrained problem. It should be mentioned that even if the space  $X$  is finite-dimensional the generalized exact penalty property is of interest. Since computational algorithms in general produce only approximate solutions of minimization problems it is important to know that approximate solutions of unconstrained penalized problems with large penalty coefficients are close enough to approximate solutions of the corresponding constrained problem.

In [18] we study the stability of the generalized exact penalty property under perturbations of the functions  $f$  and  $g$  and of the parameter  $c$ . The stability of the generalized exact penalty property is crucial in practice. One reason is that in practice we deal with a problem which consists a perturbation of the problem we wish to consider. Another reason is that the computations introduce numerical errors. In [18] we show that the generalized exact penalty property is stable if  $c$  is not a critical value of  $g$ . In the present paper our goal is to show that the generalized exact penalty property is stable under assumptions which are essentially weaker than the assumptions of [18]. In particular, we show that if the equality constrained problem

possesses a solution, then the assumption of [18] can be substituted by the following assumption:

The set of solutions of the problem does not contain a critical point of  $g$ .

The paper is organized as follows. The main results are stated in Section 2. Section 3 contains auxiliary results. The main results are proved in Section 4.

## 2. THE MAIN RESULTS

Let  $(X, \|\cdot\|)$  be a Banach space,  $(X^*, \|\cdot\|_*)$  its dual space and let  $f : X \rightarrow R^1$  be a locally Lipschitzian function.

For each  $x \in X$  let

$$f^0(x, h) = \limsup_{t \rightarrow 0+, y \rightarrow x} [f(y + th) - f(y)]/t, \quad h \in X,$$

be the Clarke generalized directional derivative of  $f$  at the point  $x$  [4], let

$$\partial f(x) = \{l \in X^* : f^0(x, h) \geq l(h) \text{ for all } h \in X\}$$

be Clarke's generalized gradient of  $f$  at  $x$ , [4] and set

$$(2.1) \quad \Xi_f(x) = \inf\{f^0(x, h) : h \in X \text{ and } \|h\| \leq 1\}$$

[16].

A point  $x \in X$  is called a critical point of  $f$  if  $0 \in \partial f(x)$  [16]. It is not difficult to see that  $x \in X$  is a critical point of  $f$  if and only if  $\Xi_f(x) = 0$ .

A real number  $c \in R^1$  is called a critical value of  $f$  if there is a critical point  $x$  of  $f$  such that  $f(x) = c$ .

It is known [4, Chapter 2, Section 2.3] that  $\partial(-f)(x) = -\partial f(x)$  for any  $x \in X$ . This equality implies that  $x \in X$  is a critical point of  $f$  if and only if  $x$  is a critical point of  $-f$  and  $c \in R^1$  is a critical value of  $f$  if and only if  $-c$  is a critical value of  $-f$ .

For each function  $h : X \rightarrow R^1$  set  $\inf(h) = \inf\{h(z) : z \in X\}$ .

For each  $x \in X$  and each nonempty set  $B \subset X$  put

$$d(x, B) = \inf\{\|x - y\| : y \in B\}.$$

We say that a locally Lipschitz function  $f : X \rightarrow R^1$  satisfies Palais-Smale (P-S) condition on a set  $A \subset X$  if for any sequence  $\{x_i\}_{i=1}^\infty \subset A$  for which the sequence  $\{f(x_i)\}_{i=1}^\infty$  is bounded and  $\lim_{i \rightarrow \infty} \Xi_f(x_i) = 0$  there exists a norm convergent subsequence in  $X$  [1, 16, 17].

*Remark 2.1.* In [17] instead of the function  $\Xi_f$  we introduced a function  $\tilde{\Xi}_f : X \rightarrow R^1$  defined by

$$\tilde{\Xi}_f(x) = \inf\{f^0(x, h) : h \in X \text{ and } \|h\| = 1\}, \quad x \in X.$$

Clearly for all  $x \in X$  we have  $\Xi_f(x) \leq \tilde{\Xi}_f(x)$  and  $\tilde{\Xi}_f(x) \geq 0$  if and only if  $\Xi_f(x) = 0$ . It is not difficult to see that for each sequence  $\{x_i\}_{i=1}^\infty \subset X$

$$\lim_{i \rightarrow \infty} \Xi_f(x_i) = 0 \text{ if and only if } \liminf_{i \rightarrow \infty} \tilde{\Xi}_f(x_i) \geq 0.$$

For each  $x \in X$  and each  $r > 0$  set

$$B(x, r) = \{z \in X : \|x - z\| \leq r\}, \quad B^0(x, r) = \{z \in X : \|x - z\| < r\}.$$

We assume that the infimum over an empty set is infinity.

For each  $c \in R^1$  and each pair of functions  $f, g : X \rightarrow R^1$  set

$$(2.2) \quad \inf(f, c, g) = \inf\{f(z) : z \in g^{-1}(c)\},$$

$$\inf(f, (-\infty, c], g) = \inf\{f(z) : z \in g^{-1}((-\infty, c])\}.$$

Let  $\phi : [0, \infty) \rightarrow R^1$  be an increasing function such that

$$(2.3) \quad \lim_{t \rightarrow \infty} \phi(t) = \infty.$$

Let  $M_0, M_1, M_2$  be positive numbers and let  $h : X \rightarrow R^1$  be a continuous function. Now we define a family of functions which are close to the function  $h$ . This family is determined by the parameters  $M_0, M_1, M_2$ .

Denote by  $\mathcal{U}(h, M_0, M_1, M_2)$  the set of all continuous functions  $g : X \rightarrow R^1$  such that

$$(2.4) \quad |g(x) - h(x)| \leq M_1 \text{ for all } x \in B(0, M_0),$$

$$(2.5) \quad |h(x) - g(x) - (h(y) - g(y))| \leq M_2 \|x - y\| \text{ for all } x, y \in B(0, M_0)$$

and denote by  $\mathcal{U}_\phi(h, M_0, M_1, M_2)$  the set of all  $g \in \mathcal{U}(h, M_0, M_1, M_2)$  such that

$$(2.6) \quad g(x) \geq \phi(\|x\|) \text{ for all } x \in X.$$

Let  $f_0 : X \rightarrow R^1$  be a locally Lipschitzian function which satisfies

$$(2.7) \quad f_0(x) \geq \phi(\|x\|) \text{ for all } x \in X,$$

$g_0 : X \rightarrow R^1$  be a locally Lipschitzian function and let

$$(2.8) \quad a_0 < c_0 < b_0.$$

We assume that

$$(2.9) \quad g_0^{-1}(c_0) \neq \emptyset$$

and consider the following constrained minimization problems

$$(P_e) \quad f_0(x) \rightarrow \min \text{ subject to } x \in g_0^{-1}(c_0)$$

and

$$(P_i) \quad f_0(x) \rightarrow \min \text{ subject to } x \in g_0^{-1}((-\infty, c_0]).$$

We suppose that there exists  $\theta \in X$  such that the following assumption holds:

(A0) In the case of the problem  $(P_e)$

$$\theta \in g_0^{-1}(c_0) \text{ and } 0 \notin \partial g_0(\theta);$$

in the case of the problem  $(P_i)$

$$\theta \in g_0^{-1}((-\infty, c_0]) \text{ and if } \theta \in g_0^{-1}(c_0), \text{ then } 0 \notin \partial g_0(\theta).$$

Assume that

$$(2.10) \quad f : X \rightarrow R^1, \quad g : X \rightarrow R^1, \quad c \in R^1.$$

If  $g^{-1}(c) \neq \emptyset$ , then we consider the equality-constraint problem

$$(P_{c,e}^{(f,g)}) \quad f(x) \rightarrow \min \text{ subject to } x \in g^{-1}(c)$$

and if  $g^{-1}((-\infty, c]) \neq \emptyset$ , then we consider the inequality-constraint problem

$$(P_{c,i}^{(f,g)}) \quad f(x) \rightarrow \min \text{ subject to } x \in g^{-1}((-\infty, c]).$$

We associate with these two problems the corresponding families of unconstrained minimization problems

$$(P_{\lambda,c,e}^{(f,g)}) \quad f(x) + \lambda |g(x) - c| \rightarrow \min, \quad x \in X,$$

$$(P_{\lambda,c,i}^{(f,g)}) \quad f(x) + \lambda \max\{g(x) - c, 0\} \rightarrow \min, \quad x \in X$$

where  $\lambda > 0$ .

By (2.3) there is a positive number  $\bar{M}$  such that

$$(2.11) \quad \bar{M} > \|\theta\| + 4 \text{ and } \phi(\bar{M} - 4) > f_0(\theta) + 4.$$

In this paper we use the following assumptions:

(A1)  $f_0$  is Lipschitz on  $B(0, \bar{M})$ ;

(A2) for each  $\epsilon > 0$  there exists  $x_\epsilon \in g_0^{-1}(c_0)$  such that  $f_0(x_\epsilon) \leq \inf(f, c_0, g_0) + \epsilon$  and  $0 \notin \partial g_0(x_\epsilon)$ ;

(A3) for each  $\epsilon > 0$  there exists  $x_\epsilon \in g_0^{-1}((-\infty, c_0])$  such that

$$f_0(x_\epsilon) \leq \inf(f_0, (-\infty, c_0], g_0) + \epsilon$$

and if  $x_\epsilon \in g_0^{-1}(c_0)$ , then  $0 \notin \partial g_0(x_\epsilon)$ ;

(A4) if  $y \in g_0^{-1}(c_0)$  satisfies  $f_0(y) = \inf(f_0, c_0, g_0)$ , then  $0 \notin \partial g_0(y)$ ;

(A5) if  $y \in g_0^{-1}(c_0)$  satisfies  $f_0(y) = \inf(f_0; (-\infty, c_0]; g_0)$ , then  $0 \notin \partial g_0(y)$ ;

(A6) each sequence  $\{y_k\}_{k=1}^\infty \subset \{z \in B^0(0, \bar{M}) : g_0(z) \in [a_0, b_0]\}$  which satisfies  $\liminf_{k \rightarrow \infty} \Xi_{g_0}(y_k) = 0$  possesses a norm convergent subsequence;

(A7) each sequence

$$\{y_k\}_{k=1}^\infty \subset \{z \in B^0(0, \bar{M}) : g_0(z) \in [c_0, b_0]\}$$

which satisfies  $\liminf_{k \rightarrow \infty} \Xi_{g_0}(y_k) = 0$  possesses a norm convergent subsequence.

**Remark 2.1.** If (A1) holds, then by (2.7) and (2.11)  $f_0$  is bounded from below on  $X$ .

**Remark 2.2.** Assumption (A6) ((A7) respectively) means that  $g_0$  satisfies (P-S) condition on the set  $g_0^{-1}([a_0, b_0]) \cap B^0(0, \bar{M})$  ( $g_0^{-1}([c_0, b_0]) \cap B^0(0, \bar{M})$  respectively).

In this paper we prove the following two results.

**Theorem 2.1.** Suppose that (A1), (A2), (A4) and (A6) hold. Let  $M > 0$ . Then there exist  $\alpha > 0$ ,  $\lambda_0 > 0$  such that for each  $\epsilon > 0$  there exists  $\delta \in (0, \epsilon)$  such that the following assertion holds:

For each  $\lambda \geq \lambda_0$ , each  $c \in R^1$  satisfying  $|c - c_0| \leq \alpha$ , each pair of functions  $f : X \rightarrow R^1$  and  $g : X \rightarrow R^1$  which satisfy

$$f \in \mathcal{U}_\phi(f_0, \bar{M}, \alpha, M), g \in \mathcal{U}(g_0, \bar{M}, \alpha, \alpha)$$

and each  $x \in X$  satisfying

$$f(x) + \lambda|g(x) - c| \leq \inf\{f(z) + \lambda|g(z) - c| : z \in X\} + \delta$$

there is  $y \in g^{-1}(c)$  such that  $\|y - x\| \leq \epsilon$  and

$$f(y) \leq f(x) + \lambda|g(x) - c| \leq \inf\{f(z) + \lambda|g(z) - c| : z \in X\} + \delta.$$

**Theorem 2.2.** Suppose that (A1), (A3), (A5) and (A7) hold. Let  $M > 0$ . Then there exist  $\alpha > 0$ ,  $\lambda_0 > 0$  such that for each  $\epsilon > 0$  there exists  $\delta \in (0, \epsilon)$  such that the following assertion holds:

For each  $\lambda \geq \lambda_0$ , each  $c \in R^1$  satisfying  $|c - c_0| \leq \alpha$ , each pair of functions  $f : X \rightarrow R^1$  and  $g : X \rightarrow R^1$  which satisfy

$$f \in \mathcal{U}_\phi(f_0, \bar{M}, \alpha, M), \quad g \in \mathcal{U}(g_0, \bar{M}, \alpha, \alpha)$$

and each  $x \in X$  which satisfies

$$f(x) + \lambda \max\{g(x) - c, 0\} \leq \inf\{f(z) + \lambda \max\{g(z) - c, 0\} : z \in X\} + \delta$$

there is  $y \in g^{-1}((-\infty, c])$  such that  $\|y - x\| \leq \epsilon$  and

$$f(y) \leq f(x) + \lambda \max\{g(x) - c, 0\} \leq \inf\{f(z) + \lambda \max\{g(z) - c, 0\} : z \in X\} + \epsilon.$$

Theorems 2.1 and 2.2 will be proved in Section 4. They imply the following result.

**Theorem 2.3.**

- (1) Suppose that (A1), (A2), (A4) and (A6) hold. Let  $M > 0$ . Then there exist  $\alpha > 0$ ,  $\lambda_0 > 0$  such that for each  $f \in \mathcal{U}_\phi(f_0, \bar{M}, \alpha, M)$ , each  $g \in \mathcal{U}(g_0, \bar{M}, \alpha, \alpha)$ , each  $c \in [c_0 - \alpha, c_0 + \alpha]$ , each  $\lambda \geq \lambda_0$  and for each sequence  $\{x_i\}_{i=1}^\infty \subset X$  which satisfies

$$\lim_{i \rightarrow \infty} [f(x_i) + \lambda |g(x_i) - c|] = \inf\{f(z) + \lambda |g(z) - c| : z \in X\}$$

there exists a sequence  $\{y_i\}_{i=1}^\infty \subset g^{-1}(c)$  such that

$$\lim_{i \rightarrow \infty} f(y_i) = \inf(f, c, g), \quad \lim_{i \rightarrow \infty} \|y_i - x_i\| = 0.$$

- (2) Suppose that (A1), (A3), (A5) and (A7) hold. Let  $M > 0$ . Then there exist  $\alpha > 0$ ,  $\lambda_0 > 0$  such that for each  $f \in \mathcal{U}_\phi(f_0, \bar{M}, \alpha, M)$ , each  $g \in \mathcal{U}(g_0, \bar{M}, \alpha, \alpha)$ , each  $c \in [c_0 - \alpha, c_0 + \alpha]$ , each  $\lambda \geq \lambda_0$  and for each sequence  $\{x_i\}_{i=1}^\infty \subset X$  which satisfies

$$\lim_{i \rightarrow \infty} [f(x_i) + \lambda \max\{g(x_i) - c, 0\}] = \inf\{f(z) + \lambda \max\{g(z) - c, 0\} : z \in X\}$$

there exists a sequence  $\{y_i\}_{i=1}^\infty \subset g^{-1}((-\infty, c])$  such that

$$\lim_{i \rightarrow \infty} f(y_i) = \inf(f; (-\infty, c], g), \quad \lim_{i \rightarrow \infty} \|y_i - x_i\| = 0.$$

**Remark 2.3.** Note that Theorems 2.1 and 2.2 are generalizations of the main results of [18]. In [18] the assumptions (A2) and (A4) for the equality-constraint problem and the assumptions (A3) and (A5) for the inequality-constraint problem



were substituted by an essentially stronger assumption that  $c_0$  is not a critical value of the function  $g_0$ . In particular, for the equality-constraint problem the assumption (A4) means that the problem does not possess solutions which are critical points of  $g_0$  while (A2) follows from (A4) if the problem has a solution.

**Remark 2.4.** We prove Theorems 2.1 and 2.2 simultaneously but if one reads carefully the proof it is not difficult to see that the equality-constraint case is more complicated than the inequality-constraint case. For example, in the proof of Theorem 2.2 we do not need Proposition 3.1 which plays a crucial role in the proof of Theorem 2.1. Actually we present here Theorem 2.2 for the completeness of our presentation because the proof of Theorem 2.1 also works for Theorem 2.2. Note that in [19] we obtain a generalization of Theorem 2.2 for an inequality-constraint problem with a number of constraints.

### 3. AUXILIARY RESULTS

In this section we use the notation and definitions introduced in Section 2.

Let  $(Y, \|\cdot\|)$  and  $(Z, \|\cdot\|)$  be Banach spaces,  $A \subset Y$ ,  $B \subset Z$ . We say that  $h : A \rightarrow B$  is an  $\mathcal{L}$ -mapping if for each  $x \in A$  there exists  $r > 0$  such that the restriction  $h : A \cap B(x, r) \rightarrow B$  is Lipschitzian.

Assume that  $g : X \rightarrow R^1$  is a locally Lipschitz function. In the sequel we use the following auxiliary result of [16, Lemma 1].

**Lemma 3.1.** *Let  $\delta$  be a positive number and let  $A \subset X$  be a nonempty closed subset of  $X$  such that  $\Xi_g(x) < -\delta$  for all  $x \in A$ . Then there exists an  $\mathcal{L}$ -mapping  $V : X \rightarrow X$  such that*

$$\begin{aligned} \|Vx\| &\leq 2 \text{ for all } x \in X, \quad g^0(x, Vx) \leq 0 \text{ for all } x \in X, \\ g^0(x, Vx) &\leq -\delta \text{ for all } x \in A. \end{aligned}$$

**Lemma 3.2.** *Let  $x_0 \in X$ ,  $\delta > 0$  and let  $\Xi_g(x_0) < -\delta$ . Then there exist  $r > 0$  and an  $\mathcal{L}$ -mapping  $V : X \rightarrow X$  such that*

$$\begin{aligned} \|Vx\| &\leq 2 \text{ for all } x \in X, \quad g^0(x, Vx) \leq 0 \text{ for all } x \in X, \\ g^0(x, Vx) &\leq -\delta \text{ for all } x \in B(x_0, r). \end{aligned}$$

*Proof.* By upper semicontinuity of the Clarke generalized directional derivative  $g^0(\xi, \eta)$  with respect to  $\xi$  there exists  $r > 0$  such that  $\Xi_g(x) < -\delta$  for all  $x \in B(x_0, r)$ . Now Lemma 3.2 follows from Lemma 3.1. ■

**Proposition 3.1.** *Let*

$$(3.1) \quad x_0 \in B^0(0, \bar{M}) \cap g_0^{-1}(c_0),$$

$$(3.2) \quad 0 \notin \partial g_0(x_0)$$

*and let  $\epsilon_0$  be a positive number. Then there exists  $\delta_0 \in (0, \epsilon_0)$  such that for each function  $g : X \rightarrow R^1$  which satisfies*

$$(3.3) \quad |g(z) - g_0(z)| \leq \delta_0, \quad z \in B^0(0, \bar{M}),$$

$$(3.4) \quad |(g - g_0)(z_1) - (g - g_0)(z_2)| \leq \delta_0 \|z_1 - z_2\| \text{ for all } z_1, z_2 \in B^0(0, \bar{M})$$

*and each  $c \in R^1$  satisfying*

$$(3.5) \quad |c - c_0| \leq \delta_0$$

*there is  $x \in B^0(0, \bar{M})$  such that*

$$(3.6) \quad \|x - x_0\| < \epsilon_0, \quad g(x) = c.$$

*Proof.* By (2.1) and (3.2),

$$\Xi_{g_0}(x_0) < 0.$$

Choose  $\delta > 0$  such that

$$(3.7) \quad \Xi_{g_0}(x_0) < -\delta.$$

By Lemma 3.2 there exist  $r > 0$  and an  $\mathcal{L}$ -mapping  $V : X \rightarrow X$  such that

$$(3.8) \quad r < \epsilon_0, \quad B(x_0, r) \subset B^0(0, \bar{M}),$$

$$(3.9) \quad \|Vx\| \leq 2 \text{ for all } x \in X, \quad g_0^0(x, Vx) \leq 0 \text{ for all } x \in X,$$

$$(3.10) \quad g_0^0(x, Vx) \leq -\delta \text{ for all } x \in B(x_0, r).$$

It is clear that there exist  $t_0 > 0$  and a differentiable function  $\phi : [-t_0, t_0] \rightarrow X$  such that

$$(3.11) \quad \phi'(t) = V\phi(t), \quad t \in [-t_0, t_0],$$

$$(3.12) \quad \phi(0) = x_0,$$

$$(3.13) \quad \phi(t) \in B^0(x_0, r), \quad t \in [-t_0, t_0].$$

It follows from the properties of the Clarke generalized directional derivative [4] that for each  $t_1, t_2 \in [-t_0, t_0]$  satisfying  $t_1 < t_2$ ,

$$g_0(\phi(t_2)) - g_0(\phi(t_1)) = l(\phi'(s))(t_2 - t_1),$$

where  $s \in [t_1, t_2]$  and  $l \in \partial g_0(\phi(s))$ , and in view of (3.10), (3.11), (3.13) and the definition of Clarke's generalized gradient

$$g_0(\phi(t_2)) - g_0(\phi(t_1)) = (t_2 - t_1)l(V(\phi(s))) \leq (t_2 - t_1)g_0^0(\phi(s), V(\phi(s))) \leq -\delta(t_2 - t_1).$$

Thus

$$(3.14) \quad \begin{aligned} g_0(\phi(t_2)) - g_0(\phi(t_1)) &\leq -\delta(t_2 - t_1) \\ \text{for each } t_1, t_2 &\in [-t_0, t_0] \text{ such that } t_1 < t_2. \end{aligned}$$

Choose a positive number  $\delta_0$  such that

$$(3.15) \quad \delta_0 < \min\{\epsilon_0, \delta t_0/8\}.$$

Assume that  $g : X \rightarrow R^1$  satisfies (3.4) and  $c \in R^1$  satisfies (3.5). Relations (3.8) and (3.13) imply that for all  $t \in [-t_0, t_0]$ ,

$$(3.16) \quad \phi(t) \in B^0(x_0, r) \subset B^0(0, \bar{M}) \cap B^0(x_0, \epsilon_0).$$

In view of (3.1), (3.12) and (3.14),

$$(3.17) \quad g_0(\phi(t_0/2)) \leq g_0(\phi(0)) - \delta t_0/2 = g_0(x_0) - \delta t_0/2 = c_0 - \delta t_0/2,$$

$$(3.18) \quad g_0(\phi(-t_0/2)) \geq g_0(\phi(0)) + \delta t_0/2 = g_0(x_0) + \delta t_0/2 = c_0 + \delta t_0/2.$$

By (3.4), (3.8) and (3.13),

$$(3.19) \quad |(g - g_0)(\phi(t_0/2))|, |(g - g_0)(\phi(-t_0/2))| \leq \delta_0.$$

By (3.19), (3.17), (3.5), (3.15) and (3.18)

$$g(\phi(t_0/2)) \leq g_0(\phi(t_0/2)) + \delta_0 \leq c_0 - \delta t_0/2 + \delta_0 \leq c - \delta t_0/2 + 2\delta_0 \leq c - 2\delta_0$$

and

$$g(\phi(-t_0/2)) \geq g_0(\phi(-t_0/2)) - \delta_0 = c_0 + \delta t_0/2 - \delta_0 \geq c + \delta_0/2 - 2\delta_0 \geq c + 2\delta_0.$$

These inequalities imply that there is  $s \in (-t_0/2, t_0/2)$  such that

$$g(\phi(s)) = c.$$

It follows from (3.13) and (3.16) that

$$\phi(s) \in B^0(0, \bar{M}), \|\phi(s) - x_0\| < \epsilon_0.$$

This completes the proof of Proposition 3.1. ■

#### 4. PROOF OF THEOREMS 2.1 AND 2.2

We prove Theorems 2.1 and 2.2 simultaneously. For each  $c \in R^1$  and each function  $g : X \rightarrow R^1$  set

$$(4.1) \quad A_{g,c} = g^{-1}(c) \text{ in the case of Theorem 2.1,}$$

$$A_{g,c} = g^{-1}((-\infty, c]) \text{ in the case of Theorem 2.2.}$$

For each  $\lambda > 0$ , each  $c \in R^1$  and each pair of functions  $f, g : X \rightarrow R^1$  define a function  $\psi_{\lambda,c}^{(f,g)} : X \rightarrow R^1$  as follow:

$$(4.2) \quad \psi_{\lambda,c}^{(f,g)}(z) = f(z) + \lambda|g(z) - c|, \quad z \in X$$

in the case of Theorem 2.1 and

$$(4.3) \quad \psi_{\lambda,c}^{(f,g)}(z) = f(z) + \lambda \max\{g(z) - c, 0\}, \quad z \in X$$

in the case of Theorem 2.2.

Let  $M > 0$ . We assume that Theorem 2.1 (Theorem 2.2 respectively) does not hold. Then for each natural number  $k$  there exist

$$\epsilon_k \in (0, 1), \quad \lambda_k \geq k, \quad c_k \in R^1$$

satisfying

$$(4.4) \quad |c_k - c_0| \leq k^{-1},$$

a pair of functions  $f_k, g_k : X \rightarrow R^1$  which satisfy

$$(4.5) \quad f_k \in \mathcal{U}_\phi(f_0, \bar{M}, k^{-1}, M), \quad g_k \in \mathcal{U}(g_0, \bar{M}, k^{-1}, k^{-1})$$

and  $x_k \in X$  satisfying

$$(4.6) \quad \psi_{\lambda_k, c_k}^{(f_k, g_k)}(x_k) \leq \inf(\psi_{\lambda_k, c_k}^{(f_k, g_k)}) + \epsilon_k 2^{-1} k^{-2},$$

$$(4.7) \quad \{y \in A_{g_k, c_k} \cap B(x_k, \epsilon_k) : \psi_{\lambda_k, c_k}^{(f_k, g_k)}(y) \leq \psi_{\lambda_k, c_k}^{(f_k, g_k)}(x_k)\} = \emptyset.$$

Set

$$(4.8) \quad \phi^{(k)} = \psi_{\lambda_k, c_k}^{(f_k, g_k)}, \quad k = 1, 2, \dots$$

Let  $k$  be a natural number. By Ekeland's variational principle [6], (4.2), (4.3), (4.5), (4.6) and (4.8) there is  $y_k \in X$  such that

$$(4.9) \quad \phi^{(k)}(y_k) \leq \phi^{(k)}(x_k),$$

$$(4.10) \quad \|y_k - x_k\| \leq (2k)^{-1} \epsilon_k,$$

$$(4.11) \quad \phi^{(k)}(y_k) \leq \phi^{(k)}(z) + k^{-1} \|z - y_k\| \text{ for all } z \in X.$$

By (4.7), (4.8), (4.9) and (4.10),

$$(4.12) \quad y_k \notin A_{g_k, c_k}.$$

In the case of Theorem 2.2 we obtain that

$$(4.13) \quad g_k(y_k) > c_k.$$

In the case of Theorem 2.1 we obtain that either  $g_k(y_k) > c_k$  or  $g_k(y_k) < c_k$ . In the case of Theorem 2.1 extracting a subsequence, re-indexing and replacing  $g_0$  by  $-g_0$ ,  $g_k$  by  $-g_k$ ,  $c_0$  by  $-c_0$  and  $c_k$  by  $-c_k$ , we may assume without loss of generality that in both cases

$$(4.14) \quad g_k(y_k) > c_k \text{ for all natural numbers } k.$$

In view of (A1) there exists a number  $L_0 > 0$  such that

$$(4.15) \quad |f_0(z_1) - f_0(z_2)| \leq L_0 \|z_1 - z_2\| \text{ for all } z_1, z_2 \in B(0, \bar{M}).$$

It is easy to see that for all natural numbers  $k$ ,  $\phi^{(k)}$  is locally Lipschitz on  $B(0, \bar{M})$ .

We continue the proof with four steps.

**Step 1.** We show that  $\|y_k\| \leq \bar{M} - 4$  for all sufficiently large natural numbers  $k$ . It follows from (1.11), (A0), Proposition 3.1, (4.1), (4.4) and (4.5) that there is a natural number  $k_0$  such that for each integer  $k \geq k_0$  there exists

$$(4.16) \quad \theta_k \in B^0(0, \bar{M}) \cap A_{g_k, c_k}$$

such that

$$(4.17) \quad \lim_{k \rightarrow \infty} \|\theta_k - \theta\| = 0.$$

It follows from (1.11), (4.2), (4.3), (4.5), (4.6), (4.8), (4.9), (4.15) and (4.16) that for each integer  $k \geq k_0$

$$\begin{aligned}
(4.18) \quad & f_k(y_k) \leq \phi^{(k)}(y_k) \leq \phi^{(k)}(x_k) \leq \inf(\phi^{(k)}) + 2^{-1}k^{-2} \\
& \leq \phi^{(k)}(\theta_k) + 2^{-1}k^{-2} = f_k(\theta_k) + 2^{-1}k^{-2} \\
& = f_0(\theta_k) + [f_k(\theta_k) - f_0(\theta_k)] + 2^{-1}k^{-2} \\
& \leq f_0(\theta_k) + k^{-1} + k^{-2} = f_0(\theta) + [f_0(\theta_k) - f_0(\theta)] + k^{-1} + k^{-2} \\
& \leq f_0(\theta) + L_0\|\theta_k - \theta\| + 2k^{-1}.
\end{aligned}$$

Since  $\lim_{k \rightarrow \infty} \|\theta_k - \theta\| = 0$  it follows from (4.18) that there exists a natural number  $k_1 \geq k_0$  such that for all integers  $k \geq k_1$ ,

$$f_k(y_k) \leq f_0(\theta) + 2^{-1}.$$

Together with (1.6), (1.11) and (4.5) this implies that

$$(4.19) \quad \|y_k\| \leq \bar{M} - 4 \text{ for all integers } k \geq k_1.$$

**Step 2.** We show that  $\lim_{k \rightarrow \infty} \Xi_{g_0}(y_k) = 0$ . Let  $k \geq k_1$  be a natural number. By (4.5) and (4.19) for each  $z \in B(y_k, 3/2)$ ,

$$(4.20) \quad |[g_k(z) - g_0(z)] - [g_k(y_k) - g_0(y_k)]| \leq k^{-1}\|z - y_k\|,$$

$$(4.21) \quad |[f_k(z) - f_0(z)] - [f_k(y_k) - f_0(y_k)]| \leq M\|z - y_k\|.$$

By (4.14) there exists a positive number  $r_k \leq 3/2$  such that

$$(4.22) \quad g_k(z) > c_k \text{ for all } z \in B(y_k, r_k).$$

In view of (4.22), (4.11), (4.8), (4.2) and (4.3) for all  $z \in B(y_k, r_k)$ ,

$$(4.23) \quad f_k(y_k) + \lambda_k(g_k(y_k) - c_k) \leq f_k(z) + \lambda_k(g_k(z) - c_k) + k^{-1}\|z - y_k\|.$$

By (4.20), (4.21), (4.22) and (4.23) for all  $z \in B(y_k, r_k)$ ,

$$\begin{aligned}
(4.24) \quad & f_0(y_k) + \lambda_k g_0(y_k) \\
& \leq f_k(y_k) + \lambda_k g_k(y_k) + [f_0(y_k) - f_k(y_k)] + \lambda_k [g_0(y_k) - g_k(y_k)] \\
& \leq f_k(z) + \lambda_k g_k(z) + k^{-1}\|z - y_k\| + [f_0(y_k) - f_k(y_k)] \\
& \quad - \lambda_k [g_0(y_k) - g_k(y_k)] \\
& \leq f_0(z) + \lambda_k g_0(z) + (f_k(z) - f_0(z)) + \lambda_k [g_k(z) - g_0(z)] \\
& \quad + k^{-1}\|z - y_k\| + [f_0(y_k) - f_k(y_k)] + \lambda_k [(g_0(y_k) - g_k(y_k))] \\
& = f_0(z) + \lambda_k g_0(z) + k^{-1}\|z - y_k\| \\
& \quad + [(f_k - f_0)(z) - (f_k - f_0)(y_k)] + \lambda_k [(g_k - g_0)(z) - (g_k - g_0)(y_k)] \\
& \leq f_0(z) + \lambda_k g_0(z) + k^{-1}\|z - y_k\| + M\|z - y_k\| + \lambda_k k^{-1}\|z - y_k\|.
\end{aligned}$$

In view of (4.24) for all  $z \in B(y_k, z_k)$

$$g_0(y_k) + \lambda_k^{-1} f_0(y_k) \leq g_0(z) + \lambda_k^{-1} f_0(z) + \lambda_k^{-1} \|z - y_k\| (k^{-1} + M + k^{-1}).$$

It follows from the inequality above and the properties of Clarke's generalized gradient [4, Chapter 2, Sect. 2.3] that

$$(4.25) \quad 0 \in \partial g_0(y_k) + \lambda_k^{-1} \partial f_0(y_k) + \lambda_k^{-1} (2/k + M) \{l \in X^* : \|l\|_* \leq 1\}.$$

By (4.15) and (4.19) for each natural number  $k \geq k_1$ ,

$$\partial f_0(y_k) \subset \{l \in X^* : \|l\|_* \leq L_0\}.$$

Together with (4.25) this implies that for each natural  $k \geq k_1$

$$0 \in \partial g_0(y_k) + \lambda_k^{-1} (L_0 + 2/k + M) \{l \in X^* : \|l\|_* \leq 1\}$$

and

$$(4.26) \quad \lim_{k \rightarrow \infty} \Xi_{g_0}(y_k) = 0.$$

**Step 3.** Let us show that  $\lim_{k \rightarrow \infty} g_0(y_k) = c_0$  and that  $\{y_j\}_{j=1}^\infty$  possesses a convergent subsequence. By (4.19), (4.5), (4.14), (4.8), (4.2), (4.3), (4.9), (4.6) and (4.16) for all natural number  $k \geq k_1$ ,

$$\begin{aligned} & -M + \lambda_k(g_k(y_k) - c_k) + f_0(y_k) \leq f_k(y_k) + \lambda_k(g_k(y_k) - c_k) \\ & = \phi^{(k)}(y_k) \leq \phi^{(k)}(x_k) \leq \inf(\phi^{(k)}) + 1 \leq \phi^{(k)}(\theta_k) + 1 \leq f_k(\theta_k) + 1 \\ & = f_0(\theta_k) + 1 + f_k(\theta_k) - f_0(\theta_k) \leq f_0(\theta_k) + 2 \end{aligned}$$

and

$$0 < g_k(y_k) - c_k \leq \lambda_k^{-1} [\sup\{f_0(z) : z \in B(0, \bar{M})\} - \inf(f_0) + 2 + M].$$

Thus  $\lim_{k \rightarrow \infty} (g_k(y_k) - c_k) = 0$ . Together with (4.4), (4.5) and (4.19) this implies that

$$(4.27) \quad \lim_{k \rightarrow \infty} g_0(y_k) = c_0.$$

By (4.26), (4.27), (4.19), (A6), (A7), (1.8) and (4.13) there exists a strictly increasing sequence of natural numbers  $\{k_j\}_{j=1}^\infty$  such that  $\{y_{k_j}\}_{j=1}^\infty$  converges in the norm topology of  $X$  to  $y_* \in X$ . In view of (4.27)

$$(4.28) \quad g_0(y_*) = c_0.$$

**Step 4.** Let us show that  $f_0(y_*) = \inf(f_0; A_{g_0, c_0})$ .

It follows from (4.26) and upper semicontinuity of the Clarke generalized directional derivative  $g^0(\xi, \eta)$  with respect to  $\xi$  [4] that  $\Xi_{g_0}(y_*) = 0$  and

$$(4.29) \quad 0 \in \partial g_0(y_*).$$

In view of (4.15) and (4.19),

$$(4.30) \quad f_0(y_*) = \lim_{j \rightarrow \infty} f_0(y_{k_j}).$$

Let  $\Delta \in (0, 1)$ . It follows from (A2) and (A3) that there exists

$$(4.31) \quad x(\Delta) \in A_{g_0, c_0}$$

such that

$$(4.32) \quad f_0(x(\Delta)) \leq \inf(f_0; A(g_0, c_0)) + \Delta,$$

$$(4.33) \quad 0 \notin \partial g_0(x(\Delta)) \text{ if } g(x(\Delta)) = c_0.$$

By (4.31) and (A0),

$$(4.34) \quad f_0(x(\Delta)) \leq f_0(\theta) + 1.$$

It follows from (4.34), (1.7) and (1.11) that

$$(4.35) \quad \|x(\Delta)\| < \bar{M} - 4.$$

By (4.35), (4.31), (4.33), (4.4), (4.5) and Proposition 3.1 there is a natural number  $k_2 \geq k_1$  such that for each integer  $k \geq k_2$  there is

$$(4.36) \quad x_k(\Delta) \in B^0(0, \bar{M}) \cap A_{g_k, c_k}$$

such that

$$(4.37) \quad \lim_{k \rightarrow \infty} \|x_k(\Delta) - x(\Delta)\| = 0.$$

In view of (4.19), (4.5), (4.8), (4.2), (4.3), (4.9), (4.6), (4.36), (4.19) and (4.15) for all integers  $k \geq k_2$ ,

$$\begin{aligned} f_0(y_k) - k^{-1} &\leq f_k(y_k) \leq \phi^{(k)}(y_k) \leq \phi^{(k)}(x_k) \leq \inf(\phi^{(k)}) + 2^{-1}k^{-2} \\ &\leq \phi^{(k)}(x_k(\Delta)) + 2^{-1}k^{-2} \\ &= f_k(x_k(\Delta)) + 2^{-1}k^{-2} = f_0(x_k(\Delta)) + [f_k(x_k(\Delta)) - f_0(x_k(\Delta))] + 2^{-1}k^{-2} \\ &\leq f_0(x_k(\Delta)) + k^{-1} + 2^{-1}k^{-2} \\ &= f_0(x(\Delta)) + [f_0(x_k(\Delta)) - f_0(x(\Delta))] + 1/k + (2k^2)^{-1} \\ &\leq f_0(x(\Delta)) + L_0\|x_k(\Delta) - x(\Delta)\| + 2k^{-1} \end{aligned}$$



and combined with (4.32) this implies that

$$f_0(y_k) \leq 3k^{-1} + L_0 \|x_k(\Delta) - x(\Delta)\| + \inf(f_0; A_{g_0, c_0}) + \Delta.$$

Together with (4.37) this implies that

$$\limsup_{k \rightarrow \infty} f_0(y_k) \leq \inf(f_0; A_{g_0, c_0}) + \Delta.$$

Since  $\Delta$  is an arbitrary number from the interval  $(0, 1)$  we conclude that

$$\limsup_{k \rightarrow \infty} f_0(y_k) \leq \inf(f_0; A_{g_0, c_0}).$$

Combined with (4.30) this implies that

$$f_0(y_*) = \lim_{j \rightarrow \infty} f_0(y_{k_j}) \leq \inf(f_0; A_{g_0, c_0})$$

and in view of (4.28),

$$(4.38) \quad f_0(y_*) = \inf(f_0; A_{g_0, c_0}).$$

Relations (4.38), (4.28) and (4.29) contradict (A4) in the case of Theorem 2.1 and (A5) in the case of Theorem 2.2. The contradiction we have reached proves Theorems 2.1 and 2.2

#### REFERENCES

1. T. Q. Bao and B. S. Mordukhovich, Variational principles for set-valued mappings with applications to multiobjective optimization, *Control and Cybernetics*, **36** (2007).
2. D. Boukari and A. V. Fiacco, Survey of penalty, exact-penalty and multiplier methods from 1968 to 1993 *Optimization*, **32** (1995), 301-334.
3. J. V. Burke, An exact penalization viewpoint of constrained optimization, *SIAM J. Control Optim.*, **29** (1991), 968-998.
4. F. H. Clarke, *Optimization and Nonsmooth Analysis*, Willey Interscience, 1983.
5. G. Di Pillo and L. Grippo, Exact penalty functions in constrained optimization, *SIAM J. Control Optim.*, **27** (1989), 1333-1360.
6. I. Ekeland, On the variational principle *J. Math. Anal. Appl.*, **47** (1974), 324-353.
7. I. I. Eremin, The penalty method in convex programming, *Soviet Math. Dokl.*, **8** (1966), 459-462.
8. S.-P. Han and O. L. Mangasarian, Exact penalty function in nonlinear programming, *Math. Programming*, **17** (1979), 251-269.

9. J.-B. Hiriart-Urruty and C. Lemarechal, *Convex Analysis and Minimization Algorithms*, Springer, Berlin, 1993.
10. D. G. Luenberger, Control problems with kinks *IEEE Trans. Automat. Control*, **15** (1970), 570-575.
11. B. S. Mordukhovich, *Variational Analysis and Generalized Differentiation, I: Basic Theory, II: Applications*, Springer, Berlin, 2006.
12. J. Outrata, A note on the usage of nondifferentiable exact penalties in some special optimization problems, *Kybernetika*, **24** (1988), 251-258.
13. R. T. Rockafellar, *Penalty methods and augmented Lagrangians in nonlinear programming*, Fifth Conference on Optimization Techniques (Rome, 1973), Part I. Lecture Notes in Comput. Sci., Springer, Berlin, 3, 1973, pp. 418-425.
14. R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Springer, Berlin, 1998.
15. W. I. Zangwill, Nonlinear programming via penalty functions, *Management Sci.*, **13** (1967), 344-358.
16. A. J. Zaslavski, On critical points of Lipschitz functions on smooth manifolds, *Siberian Math. J.*, **22** (1981), 63-68.
17. A. J. Zaslavski, A sufficient condition for exact penalty in constrained optimization, *SIAM Journal on Optimization*, **16** (2005), 250-262.
18. A. J. Zaslavski, Existence of exact penalty and its stability for nonconvex constrained optimization problems in Banach spaces, *Set-Valued Analysis*, in press.
19. A. J. Zaslavski, Stability of exact penalty for nonconvex inequality-constrained minimization problems, *Taiwanese J. Math.*, in accepted.

Alexander J. Zaslavski  
Department of Mathematics,  
The Technion-Israel Institute of Technology,  
32000 Haifa,  
Israel  
E-mail: ajzasl@tx.technion.ac.il