A counter-example to “Minimal distance between two non-convex surfaces”

M. D. Voisei[*] and C. Zălinescu[†]

The aim of this note is to show that the main results in [2] are false and discuss the manner in which the minimization problem in [2] is investigated while presenting in a particular case a thorough study of the mentioned optimization problem. To achieve our main goal we provide counterexamples.

In [2] one considers the problem

“$(\mathcal{P}): \quad \min\left\{P(x,y) = \frac{1}{2}\|x-y\|^2 : \ h(x)=0, \ g(y)=0\right\}, \quad (1)$

where $h(x) : \mathcal{X}_a \to \mathbb{R}$ and $g(y) : \mathcal{Y}_a \to \mathbb{R}$ are two given smooth functions, well-defined on their domain $\mathcal{X}_a \subset \mathbb{R}^3$ and $\mathcal{Y}_a \subset \mathbb{R}^3$, respectively. Without loss of generality, we assume that $\mathcal{U}_a := \mathcal{X}_a \times \mathcal{Y}_a$ is a closed convex set in $\mathbb{R}^3 \times \mathbb{R}^3$. However, the feasible set $\mathcal{U}_c \subset \mathcal{U}_a$, defined by

$\mathcal{U}_c = \{(x,y) \in \mathcal{U}_a \mid h(x)=0, \ g(y)=0\}$

is, in general, non-convex.

By introducing Lagrange multipliers $\lambda, \mu \in \mathbb{R}$ to relax the two equality constraints in $\mathcal{U}_c$, the classical Lagrangian associated with the constrained problem $(\mathcal{P})$ is

$L(x,y,\lambda,\mu) = \frac{1}{2}\|x-y\|^2 + \lambda h(x) + \mu g(y). \quad (2)$

If both $h(x)$ and $g(y)$ are convex functions, the problem $(\mathcal{P})$ is a convex quadratic minimization with convex constraints. In this case the Lagrangian is a saddle function, i.e. $L(x,y)$ is convex in the primal variables x, and y, etc.”

Hence, in the authors' opinion, for h and g convex functions, L is convex in (x,y) for all $\lambda, \mu \in \mathbb{R}$. Of course, this is false as seen taking $h = \|\cdot\| - 1$, $g = \|\cdot\|^2 - 2$ for $\lambda, \mu < -\frac{1}{2}$.

Later on one says in [2] that “In this article, we assume that $h(x)$ is a convex function, while $g(y)$ is a non-convex polynomial of degree four given by

$g(y) = \frac{1}{2}\alpha\left(\frac{1}{2}\|y-c\|^2 - \eta\right)^2 - f^T(y-c), \quad (8)$

where $\alpha, \eta \geq 0$ are given constant, c and f are two given vectors in $\mathbb{R}^3$”. Moreover, “the vectors c and f are properly chosen so that these two surfaces $h(x)=0$ and $g(y)=0$ are disjoint”.

To the above problem one associates “the Gao–Strang *total complementary function* form (see [19]):

$$\begin{aligned}\Xi(x,y,\lambda,\mu,\varsigma) &:= \frac{1}{2}\|x-y\|^2 + \lambda h(x) + \mu\left(\Lambda(y)\varsigma - V^*(\varsigma) - f^T(y-c)\right) \\ &= \frac{1}{2}\|x-y\|^2 + \lambda h(x) + \mu\left(\frac{1}{2}\|y-c\|^2\varsigma - \frac{1}{2\alpha}\varsigma^2 - \eta\varsigma - f^T(y-c)\right). \quad (11)\end{aligned}$$

Through this total complementary function, the canonical dual function can be defined by

$P^d(\lambda,\mu,\varsigma) = \left\{\Xi(x,y,\lambda,\mu,\varsigma) : \frac{\partial \Xi}{\partial x} = 0, \ \frac{\partial \Xi}{\partial y} = 0\right\}. \quad (12)$

[*]Towson University, U.S.A., email: mvoisei@towson.edu.

[†]University “Al.I.Cuza” Iaşi, Faculty of Mathematics, Romania, email: zalinesc@uaic.ro.

Let $\mathcal{S}_a$ be the so called dual feasible space such that the canonical dual function P^d is well defined by (12). Then the dual problem can be proposed as the following:

$(\mathcal{P}^d): \quad \max\left\{P^d(\lambda,\mu,\varsigma) : (\lambda,\mu,\varsigma)\in\mathcal{S}_a\right\}. \quad (13)$

THEOREM 1 (Complementary-dual principle) If $(\overline{x},\overline{y},\overline{\lambda},\overline{\mu},\overline{\varsigma})$ is a critical point of $\Xi(x,y,\lambda,\mu,\varsigma)$, then $(\overline{x},\overline{y})$ is a critical point of $P(x,y)$, $(\overline{\lambda},\overline{\mu},\overline{\varsigma})$ is a critical point of $P^d(\lambda,\mu,\varsigma)$, and

$P(\overline{x},\overline{y}) = \Xi(\overline{x},\overline{y},\overline{\lambda},\overline{\mu},\overline{\varsigma}) = P^d(\overline{\lambda},\overline{\mu},\overline{\varsigma}). \quad (14)$"

The reference [19] above is our reference [3]. The next result is also stated in [2, p. 708].

"THEOREM 2 Suppose that $h(x)$ is given convex function and $(\overline{x},\overline{y},\overline{\lambda},\overline{\mu},\overline{\varsigma})$ is a critical point of $\Xi(x,y,\lambda,\mu,\varsigma)$. If $\overline{\varsigma}\in\mathcal{S}_c$, then $(\overline{\lambda},\overline{\mu},\overline{\varsigma})$ is a global maximizer of P^d on $\mathcal{S}_c$ and $(\overline{x},\overline{y})$ is a global minimizer of P on $\mathcal{U}_c$, *i.e.*,

$P(\overline{x},\overline{y}) = \min_{(x,y)\in\mathcal{U}_c} P(x,y) = \max_{(\lambda,\mu,\varsigma)\in\mathcal{S}_c} P^d(\lambda,\mu,\varsigma) = P^d(\overline{\lambda},\overline{\mu},\overline{\varsigma}). \quad (17)$"

Here "$\mathcal{S}_c = \{(\lambda,\mu,\varsigma)\in\mathcal{S}_a \mid \lambda\in\mathcal{L}_c,\ 1+\mu\varsigma\geq 0\}$", where "$\mathcal{L}_c = \{\lambda\in\mathbb{R} \mid \frac{1}{2}\|x\|^2+\lambda h(x)$ is convex on $\mathcal{X}_a\}$ (4)".

In the text of the previous theorem there is a miss-print, namely "$\overline{\varsigma}\in\mathcal{S}_c$"; this must be corrected to $(\overline{\lambda},\overline{\mu},\overline{\varsigma})\in\mathcal{S}_c$.

The ambiguous statement of this theorem can be understood in two different ways:

(A) If $(\overline{x},\overline{y},\overline{\lambda},\overline{\mu},\overline{\varsigma})$ is a critical point of $\Xi(x,y,\lambda,\mu,\varsigma)$ with $(\overline{\lambda},\overline{\mu},\overline{\varsigma})\in\mathcal{S}_c$ then $(\overline{x},\overline{y})\in\mathcal{U}_c$, $(\overline{\lambda},\overline{\mu},\overline{\varsigma})$ is a global maximizer of P^d on $\mathcal{S}_c$, and $(\overline{x},\overline{y})$ is a global minimizer of P on $\mathcal{U}_c$, *i.e.*, $P(\overline{x},\overline{y}) = \min_{(x,y)\in\mathcal{U}_c} P(x,y) = \max_{(\lambda,\mu,\varsigma)\in\mathcal{S}_c} P^d(\lambda,\mu,\varsigma) = P^d(\overline{\lambda},\overline{\mu},\overline{\varsigma})$,

or

(B) If $(\overline{x},\overline{y},\overline{\lambda},\overline{\mu},\overline{\varsigma})$ is a critical point of $\Xi(x,y,\lambda,\mu,\varsigma)$ with $(\overline{\lambda},\overline{\mu},\overline{\varsigma})\in\mathcal{S}_c$ and $(\overline{x},\overline{y})\in\mathcal{U}_c$ then $(\overline{\lambda},\overline{\mu},\overline{\varsigma})$ is a global maximizer of P^d on $\mathcal{S}_c$ and $(\overline{x},\overline{y})$ is a global minimizer of P on $\mathcal{U}_c$, *i.e.*, $P(\overline{x},\overline{y}) = \min_{(x,y)\in\mathcal{U}_c} P(x,y) = \max_{(\lambda,\mu,\varsigma)\in\mathcal{S}_c} P^d(\lambda,\mu,\varsigma) = P^d(\overline{\lambda},\overline{\mu},\overline{\varsigma})$.

Clearly (A) $\Rightarrow$ (B). Thus the existence of a counter-example for (S) proves that in any form [2, Th. 2] is false.

Another ambiguity is given by the following phrase mentioned above: "Let $\mathcal{S}_a$ be the so called dual feasible space such that the canonical dual function P^d is well defined by (12)". Looking at the examples considered in the paper, one can interpret that $\mathcal{S}_a$ is the set of $(\lambda,\mu,\varsigma)\in\mathbb{R}^3$ for which the system $[D_x\Xi(x,y,\lambda,\mu,\varsigma)=0,\ D_y\Xi(x,y,\lambda,\mu,\varsigma)=0]$ has a unique solution. Another possible interpretation is that $\mathcal{S}_a$ is the set of $(\lambda,\mu,\varsigma)\in\mathbb{R}^3$ for which the set

$$\{\Xi(x,y,\lambda,\mu,\varsigma) \mid D_x\Xi(x,y,\lambda,\mu,\varsigma)=0,\ D_y\Xi(x,y,\lambda,\mu,\varsigma)=0\}$$

is a singleton. Denote by $\mathcal{S}_a^0$ the first set and by $\mathcal{S}_a^t$ the second one; of course, $\mathcal{S}_a^0\subset\mathcal{S}_a^t$.

Of course, the dimension 3 is not essential in the framework and statements above. The computation below is done for $n\geq 1$. We consider a particular case; more precisely we take $\alpha:=\eta:=1$, $c:=(1,0,\ldots,0)\in\mathbb{R}^n$ and $f:=\gamma c$ for some $\gamma>0$. Moreover (as in one of the applications in [2]), we take $h(x):=\frac{1}{2}\|x\|^2-\frac{1}{2}$, for $x\in\mathbb{R}^n$. Thus we have $\mathcal{X}_a=\mathcal{Y}_a=\mathbb{R}^n$ and $\mathcal{U}_c=\mathcal{X}_c\times\mathcal{Y}_c$ with

$$\mathcal{X}_c := \{x\in\mathbb{R}^n \mid \|x\|=1\}, \quad \mathcal{Y}_c := \left\{y\in\mathbb{R}^n \mid \tfrac{1}{2}\left(\tfrac{1}{2}\|y-c\|^2-1\right)^2-\gamma\langle c,y-c\rangle=0\right\}.$$

Note that $\mathcal{X}_c \cap \mathcal{Y}_c = \emptyset$. Indeed, assume that $x = (x_1, \ldots, x_n) \in \mathcal{X}_c \cap \mathcal{Y}_c$. Then $x_1^2 + \cdots + x_n^2 = 1$, whence $x_1 \in [-1, 1]$, and

$$\gamma(x_1 - 1) = \tfrac{1}{2}\left(\tfrac{1}{2}\left((x_1-1)^2 + x_2^2 + \cdots + x_n^2\right) - 1\right)^2 = \tfrac{1}{2}\left(\tfrac{1}{2}(2 - 2x_1) - 1\right)^2 = \tfrac{1}{2}x_1^2 \geq 0, \quad (1)$$

whence $x_1 \geq 1$. Therefore, $x_1 = 1$; from (1) we get the contradiction $0 = \frac{1}{2}$.

With the preceding data we get

$$\Xi(x, y, \lambda, \mu, \varsigma) := \tfrac{1}{2}\|x - y\|^2 + \lambda\big(\tfrac{1}{2}\|x\|^2 - \tfrac{1}{2}\big) + \mu\big(\tfrac{1}{2}\|y - c\|^2 \varsigma - \tfrac{1}{2}\varsigma^2 - \varsigma - \gamma\langle c, y - c\rangle\big)$$

for $x, y \in \mathbb{R}^n$ and $\lambda, \mu, \varsigma \in \mathbb{R}$.

For the problem $(\mathcal{P})$ set up this way note that $\mathcal{X}_c$ and $\mathcal{Y}_c$ are disjoint nonempty compact sets and so P attains its (non-zero) infimum and supremum on $\mathcal{U}_c$. Since $\nabla h(x) = x \neq 0$ for every $x \in \mathcal{X}_c$ and $\nabla g(y) = \big(\frac{1}{2}\|y - c\|^2 - 1\big)(y - c) - \gamma c \neq 0$ for every $y \in \mathcal{Y}_c$, for any local extremum point (x, y) of P on $\mathcal{U}_c$, by the the classical necessary optimality condition, there exist Lagrange multipliers $\lambda, \mu \in \mathbb{R}$ such that

$$x - y + \lambda x = 0, \quad y - x + \mu\Big(\big(\tfrac{1}{2}\|y - c\|^2 - 1\big)(y - c) - \gamma c\Big) = 0;$$

from which it follows that $\lambda \neq 0$, $\mu \neq 0$, and

$$P(x, y) = \tfrac{1}{2}\lambda^2.$$

Setting $\varsigma := \frac{1}{2}\|y - c\|^2 - 1$, because $x \in \mathcal{X}_c$ and $y \in \mathcal{Y}_c$, we obtain that $(x, y, \lambda, \mu, \varsigma)$ is a critical point of Ξ, that is,

$$D_x\Xi(x, y, \lambda, \mu, \varsigma) = x - y + \lambda x = 0, \quad (2)$$

$$D_y\Xi(x, y, \lambda, \mu, \varsigma) = y - x + \mu\varsigma(y - c) - \mu\gamma c = 0, \quad (3)$$

$$D_\lambda\Xi(x, y, \lambda, \mu, \varsigma) = \tfrac{1}{2}\|x\|^2 - \tfrac{1}{2} = 0, \quad (4)$$

$$D_\mu\Xi(x, y, \lambda, \mu, \varsigma) = \tfrac{1}{2}\|y - c\|^2\varsigma - \tfrac{1}{2}\varsigma^2 - \varsigma - \gamma\langle c, y - c\rangle = 0, \quad (5)$$

$$D_\varsigma\Xi(x, y, \lambda, \mu, \varsigma) = \mu\big(\tfrac{1}{2}\|y - c\|^2 - \varsigma - 1\big) = 0. \quad (6)$$

Hence for every local extremum point (x, y) of P on $\mathcal{U}_c$ there exists $(\lambda, \mu, \varsigma) \in \mathbb{R}^3$, with $\mu \neq 0$, $\lambda \neq 0$, such that $(x, y, \lambda, \mu, \varsigma)$ is a critical point of Ξ.

For the study of the converse of the previous fact let us first note that, because $\mathcal{X}_c$, Y_c are disjoint,

I. every critical point (x, y, λ, μ) of L has $(x, y) \in \mathcal{U}_c$, $\mu \neq 0$, $\lambda \neq 0$,

II. for every critical point $(x, y, \lambda, \mu, \varsigma)$ of Ξ we have $\mu \neq 0$ iff $\lambda \neq 0$ iff $(x, y) \in \mathcal{U}_c$,

III. and in particular every critical point $(x, y, \lambda, \mu, \varsigma)$ of Ξ with $(x, y) \in \mathcal{U}_c$ has $\mu \neq 0$, $\lambda \neq 0$.

Indeed, while I is straightforward and III is a consequence of II, for II the argument is as follows: if $\mu = 0$ then from (3), $x = y$ whence $\lambda x = 0$, by (2). But $x \neq 0$ since $\|x\| = 1$, by (4). Therefore $\lambda = 0$. If $\lambda = 0$ then from (2), $x = y$ and so $(x, y) \notin \mathcal{U}_c$ because $\mathcal{X}_c \cap \mathcal{Y}_c = \emptyset$. It remains to prove that $(x, y) \notin \mathcal{U}_c$ implies $\mu = 0$. We use its counter-positive form, namely, if $\mu \neq 0$ then from (6), $\varsigma := \frac{1}{2}\|y-c\|^2 - 1$ which combined with (5) provides $y \in \mathcal{Y}_c$, because whenever $\mu \neq 0$ (5) and (6) can be equivalently restated as

$$\varsigma = \tfrac{1}{2}\|y-c\|^2 - 1, \quad \tfrac{1}{2}\big(\tfrac{1}{2}\|y-c\|^2 - 1\big)^2 = \gamma\langle c, y-c\rangle, \tag{7}$$

Since $x \in \mathcal{X}_c$ (see (4)) we proved that $(x,y) \in \mathcal{U}_c$.

In the process of examining the possible converses we present a comparison of critical points of L and Ξ, the two main functionals involved in the study of local extrema from the classical and the authors of [2] points of view respectively.

To this end we introduce the sets

$$\begin{aligned} \mathcal{S}_{pc} &:= \big\{(\lambda,\mu,\varsigma) \in \mathbb{R}^3 \mid \exists (x,y) \in \mathbb{R}^n \times \mathbb{R}^n : (x,y,\lambda,\mu,\varsigma) \text{ is a critical point of } \Xi\big\}, \\ \mathcal{S}_{pc}^r &:= \big\{(\lambda,\mu,\varsigma) \in \mathbb{R}^3 \mid \exists (x,y) \in \mathcal{U}_c : (x,y,\lambda,\mu,\varsigma) \text{ is a critical point of } \Xi\big\}, \\ \mathcal{S}_l &:= \big\{(\lambda,\mu) \in \mathbb{R}^2 \mid \exists (x,y) \in \mathbb{R}^n \times \mathbb{R}^n : (x,y,\lambda,\mu) \text{ is a critical point of } L\big\}. \end{aligned}$$

The above considerations show that

$$\mathcal{S}_l = \big\{(\lambda,\mu) \in \mathbb{R}\times\mathbb{R} \mid \exists \varsigma \in \mathbb{R} : (\lambda,\mu,\varsigma) \in \mathcal{S}_{pc}^r\big\}. \tag{8}$$

Indeed, if $(\lambda,\mu) \in \mathcal{S}_l$ then (x,y,λ,μ) is a critical point of L for some $(x,y) \in \mathcal{U}_c$ and as seen above $(\lambda,\mu,\varsigma) \in \mathcal{S}_{pc}^r$ for $\varsigma := \frac{1}{2}\|y-c\|^2 - 1$. Conversely, if $(\lambda,\mu,\varsigma) \in \mathcal{S}_{pc}^r$ then there is $(x,y) \in \mathcal{U}_c$ such that $(x,y,\lambda,\mu,\varsigma)$ is a critical point of Ξ; whence, according to II, $\mu \neq 0$, followed by $\varsigma := \frac{1}{2}\|y-c\|^2 - 1$ (see (6)). Combined with (5) this shows that (x,y,λ,μ) is a critical point of L, that is, $(\lambda,\mu) \in \mathcal{S}_l$.

Clearly, $\mathcal{S}_{pc}^r \subset \mathcal{S}_{pc}$. Also, by the argument above we have

$$\mathcal{S}_{pc}^r = \mathcal{S}_{pc} \cap (\mathbb{R}\times\mathbb{R}^*\times\mathbb{R}) = \mathcal{S}_{pc} \cap (\mathbb{R}^*\times\mathbb{R}^*\times\mathbb{R}) = \mathcal{S}_{pc} \cap (\mathbb{R}^*\times\mathbb{R}\times\mathbb{R}), \tag{9}$$

$$\mathcal{S}_{pc} \subset \mathcal{S}_{pc}^r \cup (\{0\}\times\{0\}\times\mathbb{R}), \quad \mathcal{S}_{pc}^r \cap (\{0\}\times\{0\}\times\mathbb{R}) = \emptyset, \tag{10}$$

where $\mathbb{R}^* := \mathbb{R}\setminus\{0\}$.

To describe in more details the difference between $\mathcal{S}_{pc}$ and $\mathcal{S}_{pc}^r$ we have the following result.

Lemma 1 *Assume that all the above assumptions on α, η, c, γ hold. Then $(x,y,0,0,\varsigma)$ is a critical point of Ξ if and only if $y = x \in \mathcal{X}_c$ and ς is a real solution of*

$$\tfrac{1}{2}\varsigma^2 + \varsigma x_1 + \gamma(x_1 - 1) = 0, \tag{11}$$

where $x = (x_1, x_2, \ldots, x_n)$. In particular $(c,c,0,0,0)$ is a critical point of Ξ and $(c,c) \notin \mathcal{U}_c$.

Proof. It is easily checked that for every $x \in \mathcal{X}_c$ and ς a real solution of (11) the point $(x,x,0,0,\varsigma)$ satisfies (2)–(6), that is $(x,x,0,0,\varsigma)$ is critical for Ξ, where (11) is a restatement of (5). The converse is immediate, as well as the fact that $(c,c,0,0,0)$ is critical for Ξ. ■

Let us describe the functional P^d and its domain $\mathcal{S}_a$. Recall the definition of P^d ([2, (12)] mentioned above) and use (2) and (3) to obtain equivalently that

$$y = (1+\lambda)x, \quad (\lambda + \mu\varsigma + \lambda\mu\varsigma)x = \mu(\gamma+\varsigma)c. \tag{12}$$

It follows that

$$\mathcal{S}_a^0 = \left\{(\lambda,\mu,\varsigma)\in\mathbb{R}^3 \mid \lambda+\mu\varsigma+\lambda\mu\varsigma\neq 0\right\}.$$

Clearly, for $(\lambda,\mu,\varsigma)\in\mathcal{S}_a^0$ we have that

$$x = \frac{\mu(\gamma+\varsigma)}{\lambda+\mu\varsigma+\lambda\mu\varsigma}c,\quad y = \frac{\mu(\gamma+\varsigma)(1+\lambda)}{\lambda+\mu\varsigma+\lambda\mu\varsigma}c \tag{13}$$

and so

$$P^d(\lambda,\mu,\varsigma) = \Xi\left(\tfrac{\mu(\gamma+\varsigma)}{\lambda+\mu\varsigma+\lambda\mu\varsigma}c, \tfrac{\mu(\gamma+\varsigma)(1+\lambda)}{\lambda+\mu\varsigma+\lambda\mu\varsigma}c, \lambda,\mu,\varsigma\right).$$

If $\lambda+\mu\varsigma+\lambda\mu\varsigma = 0$ then from (12), $\mu(\gamma+\varsigma)=0$ because $c\neq 0$. Then either

- $\mu = 0$ which implies $\lambda = 0$ and (by (2)) $x = y$ making $\Xi(x,y,\lambda,\mu,\varsigma) = 0$, and so
$$P^d(0,0,\varsigma) = 0,$$
- or $\mu\neq 0$, $\varsigma = -\gamma$, $\lambda\notin\{-1,0\}$, $\mu = \frac{\lambda}{\gamma(1+\lambda)}$. In this case $x = \frac{1}{1+\lambda}y$, and so
$$\begin{aligned}\Xi\left(\tfrac{1}{1+\lambda}y, y, \lambda, \tfrac{\lambda}{\gamma(1+\lambda)}, -\gamma\right) &= \tfrac{\lambda}{2(1+\lambda)}\|y\|^2 - \tfrac{\lambda}{2} + \tfrac{\lambda}{1+\lambda}\left(1-\tfrac12\|y-c\|^2 - \tfrac12\gamma - \langle c, y-c\rangle\right)\\ &= -\tfrac{\lambda}{2(\lambda+1)}(\lambda+\gamma-2);\end{aligned}$$
hence
$$P^d(\lambda, \tfrac{\lambda}{\gamma(1+\lambda)}, -\gamma) = -\tfrac{\lambda}{2(\lambda+1)}(\lambda+\gamma-2)$$
is uniquely defined.

Therefore,

$$\mathcal{S}_a^t = \mathcal{S}_a^0 \cup \{(0,0,\varsigma)\mid \varsigma\in\mathbb{R}\} \cup \left\{\left(\lambda, \tfrac{\lambda}{\gamma(1+\lambda)}, -\gamma\right) \mid \lambda\in\mathbb{R}\setminus\{-1,0\}\right\}. \tag{14}$$

The first inclusion below is obvious, while the the second one is an immediate consequence of (10) or II:

$$\mathcal{S}_{pc}\subset\mathcal{S}_a^t,\quad \mathcal{S}_{pc}\cap\mathcal{S}_a^0\subset\mathcal{S}_{pc}^r, \tag{15}$$

Note that $\mathcal{L}_c = \{\lambda\in\mathbb{R}\mid \frac12(\lambda+1)\|x\|^2 - \frac{\lambda}{2}$ is convex on $\mathcal{X}_a\} = [-1,\infty)$. Corresponding to the choices $\mathcal{S}_a^0$ and $\mathcal{S}_a^t$ for $\mathcal{S}_a$, according to the definition of $\mathcal{S}_c$, we get the following options

$$\begin{aligned}\mathcal{S}_c^0 &= \{(\lambda,\mu,\varsigma)\in\mathbb{R}^3\mid \lambda+\mu\varsigma+\lambda\mu\varsigma\neq 0,\ \lambda\geq -1,\ 1+\mu\varsigma\geq 0\},\\ \mathcal{S}_c^t &= \mathcal{S}_c^0\cup\{(0,0,\varsigma)\mid\varsigma\in\mathbb{R}\}\cup\left\{\left(\lambda,\tfrac{\lambda}{\gamma(1+\lambda)},-\gamma\right)\mid \lambda>-1,\lambda\neq 0\right\}.\end{aligned}$$

The critical point $(c,c,0,0,0)$ of Ξ studied in Lemma 1 shows that version (A) of [2, Th. 2] is false in the case $\mathcal{S}_c := \mathcal{S}_c^t$.

Since our goal is to reveal a counter-example for (B) for any choice of $\mathcal{S}_c$, let us find a way to describe the critical points $(x,y,\lambda,\mu,\varsigma)$ of Ξ with $(\lambda,\mu,\varsigma)\in\mathcal{S}_a^0$.

By (15) and (10) we have that $\mu\neq 0$ and $(x,y)\in\mathcal{U}_c$. The second part of (13) yields

$$y - c = \frac{\gamma\mu-\lambda+\lambda\gamma\mu}{\lambda+\mu\varsigma+\lambda\mu\varsigma}c = \sigma c\quad\text{with}\quad \sigma := \frac{\gamma\mu-\lambda+\lambda\gamma\mu}{\lambda+\mu\varsigma+\lambda\mu\varsigma},$$

and so $\langle c, y-c\rangle = \sigma$, $\|y-c\| = |\sigma|$. It follows from (7) that

$$2(\varsigma+1) = \sigma^2, \quad \varsigma^2 = 2\gamma\sigma.$$

Hence $\varsigma^4 = 4\gamma^2\sigma^2 = 8\gamma^2(\varsigma+1)$. For every $\gamma > 0$ the equation $\varsigma^4 = 8\gamma^2(\varsigma+1)$ has two real solutions $\varsigma_1 \in (-1,0)$ and $\varsigma_2 > 0$.

From the above considerations we see that in this case (for given $\gamma > 0$ and c as above):

- $(x, y, \lambda, \mu, \varsigma)$ is a critical point of Ξ with $(\lambda, \mu, \varsigma) \in \mathcal{S}_a^0$ iff $\varsigma^4 = 8\gamma^2(\varsigma+1)$, $\lambda+\mu\varsigma+\lambda\mu\varsigma = \pm\mu(\gamma+\varsigma)$, $\dfrac{\gamma\mu-\lambda+\lambda\gamma\mu}{\lambda+\mu\varsigma+\lambda\mu\varsigma} = \dfrac{\varsigma^2}{2\gamma}$, $\gamma+\varsigma \neq 0$, $\mu \neq 0$, and (13) holds.

In other words if we take $\gamma > 0$ such that $\gamma^4 \neq 8\gamma^2(-\gamma+1)$, pick ς one of the two solutions of $\varsigma^4 = 8\gamma^2(\varsigma+1)$, and solve for λ and μ the two nonlinear systems $\lambda+\mu\varsigma+\lambda\mu\varsigma = \pm\mu(\gamma+\varsigma)$, $\dfrac{\gamma\mu-\lambda+\lambda\gamma\mu}{\lambda+\mu\varsigma+\lambda\mu\varsigma} = \dfrac{\varsigma^2}{2\gamma}$ then $(\pm c, \pm(1+\lambda)c, \lambda, \mu, \varsigma)$ is a critical point of Ξ with $(\lambda, \mu, \varsigma) \in \mathcal{S}_a^0$. Clearly, there are only four possibilities for such a critical point. More precisely, these critical points are $(c, (\frac{\varsigma^2}{2\gamma}+1)c, \frac{\varsigma^2}{2\gamma}, -\frac{\varsigma^2}{\varsigma^3-2\gamma^2}, \varsigma)$, $(-c, (\frac{\varsigma^2}{2\gamma}+1)c, -\frac{\varsigma^2}{2\gamma}-2, -\frac{4\gamma+\varsigma^2}{\varsigma^3-2\gamma^2}, \varsigma)$, for $\varsigma = \varsigma_1, \varsigma_2$ the two real solutions of $\varsigma^4 = 8\gamma^2(\varsigma+1)$. Also, $(\lambda, \mu, \varsigma) \in \mathcal{S}_{pc} \cap \mathcal{S}_a^0 = \mathcal{S}_{pc}^r \cap \mathcal{S}_a^0$ iff $\lambda = \frac{\varsigma^2}{2\gamma}, \mu = -\frac{\varsigma^2}{\varsigma^3-2\gamma^2}$ or $\lambda = -\frac{\varsigma^2}{2\gamma} - 2, \mu = -\frac{4\gamma+\varsigma^2}{\varsigma^3-2\gamma^2}$, for $\varsigma = \varsigma_1, \varsigma_2$.

With the goal in mind of describing the set $\mathcal{S}_{pc}^r \setminus \mathcal{S}_a^0$, let us investigate the existence of critical points $(x, y, \lambda, \mu, \varsigma)$ of Ξ with $(x, y) \in \mathcal{U}_c$ and $(\lambda, \mu, \varsigma) \notin \mathcal{S}_a^0$.

Proposition 2 *Assume that all the above assumptions on α, η, c, γ hold. There exist critical points $(x, y, \lambda, \mu, \varsigma)$ of Ξ with $(x, y) \in \mathcal{U}_c$ and $(\lambda, \mu, \varsigma) \notin \mathcal{S}_a^0$ iff $0 < \gamma \leq 2\sqrt{6}-4$, for $n \geq 2$ or $\gamma = 2\sqrt{6}-4$ for $n = 1$. In this case, for $n = 1$ the only critical points $(x, y, \lambda, \mu, \varsigma)$ of Ξ with $(x, y) \in \mathcal{U}_c$ and $(\lambda, \mu, \varsigma) \notin \mathcal{S}_a^0$ are $(1, \sqrt{6}-1, \sqrt{6}-2, \frac{1+\sqrt{6}}{10}, 4-2\sqrt{6})$ and $(-1, \sqrt{6}-1, -\sqrt{6}, \frac{9+4\sqrt{6}}{10}, 4-2\sqrt{6})$, while for $n \geq 2$, $(x, y, \lambda, \mu, \varsigma)$ is a critical point of Ξ with $(x, y) \in \mathcal{U}_c$ and $(\lambda, \mu, \varsigma) \notin \mathcal{S}_a^0$ iff*

$$\|x\| = 1,\ y = (1+\lambda)x,\ \langle c, y\rangle = \tfrac{1}{2}\gamma+1,\ \lambda = \pm\sqrt{3-\gamma}-1,\ \mu = \tfrac{\lambda}{\gamma(1+\lambda)},\ \varsigma = -\gamma. \tag{16}$$

In particular, for $n = 1$ and $\gamma \neq 2\sqrt{6}-4$ or for $n \geq 2$ and $\gamma > 2\sqrt{6}-4$, $\mathcal{S}_{pc}^r \subset \mathcal{S}_a^0$ while for $n \geq 2$ and $\gamma \in (0, 2\sqrt{6}-4]$, $(\lambda, \mu, \varsigma) \in \mathcal{S}_{pc}^r \setminus \mathcal{S}_a^0$ iff $\lambda = \pm\sqrt{3-\gamma}-1$, $\mu = \frac{\lambda}{\gamma(1+\lambda)}$, $\varsigma = -\gamma$.

Proof. Assume that $(x, y, \lambda, \mu, \varsigma)$ is a critical point of Ξ with $(x, y) \in \mathcal{U}_c$ and $(\lambda, \mu, \varsigma) \notin \mathcal{S}_a^0$. In this case $\lambda+\mu\varsigma+\lambda\mu\varsigma = 0$ and $\mu \neq 0$ (see e.g. II). It follows from (12) that $\varsigma = -\gamma$ because $\|x\| = 1$ and $\mu \neq 0$. Also, $\lambda \notin \{0, -1\}$ since $\gamma > 0$ and $\lambda+\mu\varsigma+\lambda\mu\varsigma = 0$. From (4), (2), (6), and $y \in \mathcal{Y}_c$ or (5) we get $\|x\| = 1$, $y = (1+\lambda)x$, $\|y-c\|^2 = 2(1-\gamma)$, $\langle c, y\rangle = \frac{1}{2}\gamma+1$. Combining these last two equalities with $\|y\|^2 = (1+\lambda)^2$ we find $(1+\lambda)^2 = 3-\gamma$. Therefore $0 < \gamma < 3$, $\gamma \neq 2$, and $\lambda = \pm\sqrt{3-\gamma}-1$. From $\|y\|^2 = 3-\gamma$ and $\langle c, y\rangle = \frac{1}{2}\gamma+1$ we obtain for $n = 1$ that $(\frac{1}{2}\gamma+1)^2 = 3-\gamma$, so $\gamma = 2\sqrt{6}-4$, while for $n \geq 2$ we find $(\frac{1}{2}\gamma+1)^2 \leq 3-\gamma$, from which $\gamma \leq 2\sqrt{6}-4$. Clearly, from $\lambda+\mu\varsigma+\lambda\mu\varsigma = 0$ and $\varsigma = -\gamma$ we get $\mu = \frac{\lambda}{\gamma(1+\lambda)}$, and so (16) holds.

Assume that $n = 1$ and $\gamma = 2\sqrt{6}-4 = -\varsigma$. Since $c = 1$ we have $y = \langle c, y\rangle = \frac{1}{2}\gamma+1 = \sqrt{6}-1$. From $|x| = 1$ we see that $x = \pm 1$. For $x = 1$ we find $\lambda = y/x-1 = -\sqrt{6}$,

$\mu = \frac{\lambda}{\gamma(1+\lambda)} = \frac{1+\sqrt{6}}{10}$, that is the point $(1, \sqrt{6}-1, \sqrt{6}-2, \frac{1+\sqrt{6}}{10}, 4-2\sqrt{6})$ and for $x = -1$ we find $\lambda = y/x-1 = \sqrt{6}-2$, $\mu = \frac{\lambda}{\gamma(1+\lambda)} = \frac{9+4\sqrt{6}}{10}$, that is the point $(-1, \sqrt{6}-1, -\sqrt{6}, \frac{9+4\sqrt{6}}{10}, 4-2\sqrt{6})$. It is easily verifiable that these points fulfill all conditions.

Conversely assume that $n \geq 2$, $0 < \gamma \leq 2\sqrt{6}-4$, and $(x, y, \lambda, \mu, \varsigma)$ satisfies (16). Clearly (2), (3), (4) hold, $x \in \mathcal{X}_c$, $\lambda \notin \{-1, 0\}$, $\mu \neq 0$, and $\lambda + \mu\varsigma + \lambda\mu\varsigma = 0$. From $\langle c, y\rangle = \frac{1}{2}\gamma + 1$ and $\|y\|^2 = (1+\lambda)^2 = 3-\gamma$ we find $\frac{1}{2}\|y-c\|^2 = 1-\gamma = 1+\varsigma$ and $\frac{1}{2}\|y-c\|^2\varsigma - \frac{1}{2}\varsigma^2 - \varsigma - \gamma\langle c, y-c\rangle = 0$, that is, (5), (6) are true and this yields that $y \in \mathcal{Y}_c$ (see e.g. (7)). The proof is complete. ■

Let us now fix $\gamma := \sqrt{6}/96$; of course, $\gamma < 2\sqrt{6} - 4$. For this fixed γ and every $n \geq 1$, we make a complete analysis of critical points $(x, y, \lambda, \mu, \varsigma)$ of Ξ with $\mu \neq 0$ in terms of local extremality properties for problem $(\mathcal{P})$.

The equation $\varsigma^4 = 8\gamma^2(\varsigma + 1)$ becomes $192\varsigma^4 = \varsigma + 1$ and has the solutions $\varsigma_1 = -1/4$ and $\varsigma_2 = 0.2860829239$. Note that $\gamma^4 \neq 8\gamma^2(-\gamma + 1)$. Therefore the four critical points $(x, y, \lambda, \mu, \varsigma)$ of Ξ with $(\lambda, \mu, \varsigma) \in \mathcal{S}_a^0$ are

$$\left(c, (1 + \tfrac{1}{2}\sqrt{6})c, \tfrac{1}{2}\sqrt{6}, \tfrac{48}{13}, -\tfrac{1}{4}\right), \quad \left(-c, (1 + \tfrac{1}{2}\sqrt{6})c, -2 - \tfrac{1}{2}\sqrt{6}), \tfrac{16}{13}(3 + 2\sqrt{6}), -\tfrac{1}{4}\right) \tag{17}$$

and

$$(c, 2.603797322c, 1.603797322, -3.701\,325488, 0.2860829239)\,, \tag{18}$$
$$(-c, 2.603797322c, -3.603797322, -8.317027781, 0.2860829239)\,, \tag{19}$$

found with the following Mathematica code:

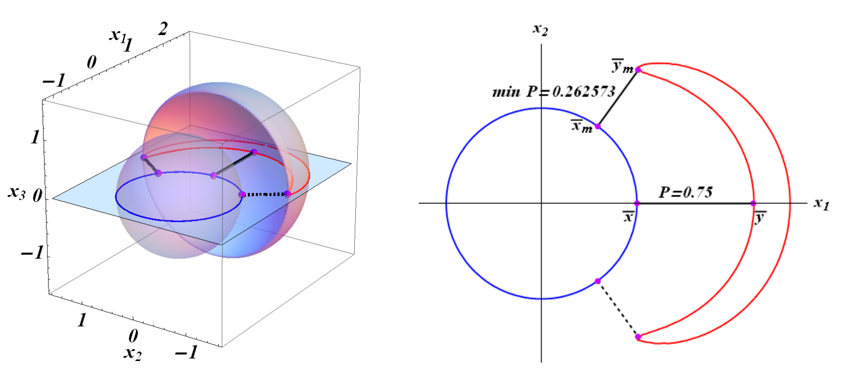
```
Clear[γ, ζ, λ, μ]; γ = √6/96; γ^4!=8γ^2(−γ + 1)
solζe = Solve [ζ^4 − 8γ^2(ζ + 1) == 0, ζ] ;
solζn = NSolve [ζ^4 − 8γ^2(ζ + 1) == 0, ζ, 10] ;
ζ1 = ζ/.solζe[[1]]
ζ2 = ζ/.solζn[[4]]

sys1 = Solve [{2γ(γμ − λ + λγμ)==ζ1^2 (λ + μζ1 + λμ ζ1) ,
 λ + μζ1 + λμ ζ1 == μ (γ + ζ1) , λ + μζ1 + λμ ζ1!=0} , {λ, μ}]
sys2 = Solve [{2γ(γμ − λ + λγμ)==ζ1^2 (λ + μζ1 + λμ ζ1) ,
 λ + μζ1 + λμ ζ1 == −μ (γ + ζ1) , λ + μζ1 + λμ ζ1!=0} , {λ, μ}]
sys3 = NSolve [{2γ(γμ − λ + λγμ)==ζ2^2 (λ + μζ2 + λμ ζ2) ,
 λ + μζ2 + λμ ζ2 == μ (γ + ζ2) , λ + μζ2 + λμ ζ1!=0} , {λ, μ}]
```

Note that only

$$(\overline{x}, \overline{y}, \overline{\lambda}, \overline{\mu}, \overline{\varsigma}) := \left(c, (1 + \tfrac{1}{2}\sqrt{6})c, \tfrac{1}{2}\sqrt{6}, \tfrac{48}{13}, -\tfrac{1}{4}\right) \tag{20}$$

has $(\overline{\lambda}, \overline{\mu}, \overline{\varsigma})$ in $\mathcal{S}_c^0$ (even in $\operatorname{int} \mathcal{S}_c^0$) because the second critical point in (17) and the critical point in (19) have $\lambda < -1$ while the point in (18) fails the condition $1 + \mu\varsigma \geq 0$.

For $n = 1$ we have seen that (16) has no solutions, and so all critical points $(x, y, \lambda, \mu, \varsigma)$ of Ξ with $(x, y) \in \mathcal{U}_c$ (equivalently $\mu \neq 0$) have $(\lambda, \mu, \varsigma) \in \mathcal{S}_a^0$, and so they are given by (17)–(19). In this case $(\overline{x}, \overline{y})$ is indeed the global minimum point of P on $\mathcal{U}_c$ with $P(\overline{x}, \overline{y}) = \frac{1}{2}\overline{\lambda}^2 = \frac{3}{4}$.



(a) $n = 3$ (b) $n = 2$

Figure 1: The constraint sets and critical points

Geometrically the case $n = 2$, corresponds to the section of $\mathcal{X}_c$ and $\mathcal{Y}_c$ with the plane Ox_1x_2 considered in the general case $n \geq 3$. For $n = 3$ the sets $\mathcal{X}_c$ and $\mathcal{Y}_c$ are obtained rotating around Ox_1 the corresponding sets from the case $n = 2$ (see Figure 1 (a), (b)).

Let us prove that, in general, that is, for $n \geq 2$, $(\overline{x}, \overline{y})$ (from (20)) is not a global minimizer of P on $\mathcal{U}_c$. This proves that [2, Th. 2] is false for $\mathcal{S}_a := \mathcal{S}_a^0$; for $\mathcal{S}_a := \mathcal{S}_a^t$ we already observed that [2, Th. 2] in its version (A) is false.

According to the previous geometric observations it suffices to show this only for $n = 2$.

Let $n = 2$. Let us study all the critical points $(x, y, \lambda, \mu, \varsigma)$ of Ξ with $(x, y) \in \mathcal{U}_c$ (equivalently $\mu \neq 0$). Besides the critical points in (17)–(19) (which correspond to $(\lambda, \mu, \varsigma) \in \mathcal{S}_a^0$), there are four other points that correspond to $(\lambda, \mu, \varsigma) \in \mathcal{S}_{pc}^r \setminus \mathcal{S}_a^0$ and are found via system (16); we mention only two of them, the other two being symmetric with respect to the Ox_1 axis. Let $\widetilde{y} := (1 + \gamma/2, \sqrt{2 - 2\gamma - \gamma^2/4}) \in \mathcal{Y}_c$ and $\widetilde{x} := \|\widetilde{y}\|^{-1}\widetilde{y}$. Then these two critical points, determined by (16), are

$$\left(\widetilde{x}, \widetilde{y}, \sqrt{3-\gamma} - 1, \frac{\sqrt{3-\gamma}-1}{\gamma\sqrt{3-\gamma}}, -\gamma\right), \quad \left(-\widetilde{x}, \widetilde{y}, -\sqrt{3-\gamma} - 1, \frac{\sqrt{3-\gamma}+1}{\gamma\sqrt{3-\gamma}}, -\gamma\right).$$

We have that

$$P(\widetilde{x}, \widetilde{y}) = \tfrac{1}{2}\left(\sqrt{3-\gamma} - 1\right)^2 = 0.2625728576 < 0.75 = P(\overline{x}, \overline{y}),$$

which proves that $(\overline{x}, \overline{y})$ is not global minimizer of P on $\mathcal{U}_c$, even though $(\overline{\lambda}, \overline{\mu}, \overline{\varsigma}) \in \mathcal{S}_c^0$. Therefore [2, Th. 2] is false, even with its weaker version (B), for all possible choices $\mathcal{S}_a^0$, $\mathcal{S}_a^t$ of $\mathcal{S}_a$.

Comparing $P(x, y)$ for all critical points of L (we look at $|\lambda|$), we conclude that

$$(\overline{x}_m, \overline{y}_m) := (\widetilde{x}, \widetilde{y}) = ((0.5872184947, 0.8094284647), (1.012757759, 1.395996491))$$

is a global minimizer of P on $\mathcal{U}_c$. Moreover, from Figure 1 (b), we observe that $(\overline{x}, \overline{y})$ is not a local minimizer; it is a local maximizer of P on $\mathcal{U}_c$.

It is worth mentioning the arguments used in the proof of [2, Th. 2]: "By the triality theory developed in [5] we know that if $h(x)$ is convex on $\mathcal{X}_a$, the totally complementary function Ξ is a saddle function over the product space $\mathcal{U}_a \times \mathcal{S}_c$, i.e. it is convex in $x \in \mathcal{X}_a$ and $y \in \mathcal{Y}_a$ and concave in λ, μ and ς such that $(\lambda, \mu, \varsigma) \in \mathcal{S}_c$".

The reference [5] above is Gao's book [1].

In fact, in our example, for a fixed $(\lambda, \mu, \varsigma) \in \mathbb{R}^3$ we have that $\Xi(\cdot, \cdot, \lambda, \mu, \varsigma)$ is convex if and only if $1 + \lambda > 0$, $1 + \mu\varsigma > 0$ and $(1 + \lambda)(1 + \mu\varsigma) \geq 1$. Moreover, for a fixed $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$, $\Xi(x, y, \cdot, \cdot, \cdot)$ cannot be discussed as being concave on $\mathcal{S}_c^0$ or $\mathcal{S}_c^t$ because the sets $\mathcal{S}_c^0$ and $\mathcal{S}_c^t$ are not convex.

In the case $\mathcal{S}_a := \mathcal{S}_a^t$ the equality $\mathcal{S}_{pc}^r = \mathcal{S}_{pc} \cap (\mathbb{R} \times \mathbb{R}^* \times \mathbb{R})$ in (9) shows that (A) $\Leftrightarrow$ (B) provided $\overline{\mu} \neq 0$. In the case $\mathcal{S}_a := \mathcal{S}_a^0$ the inclusion $\mathcal{S}_{pc} \cap \mathcal{S}_a^0 \subset \mathcal{S}_{pc}^r$ in (15) shows again that (A) $\Leftrightarrow$ (B). As seen before, for $\overline{\mu} = 0$, the statements (A) and (B) are not equivalent.

The critical point $(\overline{x}, \overline{y}, \overline{\lambda}, \overline{\mu}, \overline{\varsigma})$ of Ξ provided by (20) has $(\overline{x}, \overline{y}) \in \mathcal{U}_c$; hence $(\overline{x}, \overline{y})$ is not a critical point P, contrary to what is claimed in the first part of the conclusion in [2, Th. 1], and this happens for any of the possible choices $\mathcal{S}_a^0$, $\mathcal{S}_a^t$ of $\mathcal{S}_a$. However, for this critical point of Ξ we have that $(\overline{\lambda}, \overline{\mu}, \overline{\varsigma})$ is a critical point of P^d.

Let us consider $\mathcal{S}_a := \mathcal{S}_a^t$ and take the critical point $(c, c, 0, 0, 0)$ of Ξ. Clearly (c, c) is a critical point of P. For any $\mu \neq 0$ we have, by (14), that $(0, \mu, 0) \notin \mathcal{S}_a^t = \mathcal{S}_a$, and so $\frac{\partial P^d}{\partial \mu}$ does not exist. Hence $(0, 0, 0)$ is not critical for P^d. This shows that this claim of [2, Th. 1] is false when $\mathcal{S}_a$ is taken to be $\mathcal{S}_a^t$.

In conclusion,

- the definition of the set $\mathcal{S}_a$ in [2, p. 708] is ambiguous; from the examples considered in [2] one can deduce that $\mathcal{S}_a$ is our set $\mathcal{S}_a^0$;
- for any interpretation of $\mathcal{S}_a$ the conclusion of [2, Th. 1] is false;
- the statement of [2, Th. 2] is ambiguous;
- in both versions (A) and (B), and for the possible choices $\mathcal{S}_a^0$, $\mathcal{S}_0^t$ of $\mathcal{S}_a$, [2, Th. 2] is false as seen taking the critical point provided in (20);
- for $\mathcal{S}_a := \mathcal{S}_a^0$ (the case considered in the applications from [2]) and $n \geq 2$, one does not find the solutions of problem $(\mathcal{P})$ among the pairs $(\overline{x}, \overline{y})$ provided by the critical points with $(\overline{\lambda}, \overline{\mu}, \overline{\varsigma}) \in \mathcal{S}_c$; in our example this point is even a strict local maximum point;
- the consideration of the function Ξ is useless, at least for the problem studied in [2].

References

[1] Gao, D. Y., Duality principles in nonconvex systems. Theory, methods and applications. Kluwer Academic Publishers, Dordrecht, 2000.

[2] Gao, David Y.; Yang, Wei-Chi, Minimal distance between two non-convex surfaces. Optimization 57 (2008), 705–714.

[3] Gao, D.Y., Strang, G.: Geometric nonlinearity: Potential energy, complementary energy, and the gap function, Quart. Appl. Math. 47 (1989), 487–504.