Nonlinear Functional Analysis and Applications
Vol. 13, No. 3 (2008), pp. 503-514

http://nfaa.kyungnam.ac.kr/jour-nfaa.htm

SUFFICIENT CONDITIONS FOR THE EXACT PENALTY PROPERTY OF CONSTRAINED MINIMIZATION IN HILBERT SPACES

Alexander J. Zaslavski

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

Abstract. We use the penalty methods in order to study two constrained minimization problems in Hilbert spaces. A penalty function is said to have the exact penalty property if there is a penalty coefficient for which a solution of an unconstrained penalized problem is a solution of the corresponding constrained problem. We establish simple sufficient conditions for the exact penalty property.

1. Introduction and the main results

Penalty methods are an important and useful tool in constrained optimization. See, for example, [1], [2], [4]-[6], [8]-[13], [15]-[20] and the references mentioned there. In this paper we use the penalty approach in order to study two constrained nonconvex minimization problems with smooth cost functions. Note that classes of minimization problems with smooth objective functions and smooth constraints are considered in [3]. A penalty function is said to have the exact penalty property [1], [2], [8], [10] if there is a penalty coefficient for which a solution of an unconstrained penalized problem is a solution of the corresponding constrained problem. The notion of exact penalization was introduced in [9], [18]. For a review of the literature on exact penalization see [1], [2], [8]. We will establish simple sufficient conditions for the exact penalty property.

[0]Received February 4, 2007. Revised April 12, 2007.
[0]2000 Mathematics Subject Classification: 49M30, 90C26, 90C30.
[0]Keywords: Critical point, Frechet differentiable function, Hilbert space, minimization problem, penalty function.

Let $(X, <\cdot,\cdot>)$ be a Hilbert space with the inner product $<\cdot,\cdot>$ and the norm $||x|| = <x,x>^{1/2}$, $x \in X$. Let U be a nonempty open subset of X. Denote by $C^1(U;R^1)$ the set of all Frechet differentiable functions $f : U \to R^1$ such that the mapping $x \to f'(x)$, $x \in U$ is continuous. Here $f'(x) \in X$ is a Frechet derivative of f at $x \in U$.

Denote by $C^2(U;R^1)$ the set of all Frechet differentiable functions $f \in C^1(U;R^1)$ such that the mapping $x \to f'(x)$, $x \in U$ is also Frechet differentiable and that the mapping $x \to f''(x)$, $x \in U$ is continuous. Here $f''(x)$ is a Frechet second order derivative of f at $x \in U$. It is a linear continuous self-mapping of X.

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

$$B(x,r) = \{y \in X : \ ||x-y|| \le r\},\ B^o(x,r) = \{y \in X : \ ||x-y|| < r\},$$

$$B(r) = B(0,r),\ B^o(r) = B^o(0,r).$$

For each function $h : X \to R^1 \cup \{\infty\}$ and each $A \subset X$ put

$$\inf(h) = \inf\{h(z) : \ z \in X\} \quad \text{and} \quad \inf(h;A) = \inf\{h(x) : \ x \in A\}.$$

For each $x \in X$ and each $B \subset X$ set $d(x,B) = \inf\{||x-y|| : \ y \in B\}$.

Let $g \in C^2(U;R^1)$. A point $x \in U$ is called a critical point of g if $g'(x) = 0$. Denote by $Cr(g)$ the set of all critical points of g. A real number c is a critical value of g if there exists $x \in Cr(g)$ such that $g(x) = c$. Denote by $Cr(g,+)$ the set of all $x \in Cr(g)$ such that

$$< g''(x)u, u > \ge 0 \quad \text{for all} \quad u \in X$$

and by $Cr(g,-)$ the set of all $x \in Cr(g)$ such that

$$< g''(x)u, u > \le 0 \quad \text{for all} \quad u \in X.$$

Let $g : X \to R^1$ be a continuous function, $c \in R^1$ be such that $g^{-1}(c) \neq \emptyset$ and let $f : X \to R^1 \cup \{\infty\}$ be a lower semicontinuous function which is bounded from below and satisfies the growth condition

$$\lim_{||x|| \to \infty} f(x) = \infty. \tag{1.1}$$

Clearly, $g^{-1}(c)$ is a closed subset of $(X, ||\cdot||)$. We consider the constrained problems

$$f(x) \to \min \quad \text{subject to} \quad x \in g^{-1}(c) \tag{P_e}$$

and

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

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

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

and

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

where $\lambda > 0$.

In [19] we showed the existence of exact penalty for problems (P_i) and (P_e) with locally Lipschitzian functions f and g. In [20] assuming that $g \in C^2(X; R^1)$, $f \in C^1(X; R^1)$ and that the mapping $f' : X \to X$ is locally Lipschitzian we established that exact penalty exists for the problem (P_e) if $g^{-1}(c) \cap (Cr(g,+) \cup Cr(g,-)) = \emptyset$ and that exact penalty exists for the problem (P_i) if $g^{-1}(c) \cap Cr(g,+) = \emptyset$. In this paper we improve the results of [20]. It turns our that they remain in force if g and f are smooth only in small neighborhoods of minimizers of problems (P_e) and (P_i) respectively. (See assumptions (A2) and (A3) below).

In this paper we suppose that there exists $\gamma > 0$ such that the following assumption holds.

(A1) For each $r > 0$ the set $B(r) \cap g^{-1}([c - \gamma, c + \gamma])$ is compact.

Note that in [20] instead of (A1) we use a Palais-Smale type condition [14]. In this paper we also use the following two assumptions.

(A2) If $x \in g^{-1}(c)$ satisfies $f(x) = \inf(f; g^{-1}(c))$, then there exists $\Delta_x > 0$ such that:

the restriction of g to $B^o(x, \Delta_x)$ denoted by $\tilde{g}$ belongs to $C^2(B^o(x, \Delta_x); R^1)$ and $x \notin Cr(\tilde{g}+) \cup Cr(\tilde{g}-)$;

the restriction of f to $B^o(x, \Delta_x)$ belongs to $C^1(B^o(x, \Delta_x); R^1)$ and is Lipschitz and the mapping $f' : B^o(x, \Delta_x) \to X$ is locally Lipschitz.

(A3) If $x \in g^{-1}(c)$ satisfies $f(x) = \inf(f; g^{-1}((-\infty, c]))$, then there exists $\Delta_x > 0$ such that:

the restriction of g to $B^o(x, \Delta_x)$ denoted by $\tilde{g}$ belongs to $C^2(B^o(x, \Delta_x); R^1)$ and $x \notin Cr(\tilde{g}+)$;

the restriction of f to $B^o(x, \Delta_x)$ belongs to $C^1(B^o(x, \Delta_x); R^1)$ and is Lipschitz and the mapping $f' : B^o(x, \Delta_x) \to X$ is locally Lipschitz.

The next two theorems are the main results of the paper.

Theorem 1.1. *Assume that (A2) holds and that* $\inf(f; g^{-1}(c)) < \infty$. *Then there exists a positive number* Λ_0 *such that for each* $\epsilon > 0$ *there exists* $\delta \in (0, \epsilon)$ *such that the following assertion holds:*

If $\lambda \geq \Lambda_0$ *and if* $x \in X$ *satisfies*

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

then there exists $y \in g^{-1}(c)$ *such that*

$$||y - x|| \leq \epsilon \text{ and } f(y) \leq \inf(f; g^{-1}(c)) + \epsilon.$$

Theorem 1.2. *Assume that (A3) holds and that* $\inf(f; g^{-1}((-\infty, c])) < \infty$. *Then there is* $\Lambda_0 > 0$ *such that for each* $\epsilon > 0$ *there exists* $\delta \in (0, \epsilon)$ *such that the following assertion holds:*

If $\lambda \geq \Lambda_0$ *and if* $x \in X$ *satisfies*

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

then there exists $y \in g^{-1}((-\infty, c])$ *such that*

$$||y - x|| \leq \epsilon \text{ and } f(y) \leq \inf(f; g^{-1}((-\infty, c])) + \epsilon.$$

Theorems 1.1 and 1.2 will be proved in Section 2. In this section we present several important results which easily follow from Theorems 1.1 and 1.2.

Theorems 1.1 and 1.2 imply the following result.

Theorem 1.3. *1. Assume that (A2) holds and that* $\inf(f; g^{-1}(c)) < \infty$. *Then there exists* $\Lambda_0 > 0$ *such that for each* $\lambda \geq \Lambda_0$ *and each sequence* $\{x_i\}_{i=1}^{\infty} \subset X$ *which satisfies*

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

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

$$\lim_{i \to \infty} f(y_i) = \inf(f; g^{-1}(c)) \text{ and } \lim_{i \to \infty} ||y_i - x_i|| = 0.$$

2. Assume that (A3) holds and that $\inf(f; g^{-1}((-\infty, c])) < \infty$. *Then there exists* $\Lambda_0 > 0$ *such that for each* $\lambda \geq \Lambda_0$ *and each sequence* $\{x_i\}_{i=1}^{\infty} \subset X$ *which satisfies*

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

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

$$\lim_{i \to \infty} f(y_i) = \inf(f; g^{-1}((-\infty, c])) \text{ and } \lim_{i \to \infty} ||y_i - x_i|| = 0.$$

The next result easily follows from Theorem 1.3.

Theorem 1.4. *1. Assume that (A2) holds and that* $\inf(f; g^{-1}(c)) < \infty$. *Then there exists* $\Lambda_0 > 0$ *such that if* $\lambda \geq \Lambda_0$ *and if* $x \in X$ *satisfies*

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

then $g(x) = c$ *and* $f(x) = \inf(f; g^{-1}(c))$.

2. Assume that (A3) holds and that $\inf(f; g^{-1}((-\infty, c])) < \infty$. *Then there exists* $\Lambda_0 > 0$ *such that if* $\lambda \geq \Lambda_0$ *and if* $x \in X$ *satisfies*

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

then $g(x) \leq c$ *and* $f(x) = \inf(f; g^{-1}((-\infty, c]))$.

Theorem 1.3 implies the following result.

Theorem 1.5. *1. Assume that (A2) holds and that $\bar{x} \in g^{-1}(c)$ satisfies the following conditions:*

$$f(\bar{x}) = \inf(f; g^{-1}(c)) < \infty;$$

any sequence $\{x_n\}_{n=1}^{\infty} \subset g^{-1}(c)$ which satisfies

$$\lim_{n\to\infty} f(x_n) = \inf(f; g^{-1}(c))$$

converges to $\bar{x}$ in the norm topology.

Then there exists $\Lambda_0 > 0$ such that for each $\lambda \geq \Lambda_0$ the point $\bar{x}$ is a unique solution of the minimization problem $f(z) + \lambda|g(z) - c| \to \min,\ z \in X$.

2. Assume that (A3) holds and that $\bar{x} \in g^{-1}((-\infty, c])$ satisfies the following conditions:

$$f(\bar{x}) = \inf(f; g^{-1}((-\infty, c])) < \infty;$$

any sequence $\{x_n\}_{n=1}^{\infty} \subset g^{-1}((-\infty, c])$ which satisfies

$$\lim_{n\to\infty} f(x_n) = \inf(f; g^{-1}((-\infty, c]))$$

converges to $\bar{x}$ in the norm topology.

Then there exists $\Lambda_0 > 0$ such that for each $\lambda \geq \Lambda_0$ the point $\bar{x}$ is a unique solution of the minimization problem $f(z) + \lambda \max\{g(z) - c, 0\} \to \min,\ z \in X$.

2. Proofs of theorems 1.1 and 1.2

We prove Theorems 1.1 and 1.2 simultaneously. There exists $\bar{a} > 0$ such that

$$f(x) \geq -\bar{a} \text{ for all } x \in X. \tag{2.1}$$

Set $A = g^{-1}(c)$ in the case of Theorem 1.1 and $A = g^{-1}((-\infty, c])$ in the case of Theorem 1.2. Clearly A is a nonempty closed subset of $(X, ||\cdot||)$. For each $\lambda > 0$ we define a function $\psi_\lambda : X \to R^1 \cup \{\infty\}$ as follows:

$$\psi_\lambda(z) = f(z) + \lambda|g(z) - c|,\ z \in X \tag{2.2}$$

in the case of Theorem 1.1 and

$$\psi_\lambda(z) = f(z) + \lambda \max\{g(z) - c, 0\},\ z \in X \tag{2.3}$$

in the case of Theorem 1.2. Clearly, the function ψ_λ is lower semicontinuous for all $\lambda > 0$.

We show that there is $\Lambda_0 > 0$ such that the following property holds:

(P1) For each $\epsilon > 0$ there exists $\delta \in (0, \epsilon)$ such that for each $\lambda \geq \Lambda_0$ and each $x \in X$ which satisfies $\psi_\lambda(x) \leq \inf(\psi_\lambda) + \delta$ there is $y \in A$ for which $||y - x|| \leq \epsilon$ and $\psi_\lambda(y) \leq \inf(\psi_\lambda) + \epsilon$.

It is not difficult to see that Theorems 1.1 and 1.2 follow from (P1).

Let us assume that there is no Λ_0 for which (P1) holds. Then for each natural number k there exist

$$\epsilon_k \in (0,1),\ \lambda_k \geq k,\ x_k \in X \tag{2.4}$$

such that

$$\psi_{\lambda_k}(x_k) \leq \inf(\psi_{\lambda_k}) + (8k^2)^{-1}\epsilon_k^2, \tag{2.5}$$

$$\{z \in A \cap B(x_k, \epsilon_k) :\ \psi_{\lambda_k}(z) \leq \inf(\psi_{\lambda_k}) + \epsilon_k\} = \emptyset. \tag{2.6}$$

Let k be a natural number. Consider the function

$$\phi_{\lambda_k}(z) = \psi_{\lambda_k}(z) + (4k^2)^{-1}||z - x_k||^2,\ z \in X. \tag{2.7}$$

Clearly, the function ϕ_{λ_k} is lower semicontinuous and bounded from below. By the variational principle of Deville-Godefroy-Zizler [7] there exist $h_k \in C^2(X; R^1)$ and $y_k \in X$ such that

$$\sup\{||h_k(z)|| + ||h_k'(z)|| + ||h_k''(z)|| :\ z \in X\} \leq 32^{-1}k^{-2}\epsilon_k^2, \tag{2.8}$$

$$(\phi_{\lambda_k} + h_k)(z) > (\phi_{\lambda_k} + h_k)(y_k) \quad \text{for all} \quad z \in X \setminus \{y_k\}. \tag{2.9}$$

We show that $||x_k - y_k|| < \epsilon_k$. Let us assume the contrary. Then

$$||y_k - x_k|| \geq \epsilon_k. \tag{2.10}$$

By (2.7) and (2.5),

$$\phi_{\lambda_k}(x_k) = \psi_{\lambda_k}(x_k) \leq \inf(\psi_{\lambda_k}) + (8k^2)^{-1}\epsilon_k^2. \tag{2.11}$$

In view of (2.7), (2.10) and (2.5),

$$\begin{aligned}\phi_{\lambda_k}(y_k) &= \psi_{\lambda_k}(y_k) + (4k^2)^{-1}||y_k - x_k||^2 \geq \psi_{\lambda_k}(y_k) + (4k^2)^{-1}\epsilon_k^2 \\ &\geq \inf(\psi_{\lambda_k}) + (4k^2)^{-1}\epsilon_k^2 \geq \phi_{\lambda_k}(x_k) + (8k^2)^{-1}\epsilon_k^2.\end{aligned}$$

Combined with (2.8) this inequality implies that

$$(\phi_{\lambda_k} + h_k)(y_k) - (\phi_{\lambda_k} + h_k)(x_k) \geq \phi_{\lambda_k}(y_k) - \phi_{\lambda_k}(x_k) - (16k^2)^{-1}\epsilon_k^2 \geq (16k^2)^{-1}\epsilon_k^2.$$

This inequality contradicts (2.9). The contradiction we have reached proves that

$$||y_k - x_k|| < \epsilon_k. \tag{2.12}$$

It follows from (2.2), (2.3), (2.7), (2.8) and (2.9) that

$$\begin{aligned}f(y_k) \leq \psi_{\lambda_k}(y_k) \leq \phi_{\lambda_k}(y_k) &\leq (\phi_{\lambda_k} + h_k)(y_k) + 32^{-1}k^{-2}\epsilon_k^2 \\ &\leq \phi_{\lambda_k}(x_k) + h_k(x_k) + (32k^2)^{-1}\epsilon_k^2 \\ &\leq \phi_{\lambda_k}(x_k) + (16k^2)^{-1}\epsilon_k^2 \\ &= \psi_{\lambda_k}(x_k) + (16k^2)^{-1}\epsilon_k^2 \\ &\leq \inf(\psi_{\lambda_k}) + (4k^2)^{-1}\epsilon_k^2.\end{aligned} \tag{2.13}$$

By (2.12) and (2.6),

$$y_k \notin A. \tag{2.14}$$

In view of (2.13), (2.2) and (2.3),

$$f(y_k) \le 1 + \inf(\psi_{\lambda_k}; A) = \inf(f; A) + 1. \tag{2.15}$$

This inequality and (1.1) imply that the sequence $\{y_k\}_{k=1}^{\infty}$ is bounded

$$\sup\{||y_k|| : \ k = 1, 2, \dots\} < \infty. \tag{2.16}$$

By (2.14) in the case of Theorem 1.2 we obtain that

$$g(y_k) > c \quad \text{for all natural numbers} \quad k. \tag{2.17}$$

By (2.14) in the case of Theorem 1.1 we obtain that for each natural number k either $g(y_k) > c$ or $g(y_k) < c$. In the case of Theorem 1.1 extracting a subsequence and re-indexing we may assume that either $g(y_k) > c$ for all natural numbers k or

$$g(y_k) < c \quad \text{for all natural numbers} \quad k. \tag{2.18}$$

Define a function $\tilde{g}$ and a real number $\tilde{c}$ as follows. In the case of Theorem 1.2 put $\tilde{g} = g$, $\tilde{c} = c$. In the case of Theorem 1.1, if (2.17) holds then set $\tilde{g} = g$ and $\tilde{c} = c$, and if (2.18) is valid then put $\tilde{g} = -g$ and $\tilde{c} = -c$. Note that in all these cases we have

$$\tilde{g}(y_k) > \tilde{c} \quad \text{for all natural numbers} \quad k. \tag{2.19}$$

Let k be a natural number. It follows from (2.2), (2.3), the definition of $\tilde{g}$ and $\tilde{c}$, (2.19) and (2.13) that

$$\begin{aligned} f(y_k) + \lambda_k(\tilde{g}(y_k) - \tilde{c}) = \psi_{\lambda_k}(y_k) &\le \inf(\psi_{\lambda_k}) + 1 \\ &\le \inf(\psi_{\lambda_k}; A) + 1 \\ &\le \inf(f; A) + 1. \end{aligned} \tag{2.20}$$

Together with (2.19) and (2.4) this relation implies that

$$\begin{aligned} 0 < \tilde{g}(y_k) - \tilde{c} &\le \lambda_k^{-1}[\inf(f; A) + 1 - (f(y_k)] \\ &\le k^{-1}[\inf(f; A) + 1 - \inf(f)] \to 0 \text{ as } k \to \infty, \end{aligned} \tag{2.21}$$

$$\lim_{k \to \infty} \tilde{g}(y_k) = \tilde{c}. \tag{2.22}$$

In view of (2.22) and the definition of $\tilde{g}$ and $\tilde{c}$

$$c - \gamma \le g(y_k) \le c + \gamma \quad \text{for all sufficiently large natural numbers} \quad k. \tag{2.23}$$

By (2.23), (2.16) and (A1) extracting a subsequence and re-indexing we may assume without loss of generality that there exists

$$y_* = \lim_{k \to \infty} y_k \quad \text{in the norm topology.} \tag{2.24}$$

In view of (2.22) and (2.24)

$$g(y_*) = c \quad \text{and} \quad y_* \in A. \tag{2.25}$$

It follows from the lower semicontinuity of the function f, (2.24), (2.13), (2.2) and (2.3) that

$$\begin{aligned} f(y_*) = \liminf_{k\to\infty} f(y_k) &\le \limsup_{k\to\infty}[\inf(\psi_{\lambda_k}) + (4k^2)^{-1}\epsilon_k^2] \\ &= \limsup_{k\to\infty} \inf(\psi_{\lambda_k}) \\ &\le \limsup_{k\to\infty} \inf(\psi_{\lambda_k}; A) = \inf(f; A). \end{aligned}$$

Together with (2.25) this relation implies that

$$f(y_*) = \inf(f; A). \tag{2.26}$$

By (2.25), (2.26), (A2) and (A3) there exists $\bar{\Delta} > 0$ such that the following property holds:

(i) the restriction of g to $B^o(y_*, \bar{\Delta})$ belongs to $C^2(B^o(y_*, \bar{\Delta}); R^1)$;

(ii) the restriction of f to $B^o(y_*, \bar{\Delta})$ belongs to $C^2(B^o(y_*, \bar{\Delta}); R^1)$ and it is Lipschitz;

(iii) the mapping $f' : B^o(y_*, \bar{\Delta}) \to X$ is Lipschitz.

(iv) in the case of Theorem 1.1 $y_* \notin Cr(\tilde{g}+) \cup Cr(\tilde{g}-)$ and in the case of Theorem 1.2 $y_* \notin Cr(\tilde{g}+)$.

Relation (2.24) implies that there is a natural number k_0 such that

$$||y_k - y_*|| \le \bar{\Delta}/4 \quad \text{for all integers} \quad k \ge k_0. \tag{2.27}$$

Let $k \ge k_0$ be a natural number. By (2.19) there is an open neighborhood W of y_k in X with the norm topology such that

$$W \subset B^o(y_k, \bar{\Delta}/4) \quad \text{and} \quad \tilde{g}(z) > \tilde{c} \quad \text{for all} \quad z \in W. \tag{2.28}$$

By (2.7), (2.18), (2.2) and (2.3) for each $z \in W$

$$\begin{aligned} (\phi_{\lambda_k} + h_k)(z) &= \psi_{\lambda_k}(z) + (4k^2)^{-1}||z - x_k||^2 + h_k(z) \\ &= f(z) + \lambda_k(\tilde{g}(z) - \tilde{c}) + (4k^2)^{-1}||z - x_k||^2 + h_k(z). \end{aligned} \tag{2.29}$$

It follows from (2.27)-(2.29) and the properties (i) and (ii) that the function $\phi_{\lambda_k} + h_k$ is Frechet differentiable on the set W. In view of (2.9) and (2.29)

$$0 = (\phi_{\lambda_k} + h_k)'(y_k) = f'(y_k) + \lambda_k \tilde{g}'(y_k) + (4k^2)^{-1}2(y_k - x_k) + h_k'(y_k). \tag{2.30}$$

Together with (2.4), (2.12) and (2.8) this equality implies that

$$\begin{aligned} ||\tilde{g}'(y_k)|| &= \lambda_k^{-1}||f'(y_k) + (4k^2)^{-1}2(y_k - x_k) + h_k'(y_k)|| \\ &\le k^{-1}(||f'(y_k)|| + \epsilon_k + ||h_k'(y_k)||) \\ &\le k^{-1}(||f'(y_k)|| + 2). \end{aligned} \tag{2.31}$$

Combined with (2.24) and the property (i) this relation implies that

$$||g'(y_*)|| = \lim_{k\to\infty} ||g'(y_k)|| = \lim_{k\to\infty} ||\tilde{g}'(y_k)|| = 0 \text{ and } g'(y_*) = 0. \tag{2.32}$$

By the property (iii) there exist $\bar{L} > 0$ such that

$$||f'(z_1) - f'(z_2)|| \le \bar{L}||z_1 - z_2|| \quad \text{for each} \quad z_1, z_2 \in B^o(y_*, \bar{\Delta}). \tag{2.33}$$

Let $k \ge k_0$ be an integer. Since $h_k \in C^2(X; R^1)$ and the restriction of g to $B^o(y_*, \bar{\Delta})$ belongs to $C^2(B^o(y_*, \bar{\Delta}); R^1)$ it follows from (2.19) and (2.27) that there exists a positive number

$$r_k < \min\{\epsilon_k/2, \bar{\Delta}/8\} \tag{2.34}$$

such that

$$\tilde{g}(z) > \tilde{c} \quad \text{for each} \quad z \in B(y_k, r_k), \tag{2.35}$$

$$||h_k''(y_k) - h_k''(y_k + v)|| \le k^{-2} \quad \text{for each} \quad v \in B(0, r_k), \tag{2.36}$$

$$||g''(y_k + v) - g''(y_k)|| \le k^{-2}\lambda_k^{-1} \quad \text{for each} \quad v \in B(0, r_k). \tag{2.37}$$

By (2.7), (2.2), (2.3) and (2.35) for each $z \in B(y_k, r_k)$

$$(\phi_{\lambda_k} + h_k)(z) = f(z) + \lambda_k(\tilde{g}(z) - \tilde{c}) + (4k^2)^{-1}||z - x_k||^2 + h_k(z). \tag{2.38}$$

It follows from (2.23), (2.27) and (2.24) that for each $v \in B(0, r_k)$

$$||f'(y_k + v) - f'(y_k)|| \le \bar{L}||v||. \tag{2.39}$$

By (2.12) and (2.16) that there exists $M > 0$ such that

$$||y_k||, ||x_k|| \le M \quad \text{for all integers} \quad k \ge 1. \tag{2.40}$$

Let

$$u \in B(0, r_k) \setminus \{0\}. \tag{2.41}$$

The inclusion $h_k \in C^2(X; R^1)$ and Taylor theorem imply that there exists $t_1 \in [0, 1]$ such that

$$h_k(y_k + u) = h_k(y_k) + < h_k'(y_k), u > + < h_k''(y_k + t_1 u)u, u > /2. \tag{2.42}$$

Since the restriction of g to $B^o(y_*, \bar{\Delta})$ belongs to $C^2(B^o(y_*, \bar{\Delta}); R^1)$ it follows from the Taylor theorem that there exists $t_2 \in [0, 1]$ such that

$$g(y_k + u) = g(y_k) + < g'(y_k), u > + < g''(y_k + t_2 u)u, u > /2. \tag{2.43}$$

Since the restriction of f to $B^o(y_*, \bar{\Delta})$ belongs to $C^1(B^o(y_*, \bar{\Delta}); R^1)$ there exists $t_3 \in [0, 1]$ such that

$$f(y_k + u) = f(y_k) + < f'(y_k + t_3 u), u > . \tag{2.44}$$

Relations (2.42), (2.36) and (2.41) imply that

$$|h_k(y_k + u) - h_k(y_k) - < h_k'(y_k), u > -2^{-1} < h_k''(y_k)u, u > |$$

$$= 2^{-1}| < (h_k''(y_k + t_1 u) - h_k''(y_k))u, u > | \le k^{-2}||u||^2. \tag{2.45}$$

In view of (2.43), (2.37) and (2.41)

$$
\begin{aligned}
&|g(y_k+u)-g(y_k)+<g'(y_k),u>-<g''(y_k)u,u>/2|\\
&=2^{-1}|<(g''(y_k+t_2u)-g''(y_k))u,u>|\\
&\le ||u||^2k^{-2}\lambda_k^{-1}.
\end{aligned}
\tag{2.46}
$$

It follows from (2.44), (2.39) and (2.41) that

$$
\begin{aligned}
|f(y_k+u)-f(y_k)-&<f'(y_k),u>|\\
&=|<f'(y_k+t_3u)-f'(y_k),u>|\\
&\le ||f'(y_k+t_3u)-f'(y_k)||||u||\le \bar{L}t_3||u||^2.
\end{aligned}
\tag{2.47}
$$

By (2.9), (2.41), (2.38), (2.47), (2.46), (2.45) and (2.4),

$$
\begin{aligned}
0&<(\phi_{\lambda_k}+h_k)(y_k+u)-(\phi_{\lambda_k}+h_k)(y_k)\\
&=f(y_k+u)-f(y_k)+\lambda_k(\tilde{g}(y_k+u)-\tilde{g}(y_k))\\
&\quad+(4k^2)^{-1}[||y_k+u-x_k||^2-||y_k-x_k||^2]+h_k(y_k+u)-h_k(y_k)\\
&\le<f'(y_k),u>+\bar{L}||u||^2+\lambda_k[<\tilde{g}'(y_k),u>+<\tilde{g}''(y_k)u,u>2^{-1}]\\
&\quad+||u||^2k^{-2}+(4k^2)^{-1}[||u||^2+2<y_k-x_k,u>]+<h_k'(y_k),u>\\
&\quad+2^{-1}<h_k''(y_k)u,u>+k^{-2}||u||^2.
\end{aligned}
\tag{2.48}
$$

Set

$$
\begin{aligned}
F(u)=&<f'(y_k),u>+\bar{L}||u||^2+\lambda_k[<\tilde{g}'(y_k),u>\\
&+<\tilde{g}''(y_k)u,u>2^{-1}]+2||u||^2k^{-2}\\
&+(4k^2)^{-1}[||u||^2+2<y_k-x_k,u>]\\
&+<h_k'(y_k),u>+2^{-1}<h_k''(y_k)u,u>,\ u\in X.
\end{aligned}
\tag{2.49}
$$

Clearly $F\in C^2(X;R^1)$. It follows from (2.49) and (2.48) that

$$F(u)>F(0)\quad\text{for all}\quad u\in B(0,r_k)\setminus\{0\}.$$

The inequality above implies that

$$F'(0)=0 \text{ and } <F''(0)v,v>\ge 0\quad\text{for all}\quad v\in X. \tag{2.50}$$

Denote by I the identity operator $I:X\to X$ such that $Ix=x$ for all $x\in X$. By (2.49), (2.4) and (2.8)

$$F''(0)=2\bar{L}I+\lambda_k\tilde{g}''(y_k)+(9/2)k^{-2}I+h_k''(y_k). \tag{2.51}$$

Relations (2.50) and (2.51) imply that for each $v \in X$

$$\begin{aligned}
0 &\le < \lambda_k^{-1} F''(0)v, v > \\
&= 2\bar{L}\lambda_k^{-1}||v||^2 + < \tilde{g}''(y_k)v, v > +(9/2)k^{-2}\lambda_k^{-1}||v||^2 + \lambda_k^{-1} < h_k''(y_k)v, v > \\
&\le < \tilde{g}''(y_k)v, v > +2\bar{L}k^{-1}||v||^2 + (9/2)k^{-3}||v||^2 + k^{-1}| < h_k''(y_k)v, v > | \\
&\le < \tilde{g}''(y_k)v, v > +2\bar{L}k^{-1}||v||^2 + (9/2)k^{-3}||v||^2 + k^{-3}||v||^2 \\
&\to < \tilde{g}''(y_*)v, v >
\end{aligned}$$

as $k \to \infty$. Thus

$$< \tilde{g}''(y_*)v, v > \ge 0 \quad \text{for all} \quad v \in X. \tag{2.52}$$

Combining (2.25), (2.26) and (2.32) we obtain

$$g(y_*) = c,\ g'(y_*) = 0, \quad f(y_*) = \inf(f; A). \tag{2.53}$$

In the case of Theorem 1.1 relations (2.52) and (2.53) contradict (A2). In the case of Theorem 1.2 (2.52) and (2.53) contradict (A3). The contradiction we have reached proves that there exists $\Lambda_0 > 0$ such that property (P1) holds. This completes the proof of Theorems 1.1 and 1.2. □

References

[1] D. Boukari and A. V. Fiacco, *Survey of penalty, exact-penalty and multiplier methods from 1968 to 1993*, Optimization, **32** (1995), 301–334.

[2] J. V. Burke, *An exact penalization viewpoint of constrained optimization*, SIAM J. Control Optim., **29** (1991), 968–998.

[3] L. Cesari, *Optimization-Theory and Applications*, Springer-Verlag, New York, 1983.

[4] F. H. Clarke, *Optimization and Nonsmooth Analysis*, Willey Interscience, 1983.

[5] V. F. Demyanov, *Constrained problems of calculus of variations via penalization technique*, Equilibrium problems and variational models (Erice, 2000), Nonconvex Optim. Appl., Kluwer Acad. Publ., Norwel, MA **68**, 2003, 79–108.

[6] V. F. Demyanov, *Exact penalty functions and problems of the calculus of variations*, Avtomat. i Telemekh. (2004), 136–147.

[7] R. Deville, R. Godefroy and V. Zizler, *Smoothness and Renorming in Banach Spaces*, Longman, 1993.

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

[9] I. I. Eremin, *The penalty method in convex programming*, Soviet Math. Dokl., **8** (1966), 459–462.

[10] J.-B. Hiriart-Urruty and C. Lemarechal, *Convex Analysis and Minimization Algorithms*, Springer, Berlin, 1993.

[11] Z. Q. Luo, J.-S. Pang and D. Ralph, *Mathematical Programs with Equilibrium Constraints*, Cambridge University Press, 1996.

[12] O. L. Mangasarian and J.-S. Pang, *Exact penalty functions for mathematical programs with linear complementary constraints*, Optimization, **42** (1997), 1–8.

[13] B. S. Mordukhovich, *Penalty functions and necessary conditions for the extremum in nonsmooth and nonconvex optimization problems*, Uspekhi Math. Nauk, **36** (1981), 215–216.

[14] R. Palais, *Lusternik-Schnirelman theory of Banach manifolds*, Topology, **5** (1966), 115–132.

[15] 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, 418–425.

[16] A. M. Rubinov, B. M. Glover and X. Q. Yang, *Decreasing functions with applications to penalization*, SIAM J. Optim., **10** (1999), 289–313.

[17] A. M. Rubinov, X. Q. Yang and A. M. Bagirov, *Penalty functions with a small penalty parameter*, Optim. Methods Softw., **17** (2002), 931–964.

[18] W. I. Zangwill, *Nonlinear programming via penalty functions*, Management Sci., **13** (1967), 344–358.

[19] A. J. Zaslavski, *A sufficient condition for exact penalty in constrained optimization*, SIAM Journal on Optimization, **16** (2005), 250–262.

[20] A. J. Zaslavski, *Existence of exact penalty for constrained optimization problems in Hilbert spaces*, Nonlinear Analysis, accepted.