

# On the Merit and Penalty Functions for the D.C. Optimization

Alexander S. Strekalovsky<sup>(✉)</sup>

Institute for System Dynamics and Control Theory SB RAS,  
Lermontov street, 134, Irkutsk 664033, Russia  
`strekal@icc.ru`

**Abstract.** This paper addresses a rather general problem of nonlinear optimization with the inequality constraints and the goal function defined by the (d.c.) functions represented by the difference of two convex functions. In order to reduce the constrained optimization problem to an unconstrained one, we investigate three auxiliary problems with the max-merit, Lagrange and penalty goal functions. Further, their relations to the original problem are estimated by means of the new Global Optimality Conditions and classical Optimization Theory as well as by examples.

**Keywords:** Nonlinear programming · d.c. functions · Global optimality conditions · The lagrange function · Penalty functions

## 1 Introduction

It is well-known that the contemporary Optimality Conditions (OC) theory [1–13], a considerable part of which is presented by modern generalizations of the KKT-theorem [1–13], turns out to be ineffective when it comes to a characterization of a global solution to nonconvex problems. Meanwhile, real-life applied problems might have a lot (often a huge number!) of local extrema [15–22].

On the other hand, new attractive and promising areas for investigations in optimization in the 21st century arise from various applications. Among others, let us mention the following problems: the search for equilibrium in competitions; hierarchical optimization problems; dynamical control problems. However, as it has been shown in [17], it turns out that all these new problems are related to nonconvexity. Hence, it becomes obvious that we need new mathematical tools (optimality conditions, numerical procedures etc.) that would allow us to escape stationary or local solutions and construct numerical procedures able to jump out of local pits simultaneously improving the goal functions. The first attempts to propose such an apparatus have been undertaken in [17–22] for special d.c. optimization problems such as d.c. minimization, convex maximization, reverse-convex problems etc [15–22]. To deal with this class of problems, the Global Search Theory [17, 21] has been developed, which comprises local search methods [17–19, 21] and global search procedures [17, 21, 22] that employ classical methods.

In this paper, we address more general problems with d.c. functions in inequality constraints and goal functions and investigate the reduction of the constrained problem to three unconstrained ones formed by the max-merit, Lagrange and penalty goal functions. The study is performed, in particular, by means of the new Global Optimality Conditions for three different auxiliary unconstrained problems. After the statement of the problem in Sects. 2 and 3 considers the max-merit function and an auxiliary problem with Example 1 to illustrate inadequacy of the max-merit function to the original Problem.

Further, we study the well-known Lagrange function and show by means of the Global Optimality Conditions (GOC) and examples that  $\mathcal{L}(x, \lambda)$  seeks a saddle point (which does not always exist) but not a solution to the original problem.

In Sect. 5, we consider a penalization approach and provide some necessary information. We study the GOC for the penalized problem and use examples to demonstrate that the new tools are effective.

Section 6 provides the conclusion to summarize the content of the paper.

## 2 Statement of Problem

Let us consider the following problem:

$$(\mathcal{P}): \left. \begin{aligned} f_0(x) &:= g_0(x) - h_0(x) \downarrow \min, & x \in S, \\ f_i(x) &:= g_i(x) - h_i(x) \leq 0, & i \in I = \{1, \dots, m\}; \end{aligned} \right\} \quad (1)$$

where the functions  $g_i(\cdot)$ ,  $h_i(\cdot)$ ,  $i = 0, 1, \dots, m$ , are convex on  $\mathbb{R}^n$ , so that the functions  $f_i(\cdot)$ ,  $i = 0, 1, \dots, m$ , are the (d.c.) functions of A.D. Alexandrov represented as a difference of two convex functions [1, 2, 4, 5].

In order to avoid some singularities [1, 4, 5], we assume that

$$S \subset \left[ \bigcap_{i=0}^m \text{int}(\text{dom } g_i(\cdot)) \right] \cap \left[ \bigcap_{i=0}^m \text{int}(\text{dom } h_i(\cdot)) \right] \neq \emptyset,$$

where the set  $S \subset \mathbb{R}^n$  is convex.

Further, let the following assumptions hold:

$$\begin{aligned} D &:= \{x \in S \mid f_i(x) \leq 0, \quad i \in I\} \neq \emptyset, \\ \mathcal{V}(\mathcal{P}) &:= \inf(f_0, D) \triangleq \inf\{f_0(x) \mid x \in S, \quad f_i(x) \leq 0, \quad i \in I\} > -\infty, \\ \text{Sol}(\mathcal{P}) &:= \{z \in D \mid f_0(z) = \mathcal{V}(\mathcal{P})\} \neq \emptyset. \end{aligned}$$

## 3 The Max-Merit Function

Let us consider the following function [5–9]

$$F(x, \eta) := \max\{f_0(x) - \eta; f_1(x), \dots, f_m(x)\}, \quad (2)$$

where  $\eta \in \mathbb{R}$ . Below, for a feasible (in  $(\mathcal{P})$ ) point  $z \in D$ , we denote  $\zeta := f_0(z)$ .

**Proposition 1** ([5–9]). *Suppose that a point  $z$  is a solution to Problem  $(\mathcal{P})$ :  $z \in \text{Sol}(\mathcal{P})$ . Then, the point  $z$  is a solution to the following auxiliary problem*

$$(\mathcal{P}_\zeta): \quad F(x, \zeta) \downarrow \min_x, \quad x \in S. \tag{3}$$

**Proposition 2** ([5]). *Suppose the parameter  $\eta$  is equal to the optimal value of Problem  $(\mathcal{P})$ –(1),  $\eta = \mathcal{V}(\mathcal{P})$ . Then, the point  $z \in D$  is a solution to Problem  $(\mathcal{P})$  if and only if  $z$  is a solution to the auxiliary problem  $(\mathcal{P}_\zeta)$  with  $\zeta := f_0(z) = \mathcal{V}(\mathcal{P})$ . Under latter conditions, the equality  $\text{Sol}(\mathcal{P}) = \text{Sol}(\mathcal{P}_\zeta)$  holds.*

**Lemma 1.** *Suppose that the point  $z \in D$  is not a solution to problem  $(\mathcal{P}_\eta)$  with  $\eta = \zeta := f_0(z)$ , so that there exists a point  $u \in S$ , such that*

$$F(u, \zeta) < 0 = F(z, \zeta).$$

*Then, the point  $z \in D$  cannot be a solution to Problem  $(\mathcal{P})$ :  $z \notin \text{Sol}(\mathcal{P})$ .*

*Proof.* From the inequality  $F(u, \zeta) < 0$  it follows that  $u \in S$ ,  $f_i(u) < 0$ , and  $f_0(u) < f_0(z) = \zeta$ , so that  $u$  is feasible for Problem  $(\mathcal{P})$ . Hence,  $z \notin \text{Sol}(\mathcal{P})$ .  $\square$

Note that it is not difficult to show that the objective function  $F(x, \eta)$  of Problem  $(\mathcal{P}_\eta)$ –(3), given in (2), is a d.c. function.

*Remark 1.* Let us now pay attention to the fact that Proposition 2 provides the sufficient conditions for  $z \in S$  to be a solution to Problem  $(\mathcal{P}_\zeta)$ ,  $\zeta = f_0(z)$ , but not to the original Problem  $(\mathcal{P})$ . Only if we added the equality  $\zeta = \mathcal{V}(\mathcal{P})$  (see Proposition 2), which is, in particular, rather difficult to verify in the majority of the applied problems, then we would be able to make a conclusion about the global solution property in Problem  $(\mathcal{P})$  of the feasible point  $z \in D$  under investigation. On the other hand, the equality  $f_0(z) = \zeta = \mathcal{V}(\mathcal{P})$  for a feasible point  $z \in D$  immediately provides that  $z \in \text{Sol}(\mathcal{P})$  without any supplementary conditions. So, this condition appears to be incorrect. In order to see the adequateness of  $F(x, \eta)$  with respect to Problem  $(\mathcal{P})$  let us consider an example.  $\square$

*Example 1.* Consider the problem

$$\left. \begin{aligned} f_0(x) &= \frac{1}{2}(x_1 - 4)^2 + (x_2 + 2)^2 \downarrow \min_x, \\ f_1(x) &= (x_1 - 1)^2 - (x_2 + 1)^2 \leq 0, \\ f_2(x) &= (x_2 - 2)^2 - (x_1 + 2)^2 \leq 0. \end{aligned} \right\} \tag{4}$$

It is easy to see that the point  $z_* = (4, -2)^\top$  is the global minimum of the strongly convex function  $f_0(\cdot)$  on  $\mathbb{R}^2$ , and  $f_0(z_*) = 0$  provides a lower bound for  $\mathcal{V}(4) = \inf(f_0, D) \geq 0$ . Note that  $z_*$  is unfeasible in (4), since  $f_1(z_*) = 8 > 0$ . Let us consider another point  $z = (\frac{4}{3}, -\frac{2}{3})^\top$  which is feasible for (4), since  $f_1(z) = 0$  and  $f_2(z) = -4 < 0$ . In addition, it can be readily seen that  $z$  satisfies the KKT-conditions with  $\lambda_0 = 1$ ,  $\lambda_1 = 4 > 0$ ,  $\lambda_2 = 0$ ,  $\zeta := f_0(z) = 5\frac{1}{3}$ .

Further, since

$$f_1(x) = (x_1 - x_2 - 2)(x_1 + x_2) \leq 0, \quad f_2(x) = (x_2 - x_1 - 4)(x_2 + x_1) \leq 0,$$

it can be readily shown that the feasible set  $D := \{x \in \mathbb{R}^2 \mid f_i(x) \leq 0, \ i = 1, 2\}$  is represented by the union of the two convex parts:  $D = D_1 \cup D_2$ ,  $D_1 = \{x \mid x_1 + x_2 = 0\}$ ,  $D_2 = \{x \mid x_1 + x_2 \geq 0, \ x_2 - x_1 - 4 \leq 0, \ x_1 - x_2 - 2 \leq 0\}$ .

Hence, from the geometric view-point, it is easy to see that the point  $z_0 = (\frac{8}{3}, -\frac{8}{3})^\top$  is the global solution to (4) with the optimal value  $\mathcal{V}(4) = f_0(z_0) =: \zeta_0 = \frac{4}{3}$ . So, the point  $z = (\frac{4}{3}, -\frac{2}{3})^\top$  is not to be a global solution to (4), since  $f_0(z) = 5\frac{1}{3} = \zeta$ . However, the goal function  $F(x, \zeta)$  of Problem  $(\mathcal{P}_\zeta)$  does not distinguish between these two points. Actually,

$$F(z_0, \zeta) = \max\{f_0(z_0) - \zeta; f_1(z_0); f_2(z_0)\} = 0 = \max\{f_0(z) - \zeta; f_1(z); f_2(z)\} = F(z, \zeta),$$

because  $z \in D$ ,  $z_0 \in D = \{x \in \mathbb{R}^2 \mid f_i(x) \leq 0, \ i = 1, 2\}$ .

Moreover, for all feasible (in Problem (4)) points, which are better (in the sense of the problem (4)) than the point  $z$ , i.e.  $u \in \{x \in \mathbb{R}^2 \mid x_1 + x_2 = 0, \ f_0(x) < \zeta = 5\frac{1}{3}\}$ , we have the same results:  $F(u, \zeta) = 0$ , because  $f_1(u) = 0 = f_2(u)$ . For instance, for any point  $x(\alpha)$  of the form

$$x(\alpha) = (v(\alpha), -v(\alpha))^\top, \quad v(\alpha) = 1.1\alpha + 4.2(1 - \alpha), \quad \alpha \in [0, 1],$$

we have  $F(x(\alpha), \zeta) = 0$ . Meanwhile,  $f_0(x(\alpha)) < f_0(z) = 5\frac{1}{3} \ \forall \alpha \in [0, 1]$ .

Therefore, one can see that there exist a great number of points better than  $z$  in the sense of Problem (4). □

So, Example 1 demonstrates that Problem  $(\mathcal{P}_\eta)$  is not sufficiently adequate to Problem  $(\mathcal{P})$ . More precisely, taking into account Propositions 1 and 2, it is easy to see that the set  $Sol(\mathcal{P}_\zeta)$  might contain a lot of points not belonging to  $Sol(\mathcal{P})$ , so that the inclusion  $Sol(\mathcal{P}) \subset Sol(\mathcal{P}_\zeta)$  may be really proper. Moreover, the inequality  $\zeta > \zeta_* = \mathcal{V}(\mathcal{P})$  holds together with the inclusion  $Sol(\mathcal{P}) \subset Sol(\mathcal{P}_\zeta)$ , which is inadmissible. Therefore, we move on to another type of the merit (or penalty, in a rough sense) function.

## 4 The Lagrange Function

Consider the standard (normal) Lagrange function for Problem  $(\mathcal{P})$

$$\mathcal{L}(x, \lambda) = f_0(x) + \sum_{i=1}^m \lambda_i f_i(x). \tag{5}$$

It is common to call a pair  $(z, \lambda)$  a saddle point of the Lagrange function  $\mathcal{L}(x, \lambda)$ :  $(z, \lambda) \in Sdl(\mathcal{L})$ , if the following two inequalities are satisfied [1–9]:

$$\forall \mu \in \mathbb{R}_+^m: \quad \mathcal{L}(z, \mu) \leq \mathcal{L}(z, \lambda) \leq \mathcal{L}(x, \lambda) \quad \forall x \in S. \tag{6}$$

**Lemma 2** ([1, 2, 4–9]). *For a pair  $(z, \lambda) \in S \times \mathbb{R}_+^m$  the following two assertions are equivalent:*

$$(i) \quad \max_{\mu} \{ \mathcal{L}(z, \mu) \mid \mu \in \mathbb{R}_+^m \} = \mathcal{L}(z, \lambda); \tag{7}$$

$$(ii) \quad z \in D, \quad \lambda \in \mathbb{R}_+^m, \quad \lambda_i f_i(z) = 0, \quad i = 1, \dots, m. \tag{8}$$

Recall that a vector  $\lambda \in \mathbb{R}_+^m$  satisfying the KKT-conditions, including (8), is usually called a Lagrange multiplier [1, 2, 4–9] at a point  $z \in D$ . The set of all Lagrange multipliers at  $z$  will be denoted below by  $\mathcal{M}(z)$ .

Remember, in addition, that for a convex optimization problem (P)–(1), when  $h_i(x) \equiv 0 \ \forall i \in \{0\} \cup I$ , we have  $\mathcal{M}(z_1) = \mathcal{M}(z_2) = \mathcal{M}$ , if  $z_i \in \text{Sol}(\mathcal{P})$ ,  $i = 1, 2$  [5, Chapter VII].

**Proposition 3** ([1, 2, 4, 5, 7–9]). *If the pair  $(z, \lambda) \in S \times \mathbb{R}_+^m$  is a saddle point of the Lagrange function  $\mathcal{L}(x, \mu)$  on the set  $S \times \mathbb{R}_+^m$ , then the point  $z$  is a global solution to Problem (P).*

In what follows, we will employ this assertion in a different form.

**Proposition 4.** *Suppose  $z \in D$ ,  $z$  is a KKT-point but not a global solution to Problem (P). Then, there exists no Lagrange multiplier  $\lambda \in \mathcal{M}(z)$  such that  $(z, \lambda) \in \text{Sdl}(\mathcal{L})$ .*

Further, since  $f_i(x) = g_i(x) - h_i(x)$ ,  $i = 0, 1, \dots, m$ ,  $\mathcal{L}(x, \lambda)$  has a very simple, clear and suitable d.c. representation

$$\left. \begin{aligned} (a) \quad & \mathcal{L}(x, \lambda) = G_{\lambda}(x) - H_{\lambda}(x), \\ (b) \quad & G_{\lambda}(x) = g_0(x) + \sum_{i=1}^m \lambda_i g_i(x), \quad H_{\lambda}(x) = h_0(x) + \sum_{i=1}^m \lambda_i h_i(x). \end{aligned} \right\} \tag{9}$$

Taking into account (9), let us look at the normal Lagrange function (5) from the view-point of the Global Optimality Conditions (GOC) [17, 18, 20, 21].

**Theorem 1.** *Suppose  $(z, \lambda) \in \text{Sdl}(\mathcal{L})$ ,  $\lambda_0 = 1$ ,  $\zeta := f_0(z)$ . Then,  $\forall (y, \beta) \in \mathbb{R}^n \times \mathbb{R}$  such that*

$$H_{\lambda}(y) := \sum_{i=0}^m \lambda_i h_i(y) = \beta - \zeta, \tag{10}$$

*the following inequality holds*

$$G_{\lambda}(x) - \beta \geq \sum_{i=0}^m \lambda_i \langle h'_i(y), x - y \rangle \quad \forall x \in S, \tag{11}$$

*for any subgradients  $h'_i(y) \in \partial h_i(y)$  of the functions  $h_i(\cdot)$  at the point  $y$ ,  $i \in I \cup \{0\}$ .*

*Proof.* According to the assumption, we have the chain

$$\zeta := f_0(z) = \sum_{i=0}^m \lambda_i f_i(z) = \mathcal{L}(z, \lambda) \leq \sum_{i=0}^m \lambda_i f_i(x) = \mathcal{L}(x, \lambda) \quad \forall x \in S,$$

from which, due to (9) and (10), it follows

$$\beta - H_\lambda(y) = \zeta \leq \mathcal{L}(x, \lambda) = G_\lambda(x) - H_\lambda(x) \quad \forall x \in S.$$

Further, by the convexity of  $H_\lambda(\cdot) = \sum_{i=0}^m \lambda_i h_i(\cdot)$ ,  $\lambda_i \geq 0$ ,  $i \in I$ , we obtain

$$G_\lambda(x) - \beta \geq H_\lambda(x) - H_\lambda(y) \geq \sum_{i=0}^m \lambda_i \langle h'_i(y), x - y \rangle \quad \forall x \in S,$$

which coincides with (11). □

*Remark 2.* Due to Proposition 3, it can be readily seen that for a global solution  $z \in \text{Sol}(\mathcal{P})$ , for which one can find a multiplier  $\lambda \in \mathcal{M}(z)$  such that  $(z, \lambda) \in \text{Sdl}(\mathcal{L})$ , the conditions (10)–(11) turn out to be necessary global optimality conditions.

*Remark 3.* It is clear that Theorem 1 reduces the nonconvex problem

$$(\mathcal{L}): \quad \mathcal{L}(x, \lambda) \downarrow \min_x, \quad x \in S,$$

to the verification of the principal inequality (PI) (11) for the family of parameters  $(y, \beta)$ :  $H_\lambda(y) = \beta - \zeta$ , or, more precisely, to solving the family of the convex linearized problems

$$(\mathcal{L}L(y)): \quad \Phi_\lambda(x) = G_\lambda(x) - \langle H'_\lambda(y), x \rangle \downarrow \min_x, \quad x \in S, \tag{12}$$

with the subsequent verification of PI (11) with  $x = u = u(y, \beta) \in \text{Sol}(\mathcal{L}L(y))$ .

*Remark 4.* Furthermore, suppose that there exists a tuple  $(y, \beta, u)$ , such that  $(y, \beta)$  satisfies the equality (10) and violates the PI (11), i.e.  $G_\lambda(u) - \beta < \langle H'_\lambda(y), u - y \rangle$ . Whence, due to convexity of  $H_\lambda(\cdot) = \sum_{i=0}^m \lambda_i h_i(\cdot)$ , it follows

$$G_\lambda(u) - \beta < H_\lambda(u) - H_\lambda(y).$$

Next, on account of (9) and (10), we have

$$\mathcal{L}(u, \lambda) = G_\lambda(u) - H_\lambda(u) < \beta - H_\lambda(y) = \zeta = f_0(z) = \mathcal{L}(z, \lambda),$$

where  $\lambda \in \mathcal{M}(z)$ . Hence, the right-hand-side inequality in (6) is violated with  $u \in S$ . It means that the pair  $(z, \lambda)$  is not a saddle point:  $(z, \lambda) \notin \text{Sdl}(\mathcal{L})$ . □

Therefore, from the point-of-view of optimization theory [1, 2, 4–12, 17, 18, 20, 21] the conditions (10)–(11) of Theorem 1 possess the constructive property. Nevertheless, it is not clear whether there exists a tuple  $(y, \beta, u)$  that violates (11). The answer is given by the following result.

**Theorem 2.** *Let there be given a KKT-point  $z \in D$  with the corresponding multipliers  $\lambda \in \mathcal{M}(z)$ ,  $\lambda_0 = 1$ . In addition, let the following assumption take place*

$$(\mathcal{H}): \quad \exists v \in \mathbb{R}^n: \quad \mathcal{L}(v, \lambda) > \mathcal{L}(z, \lambda) = f_0(z) =: \zeta. \quad (13)$$

*Besides, suppose that the pair  $(z, \lambda)$  is not a saddle point of  $\mathcal{L}(z, \lambda)$  on  $S \times \mathbb{R}_+^m$ .*

*Then, one can find a tuple  $(y, \beta, u)$ , where  $(y, \beta) \in \mathbb{R}^{n+1}$ ,  $u \in S$ , and a fixed ensemble of subgradients  $\{h'_{i0}(y), h'_{i1}(y), \dots, h'_{im}(y)\}$ ,  $h'_{i0}(y) \in \partial h_i(y)$ ,  $i \in \{0\} \cup I$ , such that*

$$H_\lambda(y) \triangleq \sum_{i=0}^m \lambda_i h_i(y) = \beta - \zeta, \quad (14)$$

$$G_\lambda(y) \leq \beta, \quad (15)$$

$$G_\lambda(u) - \beta < \sum_{i=0}^m \lambda_i \langle h'_{i0}(y), u - y \rangle. \quad (16)$$

□

Let us verify the effectiveness of the constructive property of the GOC of Theorems 1 and 2 by the example.

*Example 1 (Revisited).* As it has been shown above, the point  $z = (\frac{4}{3}, -\frac{2}{3})^\top$  is the KKT-point with  $\lambda_0 = 1$ ,  $\lambda_1 = 4$ ,  $\lambda_2 = 0$ . Recall that  $\zeta := f_0(z) = 5\frac{1}{3}$ . Meanwhile, there exist points feasible in the problem (4) and better than  $z$  in the sense of the problem (4).

Now, let us apply the GOC of Theorems 1 and 2 in order to improve the point  $z$ . For this purpose, employ the Lagrange function  $\mathcal{L}(x, \lambda)$  with  $\lambda = (1, 4, 0) \in \mathcal{M}(z)$ :

$$\mathcal{L}(x, \lambda) = f_0(x) + \lambda f_1(x) = \frac{1}{2}(x_1 - 4)^2 + (x_2 + 2)^2 + 4[(x_1 - 1)^2 - (x_2 + 1)^2].$$

Then we have  $\mathcal{L}(x, \lambda) = G_\lambda(x) - H_\lambda(x)$ , where

$$G_\lambda(x) = \frac{1}{2}(x_1 - 4)^2 + (x_2 + 2)^2 + 4(x_1 - 1)^2, \quad H_\lambda(x) = 4(x_2 + 1)^2. \quad (17)$$

Let us choose  $y = (0, -\frac{1}{2})^\top$ ,  $u = (\frac{4}{3}, 0)^\top$ ,  $f_1(u) = -\frac{8}{9} < 0$ ,  $f_2(u) = -7\frac{1}{9} < 0$ .

$$\nabla H_\lambda(x) = \begin{bmatrix} 0 \\ 8(x_2 + 1) \end{bmatrix}, \quad \nabla H_\lambda(y) = \begin{bmatrix} 0 \\ 8(y_2 + 1) \end{bmatrix} = \begin{pmatrix} 0 \\ 4 \end{pmatrix}.$$

Further, we obtain that  $\beta = H_\lambda(y) + \zeta = 4(-\frac{1}{2} + 1)^2 + 5\frac{1}{3} = 6\frac{1}{3}$ ,

$$\langle \nabla H_\lambda(y), u - y \rangle = \langle \begin{pmatrix} 0 \\ 4 \end{pmatrix}, \begin{pmatrix} \frac{4}{3} \\ 0.5 \end{pmatrix} \rangle = 2, \quad \psi(y, \beta) \triangleq \beta + \langle \nabla H_\lambda(y), u - y \rangle = 8\frac{1}{3}.$$

$$G_\lambda(u) = \frac{1}{2}(\frac{4}{3} - 4)^2 + 2^2 + 4(\frac{4}{3} - 1)^2 = \frac{32}{9} + 4 + \frac{4}{9} = 8.$$

Hence, we see that

$$G_\lambda(u) = 8 < 8\frac{1}{3} = \beta + \langle \nabla H_\lambda(y), u - y \rangle,$$

and the PI (11) is violated. Due to Theorems 1 and 2, it means that  $(z, \lambda)$  is not a saddle point of the Lagrange function, which can be proved as follows:

$$\mathcal{L}(u, \lambda) \triangleq f_0(u) + 4f_1(u) = 7\frac{5}{9} + 4 \cdot (-\frac{8}{9}) = 4 < 5\frac{1}{3} = f_0(z) = \mathcal{L}(z, \lambda).$$

On the other hand, it is easy to compute that

$$f_0(u) = \frac{1}{2}(\frac{4}{3} - 4)^2 + 2^2 = \frac{32}{9} + 4 = 7\frac{5}{9} > f_0(z) = 5\frac{1}{3},$$

so that there is no improvement at all in the original problem (4). Consequently, we see that the GOC of Theorems 1 and 2 allow us to improve the point  $z$  in the sense of the Lagrange function, since they are striving to minimize the Lagrange function with respect to the variable  $x$ . However, they do not aim at minimizing the function  $f_0(x)$  over the feasible set  $D$ , i.e. at solving Problem (P). □

The next result will be useful below.

**Lemma 3** ([9,10,12]). *A point  $z$  is a solution to the problem*

$$(Q): \left. \begin{aligned} S(x) &:= \varphi(x) + f(x) \downarrow \min, & x \in S, \\ \varphi(x) &:= \max_j \{ \varphi_j(x) \mid j \in J = \{1, \dots, N\} \}. \end{aligned} \right\} \quad (18)$$

*if and only if the pair  $(z, t_*)$  is a solution to the problem*

$$(QA): \Phi(x, t) := t + f(x) \downarrow \min, \quad x \in S, t \in \mathbb{R}, \quad \varphi_j(x) \leq t, j \in J, \quad (19)$$

where

$$t_* = \varphi(z) = \max_j \{ \varphi_j(z) \mid j \in J \}. \quad (20)$$

**Lemma 4** *Let the quadratic function  $q(x) := \frac{1}{2} \langle x, Ax \rangle - \langle b, x \rangle$  with the positive definite matrix  $A = A^T > 0$ ,  $b \in \mathbb{R}^n$ , be given. Consider the optimization problem (with a parameter  $u \in \mathbb{R}^n$ )*

$$Q(y, \beta) := \beta + \langle \nabla q(y), u - y \rangle \uparrow \max_{(y, \beta)} \quad (y, \beta) \in \mathbb{R}^{n+1}: \quad q(y) = \beta - \gamma. \quad (21)$$

*Then, the solution  $(y_*, \beta_*)$  to (21) is provided by the equalities*

$$y_* = u, \quad \beta_* = q(y_*) + \gamma. \quad (22)$$

*Example 2* (of G.R. Walsh, [23], p. 67). Consider the problem

$$\left. \begin{aligned} f_0(x) &= x_1 x_2 - 2x_1^2 - 3x_2^2 \downarrow \min, \\ f_1(x) &= 3x_1 + 4x_2 - 12 \leq 0, & f_2(x) &= x_2^2 - x_1^2 + 1 \leq 0, \\ f_3(x) &= -x_1 \leq 0, & f_4(x) &= x_1 - 4 \leq 0, \\ f_5(x) &= -x_2 \leq 0, & f_6(x) &= x_2 - 3 \leq 0. \end{aligned} \right\} \quad (23)$$



Let us study the unique solution  $z = (4, 0)^\top$  to (23),  $f_0(z) = \zeta = -32$ . Clearly, the Lagrange function at  $z$  takes the form  $\mathcal{L}(x, \lambda) = x_1x_2 - 2x_1^2 - 3x_2^2 + \lambda_1(3x_1 + 4x_2 - 12) + \lambda_4(x_1 - 4) - \lambda_5x_2$ ,  $S = \mathbb{R}^2$ , because  $\lambda_2 = \lambda_3 = \lambda_6 = 0$ , due to the complementarity conditions:  $\lambda_i f_i(x) = 0$ ,  $i = \overline{1, 6}$ . In addition, with the help of the KKT-conditions at  $z = (4, 0)^\top$ , we derive that  $\lambda_1 = 2$ ,  $\lambda_4 = 10$ ,  $\lambda_5 = 12$ , so that

$$\mathcal{L}(x, \lambda) = x_1x_2 - 2x_1^2 - 3x_2^2 + 16x_1 - 4x_2 - 64. \tag{24}$$

Besides,  $\mathcal{L}(z, \lambda) = -32 = \zeta = f_0(z)$ , as it should be. Further, it can be readily seen that the function  $x \mapsto \mathcal{L}(x, \lambda)$  is a d.c. one. We will use the d.c. representation as follows:  $\mathcal{L}(x, \lambda) = G_\lambda(x) - H_\lambda(x)$ , where

$$G_\lambda(x) = 2(x_1^2 + x_2^2) + 16x_1 - 4x_2 - 64, \quad H_\lambda(x) = 4x_1^2 + 5x_2^2 - x_1x_2. \tag{25}$$

Therefore, one can see that  $\beta = H_\lambda(y) + \zeta = H_\lambda(y) - 32$ ,  $\nabla H_\lambda(x) = (8x_1 - x_2, 10x_2 - x_1)^\top$ ,  $\langle \nabla H(y), u - y \rangle = 8y_1u_1 - y_2u_1 + 2y_1y_2 - 8y_1^2 - 10y_2^2 + 10y_2u_2 - y_1u_2$ , from which it follows that

$$\begin{aligned} \theta(y, \beta) &:= \beta + \langle \nabla H(y), u - y \rangle = \\ \psi(y) &:= (8u_1 - u_2)y_1 + (10u_2 - u_1)y_2 + y_1y_2 - 4y_1^2 - 5y_2^2 - 32. \end{aligned} \tag{26}$$

Furthermore, Lemma 4 leads us to the equality:  $y_* = u$ . Now, let us choose the vector  $u$  as  $u = (-\frac{1}{5}, -\frac{4}{5})^\top$ , since  $S = \mathbb{R}^2$ . Then one can compute that  $G_\lambda(u) = -62\frac{16}{25} < -42 = \psi(y_*) = \theta(y_*, \beta_*)$ , so that the inequality (11) is violated. Therefore, due to Theorems 1 and 2, we see that the pair  $(z, \lambda)$  is not a saddle point of  $\mathcal{L}(u, \lambda)$ . The latter assertion can be easily verified by the direct calculations  $\mathcal{L}(u, \lambda) = -65\frac{21}{25} < -32 = \mathcal{L}(z, \lambda)$ .

To sum up, we see that there does not exist a Lagrange multiplier  $\lambda$  such that  $(z, \lambda) \in Sdl(\mathcal{L})$  even for the unique global solution  $z \in Sol(\mathcal{P})$ .  $\square$

So, the max-merit function  $F(x, \eta)$  defined in (2), as well as the Lagrange function, possesses some shortcomings, and both functions do not reflect completely the properties of Problem (P). Hence, we have to undertake further investigations, perhaps, with different penalty or merit functions [10–13].

## 5 Penalty Functions

Now, let us introduce now the  $l_\infty$ -penalty function [10–13] for Problem (P)–(1)

$$W(x) := \max\{0, f_1(x), \dots, f_m(x)\} =: \max\{0, f_i(x), i \in I\}. \tag{27}$$

Further, consider the penalized problem as follows

$$(\mathcal{P}_\sigma): \quad \theta_\sigma(x) = f_0(x) + \sigma W(x) \downarrow \min, \quad x \in S. \tag{28}$$

As well-known [5, 6, 11–13], if  $z \in Sol(\mathcal{P}_\sigma)$ , and  $z \in D := \{x \in S \mid f_i(x) \leq 0, \quad i \in I\}$ , then  $z \in Sol(\mathcal{P})$ . On the other hand, if  $z \in Sol(\mathcal{P})$ , then, under supplementary conditions [5, 6, 10–13], for some  $\sigma_* \geq \|\lambda_z\|_1$  (where

$\lambda_z \in \mathbb{R}^m$  is the KKT-multipliers corresponding to  $z$ ), the inclusion  $z \in \text{Sol}(\mathcal{P}_\sigma)$  holds  $\forall \sigma > \sigma_*$ .

Furthermore [5, Lemma 1.2.1, Chapter VII],  $\text{Sol}(\mathcal{P}) = \text{Sol}(\mathcal{P}_\sigma)$ , so that Problems  $(\mathcal{P})$  and  $(\mathcal{P}_\sigma)$  happen to be equivalent  $\forall \sigma \geq \sigma_*$ .

Before we move on any further, a few words should be said about “supplementary conditions”. First of all, let us mention the notion of calmness introduced by R.T. Rockafellar [1] (see F. Clarke [6] and J.V. Burke [13]). To begin with, instead of  $(\mathcal{P})$ , consider the perturbed d.c. optimization problem ( $v \in \mathbb{R}^m$ )

$$(\mathcal{P}(v)) : \quad f_0(x) \downarrow \min_x, \quad x \in S, \quad f_i(x) \leq v_i, \quad i \in I. \tag{29}$$

Let  $x_* \in S$ ,  $v_* \in \mathbb{R}^m$  be such that  $f_i(x_*) \leq v_{*i}$ ,  $i \in I$ . Then, Problem  $(\mathcal{P}(v_*))$  is said to be calm at  $x_*$ , if there exist constants  $\varkappa \geq 0$  and  $\rho > 0$  such that  $\forall (x, v) \in S \times \mathbb{R}^m$  with  $\|x - x_*\| \leq \rho$  and  $f_i(x) \leq v_i$ ,  $i \in I$ . We have

$$f_0(x_*) \leq f_0(x) + \varkappa \|v - v_*\|. \tag{30}$$

The constants  $\varkappa$  and  $\rho$  are called the modulus and the radius of calmness for  $(\mathcal{P}(v_*))$  at  $x_*$ , respectively. Observe that, if  $(\mathcal{P}(v_*))$  is calm at  $x_*$ , then  $x_*$  is a  $\rho$ -local solution to  $(\mathcal{P}(v_*))$ , i.e. ( $v = v_*$ )

$$f_0(x_*) \leq f_0(x) \quad \forall x \in S : \|x - x_*\| \leq \rho, \quad \text{and} \quad f_i(x) \leq v_{*i}, \quad i \in I,$$

i.e.  $x \in D(v_*) \cap \mathbb{B}_x(x_*, \rho)$ . For instance, when  $v_* = 0$  and  $(\mathcal{P}(0))$  is calm (with  $\varkappa \geq 0$  and  $\rho > 0$ ),  $x_*$  is the  $\rho$ -local solution to  $(\mathcal{P}(0)) := (\mathcal{P})$ .

The most fundamental finding consists in the equivalence of the calmness of  $(\mathcal{P}(v_*))$  at  $x_*$  (with  $\varkappa \geq 0$  and  $\rho > 0$ ) and the fact that  $x_*$  is a  $\rho$ -local minimum of the following penalized function with  $\sigma \geq \varkappa$

$$\theta_\sigma(x; v_*) := f_0(x) + \sigma \text{dist}(F(x) \mid \mathbb{R}_-^m + v_*), \tag{31}$$

where  $F(x) = (f_1(x), \dots, f_m(x))^\top$ ,  $\mathbb{R}_-^m = \{y \in \mathbb{R}^m \mid y_i \leq 0, \quad i \in I\}$ ,  $\text{dist}(y_0 \mid C) = \inf\{\|y - y_0\| : y \in C\}$ . If one takes, for instance,  $\|y\|_\infty \triangleq \max_i \{ |y_i| : i \in I \}$ , then it is obvious that

$$W(x) := \max\{0, f_1(x), \dots, f_m(x)\} = \text{dist}_\infty(F(x) \mid \mathbb{R}_-^m).$$

Various conditions can be found in the literature (see [13] and the references therein) that ensure that the calmness hypothesis is satisfied. All of these conditions are related to the regularity of the constraint system of Problem  $(\mathcal{P})$ :

$$x \in S, \quad f_i(x) \leq 0, \quad i \in I. \tag{32}$$

Recall that the system (32) is said to be regular at the solution  $x_0$  (i.e.  $x_0 \in S$ ,  $F(x_0) \leq 0_m$ ) if there exist some constants  $M > 0$  and  $\varepsilon > 0$  such that  $\text{dist}_x(x \mid D(v)) \leq M \text{dist}_y(F(x) \mid \mathbb{R}_-^m + v) \quad \forall x \in S \cap \mathbb{B}_x(x_0, \varepsilon)$  and  $\forall v \in \mathbb{B}_y(0, \varepsilon)$  where

$$D(v) := \{x \in S \mid F(x) \in \mathbb{R}_-^m + v\}, \quad \mathbb{B}_x(x_0, \varepsilon) := \{x \in \mathbb{R}^n \mid \|x - x_0\|_x \leq \varepsilon\}.$$

The conditions yielding the regularity of system (32) and more general systems have been studied in [14]. In the optimization literature such conditions are often called the constraint qualifications conditions, e.g. the Slater and Mangasarian-Fromovitz conditions (etc. see [10–14]).

To sum up, one can say that, under well-known regularity conditions at the global solution  $z \in \text{Sol}(\mathcal{P})$ , Problem  $(\mathcal{P})$  ( $v_* = 0_m$ ) is calm at  $z \in \text{Sol}(\mathcal{P})$  (with the corresponding  $\varkappa \geq 0$  and  $\rho > 0$ ), and, therefore, the goal function of the penalized problem ( $\forall \sigma \geq \varkappa \geq 0$ )

$$(\mathcal{P}_\sigma): \quad \theta_\sigma(x) = f_0(x) + \sigma \text{dist}_\infty(F(x) \mid \mathbb{R}_-^m) = f_0(x) + \sigma W(x) \downarrow \min_x, \quad x \in S$$

attains at  $z$  its global minimum over  $S$ .

Furthermore, it can be readily seen that the penalized function  $\theta_\sigma(\cdot)$  is a d.c. function, because the functions  $f_i(\cdot)$ ,  $i \in I \cup \{0\}$ , are the same. Actually, since  $\sigma > 0$ ,

$$\theta_\sigma(x) = G_\sigma(x) - H_\sigma(x), \tag{33}$$

$$H_\sigma(x) := h_0(x) + \sigma \sum_{i \in I} h_i(x), \tag{34}$$

$$\begin{aligned} G_\sigma(x) &:= \theta_\sigma(x) + H_\sigma(x) \\ &= g_0(x) + \sigma \max \left\{ \sum_{i=1}^m h_i(x); \max_{i \in I} [g_i(x) + \sum_{\substack{j \neq i \\ j \in I}} h_j(x)] \right\}, \end{aligned} \tag{35}$$

it is clear that  $G_\sigma(\cdot)$  and  $H_\sigma(\cdot)$  are convex functions.

For  $z \in S$ , denote  $\zeta := \theta_\sigma(z)$ .

Now we can formulate the major result of the paper.

**Theorem 3** *If  $z \in \text{Sol}(\mathcal{P}_\sigma)$ , then*

$$\forall (y, \beta) : H_\sigma(y) = \beta - \zeta \tag{36}$$

*the following inequality holds*

$$G_\sigma(x) - \beta \geq \langle \nabla h_0(y) + \sigma \sum_{i \in I} \nabla h_i(y), x - y \rangle \quad \forall x \in S. \tag{37}$$

□

It is easy to see that Theorem 3 reduces the nonconvex (d.c.) Problem  $(\mathcal{P}_\sigma)$  to solving the family of convex linearized problems of the form

$$(\mathcal{P}_\sigma \mathcal{L}(y)): \quad G_\sigma(x) - \langle \nabla H_\sigma(y), x \rangle \downarrow \min_x, \quad x \in S, \tag{38}$$

depending on the parameters  $(y, \beta)$  fulfilling (36).

If for such a pair  $(y, \beta)$  and some  $u \in S$  ( $u$  may be a solution to  $(\mathcal{P}_\sigma \mathcal{L}(y))$ ) the inequality (37) is violated, i.e.

$$G_\sigma(u) < \beta + \langle \nabla H_\sigma(y), u - y \rangle, \tag{39}$$

then, due to convexity of  $H_\sigma(\cdot)$  and with the help of (36), we obtain that

$$G_\sigma(u) < \beta + H_\sigma(u) - H_\sigma(y) = H_\sigma(u) + \zeta.$$

The latter implies that  $\theta_\sigma(u) = G_\sigma(u) - H_\sigma(u) < \zeta := \theta_\sigma(z)$ , so that  $u \in S$  is better than  $z$ , i.e.  $z \notin \text{Sol}(\mathcal{P}_\sigma)$ .

It means that the Global Optimality Conditions (36), (37) of Theorem 3 possess the constructive (algorithmic) property allowing to design local and global search methods for Problem  $(\mathcal{P}_\sigma)$  [17, 18, 20–22].

In particular, they enable us to escape a local pit of  $(\mathcal{P}_\sigma)$  to reach a global solution. The question arises whether such a tuple  $(y, \beta, u)$  exists. The answer is given by the following result.

**Theorem 4** *Let for a point  $z \in S$  there exists  $w \in \mathbb{R}^n$  such that*

$$(\mathcal{H}): \quad \theta_\sigma(w) > \theta_\sigma(z).$$

*If  $z$  is not a solution to Problem  $(\mathcal{P}_\sigma)$ , then one can find a pair  $(y, \beta) \in \mathbb{R}^{n+1}$ , satisfying (36), and a point  $u \in S$  such that the inequality (39) holds.  $\square$*

Now let us set  $y = z$  in (38). Then from (36) it follows that  $\beta = \theta_\sigma(z) + H_\sigma(z) = G_\sigma(z)$ . Furthermore, from (37) we derive

$$G_\sigma(x) - G_\sigma(z) \geq \langle \nabla H_\sigma(z), x - z \rangle \quad x \in S,$$

that yields that  $z$  is a solution to the convex linearized problem

$$(\mathcal{P}_\sigma \mathcal{L}(z)): \quad G_\sigma(x) - \langle \nabla H_\sigma(z), x \rangle \downarrow \min_x, \quad x \in S,$$

With the help of Lemma 3 and due to (33)–(35), we see that the latter problem amounts to the next one

$$\left. \begin{aligned} g_0(x) - \langle \nabla H_\sigma(z), x \rangle + \sigma t \downarrow \min_{(x,t)}, \quad x \in S, \quad t \in \mathbb{R}, \\ \sum_{i \in I} h_i(x) \leq t, \quad g_i(x) + \sum_{j \neq i} h_j(x) \leq t, \quad i \in I. \end{aligned} \right\} \quad (40)$$

Moreover, the KKT-conditions to Problem (40) provide for the KKT-conditions at  $z$  for the original Problem  $(\mathcal{P})$ .

So, the Global Optimality Conditions (36), (37) of Theorems 3 and 4 are connected with classical optimization theory [1–13, 15].

*Example 1 (Revisited).* Let us return to problem (4), where the point  $z = (\frac{4}{3}, -\frac{2}{3})^\top$ , with  $f_1(z) = 0$  and  $f_2(z) = -4 < 0$ ,  $\zeta := f_0(z) = 5\frac{1}{3}$ , satisfies the KKT-conditions with  $\lambda_0 = 1$ ,  $\lambda_1 = 4 > 0$ ,  $\lambda_2 = 0$ .

On the other hand, the point  $z_0 = (\frac{8}{3}, -\frac{8}{3})^\top$  is the global solution to (4) with the optimal value  $\mathcal{V}(4) = f_0(z_0) =: \zeta_0 = \frac{4}{3}$ .

In this example,  $h_0(x) \equiv 0$ ,  $g_0(x) = f_0(x) = \frac{1}{2}(x_1 - 4)^2 + (x_2 + 2)^2$ ,  $g_1(x) = (x_1 - 1)^2$ ,  $h_1(x) = (x_2 + 1)^2$ ,  $g_2(x) = (x_2 - 2)^2$ ,  $h_2(x) = (x_1 + 2)^2$ . Therefore, taking into account (34) and (35), one can see that

$$\begin{aligned}
 H_\sigma(x) &:= h_0(x) + \sigma \sum_{i \in I} h_i(x) = \sigma \left[ (x_1 + 2)^2 + (x_2 + 1)^2 \right], \\
 G_\sigma(x) &:= g_0(x) + \sigma \max \left\{ \sum_{i=1}^2 h_i(x); g_1(x) + h_2(x); g_2(x) + h_1(x) \right\} = \\
 &\sigma \max \left\{ (x_1 + 2)^2 + (x_2 + 1)^2; \frac{1}{2}(x_1 - 4)^2 + (x_2 + 2)^2 + \right. \\
 &\left. (x_1 - 1)^2 + (x_1 + 2)^2; (x_2 - 2)^2 + (x_2 + 1)^2 \right\}.
 \end{aligned}$$

Set  $\sigma := 5 = \|\lambda\|_1$ ,  $y = (1.5, -2)^\top \notin D$ . Then, we have  $\beta = H_\sigma(y) + \zeta = 5[(1.5 + 2)^2 + (-2 + 1)^2] + 5\frac{1}{3} = 71\frac{7}{12}$ . Now set  $u = (2, -2)^\top \in D$ .

Then, according to Theorem 3, it is not difficult to compute that

$$\beta + \langle \nabla H_\sigma(y), u - y \rangle = 89\frac{1}{12}.$$

On the other hand, we see that

$$G_\sigma(u) = 87 < 89\frac{1}{12} = \beta + \langle \nabla H_\sigma(y), u - y \rangle.$$

It means that the principal inequality (37) of the GOC is violated, so that  $z \notin \text{Sol}(\mathcal{P}_\sigma)$ . Consequently,  $z \notin \text{Sol}(\mathcal{P})$ , since  $2 = f_0(u) = \theta_\sigma(u) < \theta_\sigma(z) = f_0(z) = 5\frac{1}{3}$ , because  $u$  and  $z$  are feasible.

So, the Global Optimality Conditions (GOC) of Theorems 3 and 4 allow us not only to show that the KKT-point  $z = (\frac{4}{3}, -\frac{2}{3})^\top$  is not a global solution to the problem (7), but, in addition, to construst a feasible point  $u = (2, 2)^\top \in D$  which is better than  $z$  and closer to the global solution  $z_0 = (\frac{8}{3}, -\frac{8}{3})^\top$ . Remember, the max-merit function  $F(x, \zeta)$  does not differ between these two points:

$$F(z, \zeta) = 0 = F(z_0, \zeta).$$

Besides, in order to find a saddle point, the Lagrange function aims at improving exactly itself but not at solving Problem (P). □

Hence, the exact penalization approach demonstrated some advantages in comparison with the max-merit and Lagrange functions.

In addition, employing the constructive property of the GOC of Theorems 3 and 4, we can design the Special Local Search and Global Search Methods. The latter one can escape local pits and attain global solutions in general d.c. optimization problems.

## 6 Conclusion

We considered the reduction of the constrained optimization problem with the d.c. goal function and d.c. inequality constraints to three auxiliary unconstrained problems with different objective functions: the max-merit function, the Lagrange function and an exact penalty function.

The comparison was based on the level of adequacy to the original problem and has been carried out by means of the classical tools, the new Global Optimality Conditions (GOC) and a few examples.

The results showed certain advantages of the exact penalization approach that facilitates development of the new local and global search methods for solving the original problem [17–22].

**Acknowledgments.** This work has been supported by the Russian Science Foundation, project No. 15-11-20015.

## References

1. Rockafellar, R.T.: *Convex Analysis*. Princeton University Press, Princeton (1970)
2. Rockafellar, R.T., Wets, R.J.-B.: *Variational Analysis*. Springer, New York (1998)
3. Demyanov, V.F., Rubinov, A.M.: *Constructive Nonsmooth Analysis*. Peter Lang, Frankfurt (1995)
4. Hiriart-Urruty, J.-B.: Generalized differentiability, duality and optimization for problems dealing with difference of convex functions. In: Ponstein, J. (ed.) *Convexity and Duality in Optimization*. Lecture Notes in Economics and Mathematical Systems, vol. 256, pp. 37–70. Springer, Heidelberg (1985)
5. Hiriart-Urruty, J.-B., Lemaréchal, C.: *Convex Analysis and Minimization Algorithms*. Springer, Heidelberg (1993)
6. Clarke, F.H.: *Optimization and Nonsmooth Analysis*. Wiley, New York (1983)
7. Ioffe, A.D., Tihomirov, V.M.: *Theory of Extremal Problems*. Elsevier Science Ltd., Amsterdam (1979)
8. Alekseev, V.M., Tikhomirov, V.M., Fomin, S.V.: *Optimal Control*. Springer, New York (1987)
9. Hiriart-Urruty, J.-B.: *Optimisation et Analyse Convex*. Presses Universitaires de France, Paris (1998)
10. Nocedal, J., Wright, S.J.: *Numerical Optimization*. Springer, New York (2006)
11. Bonnans, J.-F., Gilbert, J.C., Lemaréchal, C., Sagastizábal, C.A.: *Numerical Optimization: Theoretical and Practical Aspects*, 2nd edn. Springer, Heidelberg (2006)
12. Izmailov, A.F., Solodov, M.V.: *Newton-Type Methods for Optimization and Variational Problems*. Springer, New York (2014)
13. Burke, J.V.: An exact penalization viewpoint of constrained optimization. *SIAM J. Control Optim.* **29**(4), 968–998 (1991)
14. Borwein, J.M.: Stability and regular points of inequality systems. *J. Optim. Theory Appl.* **48**, 9–52 (1986)
15. Horst, R., Tuy, H.: *Global Optimization Deterministic Approaches*. Springer, Heidelberg (1993)
16. Tuy, H.: D.c. optimization: theory, methods and algorithms. In: Horst, R., Pardalos, P.M. (eds.) *Handbook of Global Optimization*, pp. 149–216. Kluwer Academic Publisher, Dordrecht (1995)
17. Strekalovsky, A.S.: On solving optimization problems with hidden nonconvex structures. In: Rassias, T.M., Floudas, C.A., Butenko, S. (eds.) *Optimization in Science and Engineering*, pp. 465–502. Springer, New York (2014)
18. Strekalovsky, A.S.: On local search in D.C. optimization problems. *Appl. Math. Comput.* **255**, 73–83 (2015)

19. Strekalovsky, A.S., Gruzdeva, T.V.: Local search in problems with nonconvex constraints. *Comput. Math. Math. Phys.* **47**(3), 381–396 (2007)
20. Strekalovsky, A.S.: On problem of global extremum. *Proc. USSR Acad. Sci.* **292**(5), 1062–1066 (1987)
21. Strekalovsky, A.S.: *Elements of Nonconvex Optimization*. Nauka, Novosibirsk (2003). [In Russian]
22. Strekalovsky, A.S.: Minimizing sequences in problems with D.C. constraints. *Comput. Math. Math. Phys.* **45**(3), 418–429 (2005)
23. Walsh, G.R.: *Methods of Optimization*. Wiley, New York (1975)