

Smooth Exact Penalty Functions: A General Approach

M.V. Dolgopolik

January 30, 2018

Abstract

In this article we present a new perspective on the smooth exact penalty function proposed by Huyer and Neumaier that is becoming more and more popular tool for solving constrained optimization problems. Our approach to Huyer and Neumaier's exact penalty function allows one to apply previously unused tools (namely, parametric optimization) to the study of the exactness of this function. We give a new simple proof of the local exactness of Huyer and Neumaier's penalty function that significantly generalizes all similar results existing in the literature. We also obtain new necessary and sufficient conditions for the global exactness of this penalty function.

1 Introduction

The method of exact penalty functions [8, 4, 7] is a powerful tool for solving various constrained optimization problems. However, as it is well-known, exact penalty functions are usually nonsmooth, even in the case when the original problem is smooth. This obstacle makes it impossible to apply (without some transformation of the problem) well-developed and extensively studied methods of smooth unconstrained optimization to minimization of an exact penalty function.

Huyer and Neumaier in [9] proposed a new approach to exact penalization that allows one to overcome nonsmoothness of exact penalty functions. Namely, let the original constrained optimization problem have the form

$$\min f(x) \quad \text{subject to} \quad F(x) = 0, \quad x \in [\underline{x}, \bar{x}], \quad (1)$$

where $f: \mathbb{R}^n \rightarrow \mathbb{R}$ and $F: \mathbb{R}^n \rightarrow \mathbb{R}^m$ are smooth functions, $\underline{x}, \bar{x} \in \mathbb{R}^n$ are given vectors, and

$$[\underline{x}, \bar{x}] = \{x = (x_1, \dots, x_n) \in \mathbb{R}^n \mid \underline{x}_i \leq x_i \leq \bar{x}_i \quad \forall i \in \{1, \dots, n\}\}.$$

The new approach consists in the introduction of an additional variable $\varepsilon \geq 0$ in the following way. Choose $w \in \mathbb{R}^m$, and note that problem (1) is equivalent to the problem

$$\min_{x, \varepsilon} f(x) \quad \text{subject to} \quad F(x) = \varepsilon w, \quad \varepsilon = 0, \quad x \in [\underline{x}, \bar{x}]. \quad (2)$$

Then one defines the new “smooth” penalty function for the augmented problem (2) as follows

$$F_\lambda(x, \varepsilon) = \begin{cases} f(x), & \text{if } \varepsilon = \Delta(x, \varepsilon) = 0, \\ f(x) + \frac{1}{2\varepsilon} \frac{\Delta(x, \varepsilon)}{1 - q\Delta(x, \varepsilon)} + \lambda\beta(\varepsilon), & \text{if } \varepsilon > 0, \Delta(x, \varepsilon) < q^{-1}, \\ +\infty, & \text{otherwise.} \end{cases} \quad (3)$$

where $\lambda \geq 0$ is the penalty parameter, $\Delta(x, \varepsilon) = \|F(x) - \varepsilon w\|^2$ is the constraint violation measure, $\beta: [0, \bar{\varepsilon}] \rightarrow [0, +\infty)$ with $\beta(0) = 0$ is the penalty term, $q > 0$ and $\bar{\varepsilon} > 0$ are some prespecified thresholds. Finally, one replaces the augmented problem (2) with the penalized problem

$$\min_{x, \varepsilon} F_\lambda(x, \varepsilon) \quad \text{subject to} \quad (x, \varepsilon) \in [\underline{x}, \bar{x}] \times [0, \bar{\varepsilon}]. \quad (4)$$

Observe that the penalty function $F_\lambda(x, \varepsilon)$ is smooth for any $\varepsilon \in (0, \bar{\varepsilon})$ and x such that $0 < \Delta(x, \varepsilon) < q^{-1}$ provided the function β is smooth on $(0, \bar{\varepsilon})$. Furthermore, it was proven in [9] that under a standard assumption (namely, constraint qualification) the penalty function $F_\lambda(x, \varepsilon)$ is locally exact. In other words, (x^*, ε^*) is a point of local minimum of problem (4) if and only if $\varepsilon^* = 0$ and x^* is a point of local minimum of problem (1). Consequently, one can apply methods of smooth unconstrained minimization to penalized problem (4) in order to find a solution of initial constrained optimization problem (1).

Later on, Huyer and Neumaier’s approach was generalized [18, 2] and successfully applied to various constrained optimization problems [15, 13], including some optimal control problems [12, 10, 14]. However, it should be noted that the existing proofs of the exactness of the smooth penalty function (3) and its various generalizations are quite complicated, and overburdened by technical details that overshadow the understanding of the technique of smooth exact penalty functions. Also, the question of when problem (1) is actually equivalent to problem (4) in terms of globally optimal solutions (in this case the penalty function $F_\lambda(x, \varepsilon)$ is called exact) has not been discussed in the literature.

The aim of this article is to present a new perspective on the method of smooth exact penalty functions proposed by Huyer and Neumaier. This perspective allows one to apply previously unused tools (namely, parametric optimization) to the study and construction of smooth exact penalty functions. It also helped us to essentially simplify the proof of exactness of these functions. Another aim of this articles is to provide first necessary and sufficient conditions for Huyer and Neumaier’s penalty function to be (globally) exact.

The paper is organised as follows. In Section 2 we describe a new approach to smooth exact penalty functions. Some general results that draw a connection between the exactness of the new penalty function and some properties of a perturbed optimization problem are presented in Section 3. In Section 4 we give a new simple proof of the local exactness of the new penalty function that significantly generalizes all results on the local exactness of Huyer and Neumaier’s penalty functions existing in the literature. We also provide new simple sufficient conditions for this penalty function to be globally exact.

2 How to Construct a Smooth Exact Penalty Function?

Two main approaches are usually used for the study of exact penalty functions: the direct approach that is based on the use of error bounds and metric regularity, and the indirect one that relies on the analysis of a perturbed optimization problem. In the indirect approach [5, 6, 3, 17], a perturbation of the initial problem is introduced as a tool for the study of a penalty function that has already been defined. However, one can introduce perturbation in order to construct a penalty function.

Namely, define the perturbed objective function for problem (1) as follows:

$$g(x, \mu) = \begin{cases} f(x), & \text{if } \mu = \Delta(x, \mu) = 0, \\ f(x) + \frac{1}{\mu} \frac{\Delta(x, \mu)}{1 - q\Delta(x, \mu)}, & \text{if } \mu > 0, \Delta(x, \mu) < q^{-1}, \\ +\infty, & \text{otherwise,} \end{cases}$$

where $\mu \geq 0$ is a perturbation parameter. Consider the perturbed optimization problem

$$\min_x g(x, \mu) \quad \text{subject to} \quad x \in [\underline{x}, \bar{x}]. \quad (5)$$

It is clear that the problem above with $\mu = 0$ is equivalent to problem (1). Moreover, the penalization of the constraint $F(x) = 0$ is achieved via the introduced perturbation. As a second step, note that the perturbed problem is equivalent to the problem

$$\min_{x, \mu} g(x, \mu) \quad \text{subject to} \quad x \in [\underline{x}, \bar{x}], \quad \mu = 0.$$

Finally, introduce a penalty function for the problem above that penalizes only the constraint on the perturbation parameter, i.e. $\mu = 0$. This penalty function has the form

$$F_\lambda(x, \mu) = g(x, \mu) + \lambda\beta(\mu), \quad (6)$$

and it is smooth for any $\mu \in (0, \bar{\epsilon})$ and $x \in \mathbb{R}^n$ such that $0 < \Delta(x, \mu) < q^{-1}$. Thus, the fact that the nonlinear constraints are taken into account via perturbation (not penalization), and the fact that the penalty term is constructed only for a simple one dimensional constraint on the perturbation parameter helped us to avoid nonsmoothness that usually arises due to the restrictive requirements on the penalty term of an exact penalty function.

In the following section, we develop the approach discussed above in the general case, and demonstrate that the exactness of penalty function (6) is directly connected with some properties of perturbed problem (5).

3 Exact Penalty Function for a Perturbed Optimization Problem

Let X be a topological space, $f: X \rightarrow \mathbb{R} \cup \{+\infty\}$ be a given function, and $M, A \subset X$ be nonempty sets such that $M \cap A \neq \emptyset$. Hereafter, we study the following optimization problem:

$$\min f(x) \quad \text{subject to} \quad x \in M, \quad x \in A. \quad (\mathcal{P})$$

Denote by $\Omega = M \cap A$ the set of feasible points of this problem. Denote also $\mathbb{R}_+ = [0, +\infty)$ and $\text{dom } f = \{x \in X \mid f(x) < +\infty\}$. We suppose that f is bounded below on Ω .

Introduce a metric space of perturbation parameters (P, d) , and a perturbed objective function $g: X \times P \rightarrow \mathbb{R} \cup \{+\infty\}$ such that there exists $\mu_0 \in P$ for which the following conditions are satisfied:

1. $g(x, \mu_0) = f(x)$ for any $x \in \Omega$ (**consistency condition**);
2. $\arg \min_{x \in A} g(x, \mu_0) = \arg \min_{x \in \Omega} f(x)$ (**exact penalization condition**).

With the use of the exact penalization condition one gets that the problem (\mathcal{P}) is equivalent (in terms of globally optimal solutions) to the following optimization problem:

$$\min_{x, \mu} g(x, \mu) \quad \text{subject to} \quad x \in A, \quad \mu = \mu_0. \quad (7)$$

Furthermore, the consistency condition guarantees that if (x_0, μ_0) with $x_0 \in \Omega$ is a point of local minimum of the above problem, then x_0 is a point of local minimum of the problem (\mathcal{P}) .

We apply the exact exact penalization technique to treat the problem (7). Namely, choose a function $\beta: \mathbb{R}_+ \rightarrow \mathbb{R}_+ \cup \{+\infty\}$ such that $\beta(t) = 0$ iff $t = 0$. For any $\lambda \geq 0$ define the *penalty function*

$$F_\lambda(x, \mu) = g(x, \mu) + \lambda \beta(d(\mu, \mu_0))$$

and consider the following penalized problem

$$\min_{x, \mu} F_\lambda(x, \mu) \quad \text{subject to} \quad x \in A. \quad (8)$$

Observe that the function $\lambda \rightarrow F_\lambda(x, \mu)$ is non-decreasing.

Our aim is to study a relation between local/global minimizers of the initial problem (\mathcal{P}) and local/global minimizers of the problem (8) in the context of the theory of exact penalty functions. To this end, recall the definition of exact penalty function.

Definition 1. Let $x^* \in \text{dom } f$ be a point of local minimum of the problem (\mathcal{P}) . The penalty function F_λ is called (locally) *exact* at the point x^* (or, to be more precise, at the point (x^*, μ_0)) if there exists $\lambda \geq 0$ such that (x^*, μ_0) is a point of local minimum of the problem (8). The greatest lower bound of all such λ is denoted by $\lambda(x^*)$.

Definition 2. The penalty function F_λ is said to be (globally) *exact*, if there exists $\lambda \geq 0$ such that F_λ attains a global minimum on the set $A \times P$, and if $(x^*, \mu^*) \in A \times P$ is a globally optimal solution of the problem (8), then $\mu^* = \mu_0$. The greatest lower bound of all such $\lambda \geq 0$ is denoted by $\lambda^*(g, \beta)$.

Remark 1. (i) Note that if $(x^*, \mu^*) \in A \times P$ is a globally optimal solution of the problem (8) with $\mu^* = \mu_0$, then x^* is a globally optimal solution of the problem (\mathcal{P}) by virtue of the exact penalization condition on the function $g(x, \mu)$, and the fact that $F_\lambda(x, \mu_0) = g(x, \mu_0)$. Thus, the penalty function F_λ is globally exact iff there exists $\lambda \geq 0$ such that the problem (8) is equivalent to the problem (\mathcal{P}) in terms of globally optimal solutions.

(ii) It is easy to verify that if F_λ is globally exact, then for any $\lambda > \lambda^*(g, \beta)$ the function F_λ attains a global minimum on the set $A \times P$, and if $(x^*, \mu^*) \in A \times P$ is a globally optimal solution of the problem (8), then $\mu^* = \mu_0$.

Our aim is to show that the exactness of the penalty function F_λ is closely related to some properties of the perturbed optimization problem

$$\min_x g(x, \mu) \quad \text{subject to} \quad x \in A \quad (\mathcal{P}_\mu).$$

Denote by $h(\mu) = \inf_{x \in A} g(x, \mu)$ the optimal value function of this problem.

Recall that the problem $(\mathcal{P}_{\bar{\mu}})$, $\bar{\mu} \in P$, is said to be β -calm at a point $x^* \in A$ if there exist $\lambda \geq 0$, $r > 0$, and a neighbourhood U of x^* such that

$$g(x, \mu) - g(x^*, \bar{\mu}) \geq -\lambda\beta(d(\mu, \bar{\mu})) \quad \forall x \in U \cap A \quad \forall \mu \in B(\bar{\mu}, r),$$

where $B(\bar{\mu}, r) = \{\mu \in P \mid d(\mu, \bar{\mu}) \leq r\}$.

Remark 2. If $\beta(t) \equiv t$, then the concept of β -calmness coincides with the well-known concept of calmness of a perturbed optimization problem [5, 6, 3, 17].

The following propositions describes a connection between the exactness of the penalty function F_λ and the calmness of the perturbed problem (\mathcal{P}_μ) (cf. an analogous result for classical exact penalty functions in [3]).

Proposition 1. *Let $x^* \in \text{dom } f$ be a point of local minimum of the problem (\mathcal{P}) . Then the penalty function F_λ is exact at x^* if and only if the problem (\mathcal{P}_{μ_0}) is β -calm at x^* .*

Proof. The validity of the proposition follows from the fact that the inequality

$$g(x, \mu) - g(x^*, \mu_0) \geq -\lambda\beta(d(\mu, \mu_0)) \quad \forall x \in U \cap A \quad \forall \mu \in B(\bar{\mu}, r),$$

holds true for some $\lambda \geq 0$, $r \geq 0$, and a neighbourhood U of x^* iff for any $x \in U \cap A$ and $\mu \in B(\mu_0, r)$ one has

$$F_\lambda(x, \mu) = g(x, \mu) + \beta(d(\mu, \mu_0)) \geq g(x, \mu_0) = F_\lambda(x, \mu_0),$$

i.e. iff (x^*, μ_0) is a point of local minimum of F_λ on the set A . \square

Let us turn to the study of global exactness. We need an auxiliary definition. The optimal value function h is called β -calm from below at a point $\bar{\mu} \in P$, if

$$\liminf_{\mu \rightarrow \mu_0} \frac{h(\mu) - h(\bar{\mu})}{\beta(d(\mu, \bar{\mu}))} > -\infty$$

(cf. the definition of calmness from below in [16]).

Theorem 1. *Let β be strictly increasing. For the penalty function F_λ to be globally exact it is necessary and sufficient that the following assumptions hold true:*

1. *there exists a globally optimal solution of the problem (\mathcal{P}) ;*
2. *the optimal value function $h(\mu)$ is β -calm from below at μ_0 ;*
3. *there exists $\lambda_0 \geq 0$ such that F_{λ_0} is bounded below on the set $A \times P$.*

Proof. Necessity. Since the penalty function F_λ is globally exact, then it attains a global minimum on the set $A \times P$ for some $\lambda \geq 0$. Therefore, in particular, F_λ is bounded below on $A \times P$.

Fix a sufficiently large $\lambda \geq 0$, and a global minimizer (x^*, μ^*) of F_λ on the set $A \times P$. Due to the exactness of F_λ one has $\mu^* = \mu_0$. Therefore, as it was mentioned above, x^* is a globally optimal solution of the problem (\mathcal{P}) . Thus, the first assumption is valid as well.

Observe that

$$g(x^*, \mu_0) \geq h(\mu_0) := \inf_{x \in A} g(x, \mu_0) \geq \inf_{(x, \mu) \in A \times P} F_\lambda(x, \mu) = g(x^*, \mu_0).$$

Hence $h(\mu_0) = g(x^*, \mu_0)$. The exactness of F_λ implies that for all $x \in A$ and $\mu \in P$ the one has

$$g(x, \mu) + \lambda\beta(d(\mu, \mu_0)) = F_\lambda(x, \mu) \geq F_\lambda(x^*, \mu_0) = g(x^*, \mu_0) = h(\mu_0)$$

or, equivalently,

$$g(x, \mu) - h(\mu_0) \geq -\lambda\beta(d(\mu, \mu_0)) \quad \forall x \in A \quad \forall \mu \in P.$$

Since the inequality above holds true for all $x \in A$, one obtains that

$$h(\mu) - h(\mu_0) \geq -\lambda\beta(d(\mu, \mu_0)) \quad \forall \mu \in P,$$

which implies the β -calmness from below of h at μ_0 .

Sufficiency. Let x^* be a globally optimal solution of the problem (\mathcal{P}) . Then taking into account the exact penalization condition on the function $g(x, \mu)$ one gets that $h(\mu_0) = g(x^*, \mu_0)$, i.e. x^* is a point of global minimum of the function $x \rightarrow g(x, \mu_0)$.

From the fact that the optimal value function h is β -calm from below at μ_0 it follows that there exist $\lambda_1 \geq 0$ and $\delta > 0$ such that

$$h(\mu) - h(\mu_0) \geq -\lambda_1\beta(d(\mu, \mu_0)) \quad \forall \mu \in B(\mu_0, \delta).$$

Hence for any $x \in A$ one has

$$g(x, \mu) - g(x^*, \mu_0) \geq h(\mu) - h(\mu_0) \geq -\lambda_1\beta(d(\mu, \mu_0)) \quad \forall \mu \in B(\mu_0, \delta)$$

or, equivalently, for all $(x, \mu) \in A \times B(\mu_0, \delta)$ one has

$$F_{\lambda_1}(x, \mu) = g(x, \mu) + \lambda_1\beta(d(\mu, \mu_0)) \geq g(x^*, \mu_0) = F_{\lambda_1}(x^*, \mu_0).$$

On the other hand, if $x \in A$ and $\mu \notin B(\mu_0, \delta)$, then for any $\lambda \geq \lambda_2$, where

$$\lambda_2 = \lambda_0 + \frac{g(x^*, \mu_0) - c}{\beta(\delta)}, \quad c = \inf_{(x, \mu) \in A \times P} F_{\lambda_0}(x, \mu) > -\infty,$$

one has

$$\begin{aligned} F_\lambda(x, \mu) &= F_{\lambda_0}(x, \mu) + (\lambda - \lambda_0)\beta(d(\mu, \mu_0)) \geq \\ &\geq c + (\lambda - \lambda_0)\beta(\delta) \geq g(x^*, \mu_0) = F_\lambda(x^*, \mu_0). \end{aligned}$$

Thus, for any $\lambda \geq \bar{\lambda} := \max\{\lambda_1, \lambda_2\}$ one has

$$F_\lambda(x, \mu) \geq F_\lambda(x^*, \mu_0) \quad \forall (x, \mu) \in A \times P$$

or, in other words, the penalty function F_λ attains a global minimum on $A \times P$ at the point (x^*, μ_0) . Let $(\bar{x}, \bar{\mu}) \in A \times P$ be a different global minimizer of F_λ on $A \times P$. Let us show that $\bar{\mu} = \mu_0$, provided $\lambda > \bar{\lambda}$, then one concludes that the penalty function F_λ is globally exact.

Indeed, for any $\lambda > \bar{\lambda}$, $x \in A$ and $\mu \neq \mu_0$ one has

$$F_\lambda(x^*, \mu_0) = F_{\bar{\lambda}}(x^*, \mu_0) \leq F_{\bar{\lambda}}(x, \mu) < F_\lambda(x, \mu),$$

since $\beta(d(\mu, \mu_0)) > 0$. Hence $\bar{\mu} = \mu_0$ by virtue of the fact that $(\bar{x}, \bar{\mu})$ is a global minimizer of F_λ on $A \times P$. \square

Let us also point out a connection between the calmness of the optimal value function h and the calmness of the perturbed problem (\mathcal{P}_{μ_0}) at globally optimal solutions of the problem (\mathcal{P}) in the case when the set A is compact.

Theorem 2. *Let the set A be compact, and the function $g(x, \mu)$ be lower semi-continuous (l.s.c.) on $A \times B(\mu_0, r)$ for some $r > 0$. Then for the optimal value function h to be β -calm from below at μ_0 it is necessary and sufficient that the problem (\mathcal{P}_{μ_0}) is β -calm at every globally optimal solution of the problem (\mathcal{P}) .*

Proof. Necessity. Fix a globally optimal solution x^* of the problem (\mathcal{P}) . From the definition of β -calmness it follows that there exist $\lambda \geq 0$ and $r > 0$ such that

$$h(\mu) - h(\mu_0) \geq -\lambda\beta(d(\mu, \mu_0)) \quad \forall \mu \in B(\mu_0, r).$$

From the fact that x^* is a globally optimal solution of the problem (\mathcal{P}) it follows that $h(\mu_0) = g(x^*, \mu_0)$ due to the exact penalization condition on $g(x, \mu)$. Therefore

$$g(x, \mu) - g(x^*, \mu_0) \geq h(\mu) - h(\mu_0) \geq -\lambda\beta(d(\mu, \mu_0)) \quad \forall (x, \mu) \in A \times B(\mu_0, r).$$

Thus, the problem (\mathcal{P}_{μ_0}) is β -calm at x^* .

Sufficiency. Taking into account the facts that A is compact, and $g(x, \mu)$ is l.s.c. one gets that the function $g(\cdot, \mu_0)$ attains a global minimum on the set A , and the set A^* of all points of global minimum of $g(\cdot, \mu_0)$ on A is compact. Furthermore, from the exact penalization condition it follows that A^* is also the set of all globally optimal solutions of the problem (\mathcal{P}) . Hence the problem (\mathcal{P}_{μ_0}) is β -calm at every $x^* \in A^*$. Therefore for any $x^* \in A^*$ there exist $\lambda(x^*) \geq 0$, $r(x^*) \geq 0$ and a neighbourhood $U(x^*)$ of x^* such that

$$g(x, \mu) - g(x^*, \mu_0) \geq -\lambda(x^*)\beta(d(\mu, \mu_0)) \quad \forall (x, \mu) \in U(x^*) \times B(\mu_0, r(x^*)).$$

Applying the compactness of A^* one obtains that there exist $x_1^*, x_2^*, \dots, x_n^* \in A^*$ such that $A^* \subset \bigcup_{k=1}^n U(x_k^*)$. Denote

$$U = \bigcup_{k=1}^n U(x_k^*), \quad \bar{\lambda} = \max_{k \in 1:n} \lambda(x_k^*), \quad \bar{r} = \min_{k \in 1:n} r(x_k^*).$$

Then for any $x \in U$ one has

$$g(x, \mu) - h(\mu_0) \geq -\bar{\lambda}\beta(d(\mu, \mu_0)) \quad \forall \mu \in B(\mu_0, \bar{r}), \quad (9)$$

due to the fact that $g(x^*, \mu_0) = h(\mu_0) = \inf_{x \in A} g(x, \mu_0)$ for any $x^* \in A^*$.

Set $K = A \setminus U$. Since $A^* \subset U$, for any $x \in K$ one has $g(x, \mu_0) > h(\mu_0)$ (recall that A^* is the set of all global minimizers of $g(\cdot, \mu_0)$ on A). Consequently, applying the lower semicontinuity of the function $g(x, \mu)$ one gets that for any $x \in K$ there exist $\delta(x) > 0$ and a neighbourhood $V(x)$ of x such that $g(y, \mu) > h(\mu_0)$ for all $(y, \mu) \in V(x) \times B(\mu_0, \delta(x))$. Observe that from the facts that U is open and A is compact it follows that $K = A \setminus U$ is also compact. Applying this fact and the inequality above it is easy to verify that there exists $\delta > 0$ such that $g(x, \mu) > h(\mu_0)$ for all $(x, \mu) \in K \times B(\mu_0, \delta)$. Therefore taking into account (9) one gets that

$$g(x, \mu) - h(\mu_0) \geq -\bar{\lambda}\beta(d(\mu, \mu_0)) \quad \forall (x, \mu) \in A \times B(\mu_0, \min\{\delta, \bar{r}\}),$$

which yields

$$h(\mu) - h(\mu_0) \geq -\bar{\lambda}\beta(d(\mu, \mu_0)) \quad \forall \mu \in B(\mu_0, \min\{\delta, \bar{r}\}).$$

Thus, the optimal value function h is β -calm from below at μ_0 . \square

Combining Theorems 1 and 2, and Proposition 1 one obtains that the following result holds true.

Corollary 1. *Let A be compact, $g(x, \mu)$ be l.s.c. on $A \times B(\mu_0, r)$ for some $r > 0$, and the function β be strictly increasing. Then the penalty function F_λ is globally exact if and only if it is exact at every globally optimal solution of the problem (\mathcal{P}) , and there exists $\lambda_0 \geq 0$ such that $F_{\lambda_0}(x, \mu)$ is bounded below on $A \times P$.*

4 Smooth Exact Penalty Functions

Let us apply the theory developed in the previous section to the study of Huyer and Neumaier's exact penalty functions. Let X and Y be metric spaces, $A \subset X$ be a nonempty set, and $\Phi: X \rightrightarrows Y$ be a given set-valued mapping with closed images. For any subset $C \subset X$ and $x_0 \in X$ denote by $d(x_0, C) = \inf_{x \in C} d(x_0, x)$ the distance between C and x_0 . For any $y \in Y$ denote, as usual, $\Phi^{-1}(y) = \{x \in X \mid y \in \Phi(x)\}$.

Fix an element $y_0 \in Y$, and consider the following optimization problem:

$$\min f(x) \quad \text{subject to} \quad y_0 \in \Phi(x), \quad x \in A. \quad (10)$$

Note that the set of feasible points of this problem has the form $\Omega = \Phi^{-1}(y_0) \cap A$.

Following the general technique proposed above and the method of smooth exact penalty functions [18], define $P = \mathbb{R}_+$, fix a non-decreasing function $\phi: \mathbb{R}_+ \cup \{+\infty\} \rightarrow \mathbb{R}_+ \cup \{+\infty\}$ such that $\phi(t) = 0$ iff $t = 0$, and introduce the perturbed objective function

$$g(x, \mu) = \begin{cases} f(x), & \text{if } x \in \Omega, \mu = 0, \\ +\infty, & \text{if } x \notin \Omega, \mu = 0, \\ f(x) + \frac{1}{\mu}\phi(d(y_0, \Phi(x))^2), & \text{if } \mu > 0. \end{cases}$$

Clearly, the function $g(x, \mu)$ satisfies the consistency condition and the exact penalization condition with $\mu_0 = 0$.

Introduce the penalty function

$$F_\lambda(x, \mu) = g(x, \mu) + \lambda\beta(\mu),$$

where $\beta: \mathbb{R}_+ \rightarrow \mathbb{R}_+ \cup \{+\infty\}$ is a non-decreasing function such that $\beta(\mu) = 0$ iff $\mu = 0$. Let us obtain sufficient conditions for $F_\lambda(x, \mu)$ to be exact. In order to formulate these conditions, recall that a set valued mapping Φ is said to be *metrically subregular* with respect to the set A with constant $a > 0$ at a point $(\bar{x}, \bar{y}) \in X \times Y$ with $\bar{y} \in \Phi(\bar{x})$ and $\bar{x} \in A$, if there exists a neighbourhood U of \bar{x} such that

$$d(\Phi(x), \bar{y}) \geq ad(x, \Phi^{-1}(\bar{y}) \cap A) \quad \forall x \in U \cap A.$$

Thus, Φ is metrically subregular with respect to the set A iff the restriction of Φ to A is metrically subregular in the usual sense. See [1, 11] and the references therein for the extensive study of metric subregularity.

Theorem 3. *Let $x^* \in \text{dom } f$ be a point of local minimum of the problem (10), the function f be Lipschitz continuous near x^* , and the set-valued mapping Φ be metrically subregular with respect to the set A with constant $a > 0$ at (x^*, y_0) . Suppose also that the following assumptions are satisfied:*

1. *there exist $\phi_0 > 0$ and $t_0 > 0$ such that $\phi(t) \geq \phi_0 t$ for any $t \in [0, t_0]$;*
2. *there exist $\beta_0 > 0$ and $\bar{\mu} > 0$ such that $\beta(\mu) \geq \beta_0 \mu$ for any $\mu \in [0, \bar{\mu}]$.*

Then the penalty function $F_\lambda(x, \mu)$ is exact at x^ . Moreover, one has*

$$\lambda(x^*) \leq \frac{L^2}{4\phi_0\beta_0a^2}, \quad (11)$$

where $L \geq 0$ is a Lipschitz constant of f near x^ .*

Proof. Since x^* is a point of local minimum of the problem (10), there exists $\rho > 0$ such that $f(x) \geq f(x^*)$ for any $x \in B(x^*, \rho) \cap \Omega$. Suppose, for a moment, that there exists $\delta \in (0, \rho)$ such that

$$f(x) - f(x^*) \geq -Ld(x, \Omega) \quad \forall x \in B(x^*, \delta) \setminus \Omega, \quad (12)$$

where $L \geq 0$ is a Lipschitz constant of f near x^* . Then applying the metric subregularity of Φ with respect to A , and the fact that the function ϕ is non-decreasing, one gets that for any $\mu > 0$ and $x \in B(x^*, r) \cap A$, where $r = \min\{\delta, t_0\}$, the following inequalities hold true:

$$\begin{aligned} g(x, \mu) - g(x^*, 0) &= f(x) - f(x^*) + \frac{1}{\mu}\phi(d(y_0, \Phi(x)))^2 \geq \\ &\geq -Ld(x, \Omega) + \frac{\phi_0a^2}{\mu}d(x, \Phi^{-1}(y_0) \cap A)^2 = -Ld(x, \Omega) + \frac{\phi_0a^2}{\mu}d(x, \Omega)^2. \end{aligned}$$

Note that the function $h(t) = -Lt + \phi_0a^2t^2/\mu$ attains a global minimum at the point $\mu L/2\phi_0a^2$, and

$$h\left(\frac{\mu L}{2\phi_0a^2}\right) = -\frac{L^2}{4\phi_0a^2}\mu.$$

Hence for any $\mu > 0$ and $x \in B(x^*, r) \cap A$ one has

$$g(x, \mu) - g(x^*, 0) \geq -\frac{L^2}{4\phi_0 a^2} \mu. \quad (13)$$

On the other hand, if $x \in B(x^*, r) \cap A$ and $\mu = 0$, then either $x \notin \Omega$ and $g(x, \mu) = +\infty \geq g(x^*, 0)$ or $x \in \Omega$ and $g(x, \mu) = f(x) \geq f(x^*) = g(x^*, 0)$ (recall that $r \leq \delta < \rho$). Therefore the inequality (13) is satisfied for any $x \in B(x^*, r) \cap A$ and $\mu \geq 0$, which yields that

$$F_\lambda(x, \mu) = g(x, \mu) + \lambda\beta(\mu) \geq g(x, \mu) + \lambda\beta_0\mu \geq g(x^*, 0) = F_\lambda(x^*, 0)$$

for all $(x, \mu) \in B(x^*, r) \times [0, \bar{\mu}]$, and for any $\lambda \geq L^2/4\phi_0\beta_0a^2$. Thus, the penalty function $F_\lambda(x, \mu)$ is exact at x^* , and (11) holds true.

It remains to show that inequality (12) is valid for some $\delta > 0$. Indeed, fix $x \in B(x^*, \rho/2) \setminus \Omega$. By the definition of the distance between a point and a set there exists a sequence $\{x_n\} \subset \Omega$ such that $d(x, x_n) \rightarrow d(x, \Omega)$ as $n \rightarrow \infty$. Moreover, without loss of generality one can suppose that $d(x, x_n) \leq \rho/2$ for any $n \in \mathbb{N}$, since $d(x, x^*) \leq \rho/2$ and $x^* \in \Omega$. Consequently, one has

$$d(x_n, x^*) \leq d(x_n, x) + d(x, x^*) \leq \frac{\rho}{2} + \frac{\rho}{2} = \rho,$$

which implies that $f(x_n) \geq f(x^*)$. Therefore applying the Lipschitz continuity of f near x^* one obtains that for any $n \in \mathbb{N}$ the following inequality holds true

$$f(x) - f(x^*) = f(x) - f(x_n) + f(x_n) - f(x^*) \geq f(x) - f(x_n) \geq -Ld(x, x_n).$$

Passing to the limit as $n \rightarrow \infty$ one obtains the desired result. \square

Applying Corollary 1, and the theorem above one can easily obtain sufficient conditions for the penalty function F_λ to be globally exact.

Theorem 4. *Let the set A be compact. Suppose that the following assumptions are satisfied:*

1. *f is l.s.c. on A , and locally Lipschitz continuous near globally optimal solutions of the problem (\mathcal{P}) ;*
2. *Φ is metrically subregular with respect to the set A at (x^*, y_0) for any globally optimal solution x^* of the problem (\mathcal{P}) ;*
3. *the mapping $x \rightarrow d(y_0, \Phi(x))$ is continuous on A ;*
4. *ϕ is l.s.c., and there exist $\phi_0 > 0$ and $t_0 > 0$ such that $\phi(t) \geq \phi_0 t$ for any $t \in [0, t_0]$;*
5. *β is strictly increasing and there exist $\beta_0 > 0$ and $\bar{\mu} > 0$ such that $\beta(\mu) \geq \beta_0 \mu$ for any $\mu \in [0, \bar{\mu}]$.*

Then the penalty function F_λ is globally exact.

Proof. Applying Theorem 3, and taking into account the assumptions of the theorem one obtains that the penalty function F_λ is exact at every globally optimal solution of the problem (\mathcal{P}) . Since f is l.s.c. on A , and the set A is

compact, then f is bounded below on this set. Hence the function $g(x, \mu)$ is bounded below on $A \times \mathbb{R}_+$, which implies that the penalty function F_λ is also bounded below on $A \times \mathbb{R}_+$ for any $\lambda \geq 0$.

Let us show that the function $g(x, \mu)$ is l.s.c. on $A \times \mathbb{R}_+$. Then with the use of Corollary 1 one obtains the desired result.

For any $\varepsilon > 0$ introduce the function

$$g_\varepsilon(x, \mu) = f(x) + \frac{1}{\mu + \varepsilon} \phi(d(y_0, \Phi(x))^2) \quad (x, \mu) \in A \times \mathbb{R}_+.$$

Taking into account the fact that the function $x \rightarrow d(y_0, \Phi(x))$ is continuous on A , and ϕ is l.s.c., one gets that the function $x \rightarrow \phi(d(y_0, \Phi(x))^2)$ is l.s.c. on A as well. Hence and from the lower semicontinuity of f it follows that the function $g_\varepsilon(x, \mu)$ is l.s.c. on $A \times \mathbb{R}_+$. Note that

$$g(x, \mu) = \sup_{\varepsilon > 0} g_\varepsilon(x, \mu) \quad \forall (x, \mu) \in A \times \mathbb{R}_+.$$

Therefore the function $g(x, \mu)$ is l.s.c. on $A \times \mathbb{R}_+$ as the supremum of a family of l.s.c. functions. \square

Theorem 3 can be modified to the case of more general functions $g(x, \mu)$ and $F_\lambda(x, \mu)$. In particular, let

$$g(x, \mu) = \begin{cases} f(x), & \text{if } x \in \Omega, \mu = 0, \\ +\infty, & \text{if } x \notin \Omega, \mu = 0, \\ f(x) + \frac{1}{\mu^\alpha} \phi(d(y_0, \Phi(x))^2), & \text{if } \mu > 0. \end{cases}$$

and

$$F_\lambda(x, \mu) = g(x, \mu) + \lambda \beta(\mu),$$

where $\alpha > 0$ (cf. [13, 12, 14]). The following result holds true.

Theorem 5. *Let $x^* \in \text{dom } f$ be a point local minimum of the problem (10), the function f be Lipschitz continuous near x^* , and the set-valued mapping Φ be metrically subregular with respect to the set A at (x^*, y_0) . Suppose that the following assumptions are satisfied:*

1. $\phi(t) \geq \phi_0 t^\gamma$ for any $t \in [0, t_0]$, and for some $\phi_0 > 0$, $\gamma > 0$ and $t_0 > 0$;
2. $\beta(\mu) \geq \beta_0 \mu^\sigma$ for any $\mu \in [0, \bar{\mu}]$, and for some $\beta_0 > 0$, $\sigma > 0$ and $\bar{\mu} > 0$.

Suppose also that

$$\gamma > \frac{1}{2}, \quad \sigma \leq \frac{\alpha}{2\gamma - 1}.$$

Then the penalty function $F_\lambda(x, \mu)$ is exact at x^* .

Proof. Arguing in the same way as in the proof of Theorem 3 one can easily verify that there exists $\Theta > 0$ such that

$$g(x, \mu) - g(x^*, 0) \geq -\Theta \mu^{\frac{\alpha}{2\gamma - 1}} \quad \forall x \in B(x^*, r) \cap A \quad \forall \mu \geq 0,$$

where $r > 0$ is sufficiently small. Then applying the assumption on the function β one can check that F_λ is exact at x^* , provided $\sigma \leq \alpha/(2\gamma - 1)$. \square

References

- [1] D. Azé. Unified theory for metric regularity of multifunctions. *J. Convex Anal.*, 13:225–252, 2006.
- [2] L. Bingzhuang and W. Z. A modified exact smooth penalty function for nonlinear constrained optimization. *J. Inequal. Appl.*, 2012:1–11, 2012.
- [3] J. Burke. Calmness and exact penalization. *SIAM J. Control Optim.*, 29:493–497, 1991.
- [4] J. Burke. An exact penalization viewpoint on constrained optimization. *SIAM J. Control Optim.*, 29:968–998, 1991.
- [5] F. Clarke. A new approach to Lagrange multipliers. *Math. Oper. Res.*, 1:165–174, 1976.
- [6] F. Clarke. *Optimization and Nonsmooth Analysis*. John Wiley & Sons, New York, 1983.
- [7] V. Demyanov. Nonsmooth optimization. In G. Di Pillo and F. Schoen, editors, *Nonlinear Optimization. Lecture Notes in Mathematics, vol. 1989*, pages 55–164. Springer-Verlag, Berlin Heidelberg, 2010.
- [8] G. Di Pillo and L. Grippo. Exact penalty functions in constrained optimization. *SIAM J. Control Optim.*, 27:1333–1360, 1989.
- [9] W. Huyer and A. Neumaier. A new exact penalty function. *SIAM J. Optim.*, 13:1141–1158, 2003.
- [10] C. Jiang, Q. Lin, C. Yu, K. Teo, and G.-R. Duan. An exact penalty method for free terminal time optimal control problem with continuous inequality constraints. *J. Optim. Theory Appl.*, 154:30–53, 2012.
- [11] A. Kruger. Error bounds and metric subregularity. *Optim.*, 64:49–79, 2014.
- [12] B. Li, C. J. Yu, K. Teo, and G. Duan. An exact penalty function method for continuous inequality constrained optimal control problem. *J. Optim. Theory Appl.*, 151:260–291, 2011.
- [13] Q. Lin, R. Loxton, K. Teo, Y. Wu, and C. Yu. A new exact penalty method for semi-infinite programming problems. *J. Comput. Appl. Math.*, 261:271–286, 2014.
- [14] Q. Lin, R. Loxton, K. L. Teo, and Y. Wu. Optimal feedback control for dynamic systems with state constraints: An exact penalty approach. *Optim. Lett.*, 8:1535–1551, 2014.
- [15] C. Ma, X. Li, K.-F. Cedric Yiu, and L.-S. Zhang. New exact penalty function for solving constrained finite min-max problems. *Appl. Math. Mech.-Engl. Ed.*, 33:253–270, 2012.
- [16] J.-P. Penot. Calmness and stability properties of marginal and performance functions. *Numer. Funct. Anal. Optim.*, 25:287–308, 2004.

- [17] A. Uderzo. Exact penalty functions and calmness for mathematical programming under nonlinear perturbations. *Nonlinear Anal.*, 73:1596–1609, 2010.
- [18] C. Wang, C. Ma, and J. Zhou. A new class of exact penalty functions and penalty algorithms. *J. Glob. Optim.*, 58:51–73, 2014.