On unconstrained optimization problems solved using CDT and triality theory

C. Zălinescu
Institute of Mathematics “Octav Mayer”, Iasi, Romania

Abstract DY Gao solely or together with some of his collaborators applied his Canonical duality theory (CDT) for solving a class of unconstrained optimization problems, getting the so-called “triality theorems”. Unfortunately, the “double-min duality” from these results published before 2010 revealed to be false, even if in 2003 DY Gao announced that “certain additional conditions” are needed for getting it. After 2010 DY Gao together with some of his collaborators published several papers in which they added additional conditions for getting “double-min” and “double-max” dualities in the triality theorems. The aim of this paper is to treat rigorously this kind of problems and to discuss several results concerning the “triality theory” obtained up to now.

1 Introduction

In the preface of the book Canonical Duality Theory. Advances in Mechanics and Mathematics, vol 37, Springer, Cham (2017), edited by DY Gao, V Latorre and N Ruan, one says:

“Canonical duality theory is a breakthrough methodological theory that can be used not only for modeling complex systems within a unified framework, but also for solving a large class of challenging problems in multidisciplinary fields of engineering, mathematics, and sciences. ...

This theory is composed mainly of

(1) a canonical dual transformation, which can be used to formulate perfect dual problems without duality gap;

(2) a complementary-dual principle, which solved the open problem in finite elasticity and provides a unified analytical solution form for general nonconvex/nonsmooth/discrete problems;

(3) a triality theory, which can be used to identify both global and local optimality conditions and to develop powerful algorithms for solving challenging problems in complex systems.”

In the period 2009–2013 we published several papers in which we showed, mainly providing counterexamples, that practically all results by DY Gao and his collaborators called “triality theorem” and published or submitted until 2010 are false. Moreover, in the case in which the dual function has one variable, we showed in [30] that the “double-min duality” in the “triality theorem” might be true only when the primal function has also one variable. As a result, DY Gao and C Wu in [16] (and [17], [19]), for a particular class of unconstrained
problems, showed that the “double-min duality” is true only when the number of variables of the primal and dual functions are equal; they treat the general case in [18] (and [20]).

It is our aim in this work to present rigorously this “methodological theory” for unconstrained optimization problems in finite dimensional spaces. It is not the most general framework, but it covers all the situations met in the examples provided in DY Gao and his collaborators’ works on unconstrained optimization problems in finite dimensions. We also point out some drawbacks and not convincing arguments from some of those papers.

2 Preliminaries

We study the following unconstrained minimization problem

\((P)\) \(\min f(x) \text{ s.t. } x \in \mathbb{R}^n\)

where \(f := q_0 + V \circ q\) with \(q(x) := (q_1(x), \ldots, q_m(x))^T\), \(q_i (i \in \{0, m\})\) being quadratic functions defined on \(\mathbb{R}^n\), and \(V \in \Gamma, \Gamma := \Gamma(\mathbb{R}^m)\) being the class of proper convex lower semicontinuous (lsc for short) functions \(g : \mathbb{R}^m \to \overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}\). Recall that for \(g : \mathbb{R}^m \to \overline{\mathbb{R}}\), \(\text{dom } g := \{y \in \mathbb{R}^m | g(y) < \infty\}\), and \(g\) is proper when \(\text{dom } g \neq \emptyset\) and \(g(y) \neq -\infty\) for \(y \in \mathbb{R}^m\). The Fenchel conjugate \(g^* : \mathbb{R}^m \to \overline{\mathbb{R}}\) of the proper function \(g : \mathbb{R}^m \to \overline{\mathbb{R}}\) is defined by

\[ g^*(\sigma) := \sup\{\langle y, \sigma \rangle - g(y) | y \in \mathbb{R}^m\} = \sup\{\langle y, \sigma \rangle - g(y) | y \in \text{dom } g\} \quad (\sigma \in \mathbb{R}^m), \]

while its subdifferential at \(y \in \text{dom } g\) is

\[ \partial g(y) := \{\sigma \in \mathbb{R}^m | \langle y' - y, \sigma \rangle \leq g(y') - g(y) \forall y' \in \mathbb{R}^m\}, \]

and \(\partial g(y) := \emptyset\) if \(y \notin \text{dom } g\); clearly,

\[ g(y) + g^*(\sigma) \geq \langle y, \sigma \rangle \quad \text{ and } \quad [\sigma \in \partial g(y) \iff g(y) + g^*(\sigma) = \langle y, \sigma \rangle \quad \forall (y, \sigma) \in \mathbb{R}^m \times \mathbb{R}^m]. \]

It is well known that for \(g \in \Gamma\) one has \(g^* \in \Gamma\), and \(\sigma \in \partial g(y)\) iff \(y \in \partial g^*(\sigma)\); moreover, \(\partial g(y) \neq \emptyset\) for every \(y \in \text{ri}(\text{dom } g)\) and \(g(\overline{\mathbb{R}}) = \inf_{y \in \mathbb{R}^m} g(y)\) iff \(0 \notin \partial g(\overline{\mathbb{R}})\). Because \(q_i\) are quadratic functions, \(q_i(x) := \frac{1}{2} \langle x, A_i x \rangle - \langle b_i, x \rangle + c_i\) for \(x \in \mathbb{R}^n\) with \(A_i \in \mathcal{S}_n, b_i \in \mathbb{R}^n\) (seen as column matrices), and \(c_i \in \mathbb{R}\ (i \in \{0, m\})\), where \(\mathcal{S}_n\) denotes the set of \(n \times n\) real symmetric matrices; of course, \(c_0\) can be taken to be \(0\).

Consider the so called “total complementary function” (see [20, p. 134]), “Gao–Strang generalized complementary function” (see [14, p. 42]), “extended Lagrangian” (see [2, p. 275], [21]), associated to \((P)\)

\[ \Xi : \mathbb{R}^n \times \mathbb{R}^m \to \overline{\mathbb{R}}, \quad \Xi(x, \sigma) = q_0(x) + \langle q(x), \sigma \rangle - V^*(\sigma) = L(x, \sigma) - V^*(\sigma), \]

where \(L\) is the (usual) Lagrangian associated to \((q_k)_{k=0}^m\), that is \(L\) is the function

\[ L : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}, \quad L(x, \sigma) := q_0(x) + \langle q(x), \sigma \rangle. \]

It follows that

\[ \Xi(x, \sigma) = \frac{1}{2} \langle x, A(\sigma) x \rangle - \langle b(\sigma), x \rangle + c(\sigma) - V^*(\sigma), \]

where, for \(\sigma_0 := 1\) and \(\sigma := (\sigma_1, \ldots, \sigma_m)^T \in \mathbb{R}^m\),

\[ A(\sigma) := \sum_{k=0}^m \sigma_k A_k, \quad b(\sigma) := \sum_{k=0}^m \sigma_k b_k, \quad c(\sigma) := \sum_{k=0}^m \sigma_k c_k; \]
clearly, \( A(\cdot), b(\cdot), c(\cdot) \) are affine functions. Hence, \( \Xi(\cdot, \sigma) \) is quadratic for each \( \sigma \in \text{dom} V^* \) and \( \Xi(x, \cdot) \) is concave for each \( x \in \mathbb{R}^n \). Since \( V^{**} := (V^*)^* = V \), from the definition of the conjugate of \( V^* \) and \([2]\) we obtain that

\[
f(x) = \sup_{\sigma \in \text{dom} V^*} \Xi(x, \sigma) = \sup_{\sigma \in \text{ri}(\text{dom} V^*)} \Xi(x, \sigma) \quad \forall x \in \mathbb{R}^n, \tag{5}\]

because for a proper convex function \( g : \mathbb{R}^m \to \mathbb{R} \) one has \( g^* = (g + \iota_{\text{ri}(\text{dom} g)})^* \) (see \([23]\) p. 259), where the indicator function \( \iota_C : Z \to \mathbb{R} \) of the subset \( C \) of a nonempty set \( Z \) is defined by \( \iota_C(z) := 0 \) for \( z \in C \) and \( \iota_C(z) := \infty \) for \( z \in Z \setminus C \). Moreover,

\[
\nabla_x \Xi(x, \sigma) = A(\sigma)x - b(\sigma), \quad \nabla^2_{xx} \Xi(x, \sigma) = A(\sigma), \tag{6}
\]

\[
\partial (\Xi(x, \cdot)) (\sigma) = \partial V^*(\sigma) - q(x), \tag{7}
\]

for all \((x, \sigma) \in \mathbb{R}^n \times \text{dom} V^* \). Hence, for \((x, \sigma) \in \mathbb{R}^n \times \text{dom} V^* \) one has

\[
\nabla_x \Xi(x, \sigma) = 0 \iff A(\sigma)x = b(\sigma), \tag{8}
\]

\[
0 \in \partial (\Xi(x, \cdot)) (\sigma) \iff q(x) \in \partial V^*(\sigma) \iff \sigma \in \partial V((q(x)). \tag{9}
\]

Consider the following sets in which \( \sigma \) is taken from \( \mathbb{R}^n \) if not specified otherwise:

\[
Y_0 := \{ \sigma \mid \det A(\sigma) \neq 0 \}, \quad Y^+ := \{ \sigma \mid A(\sigma) > 0 \}, \quad Y^- := \{ \sigma \mid A(\sigma) < 0 \},
\]

\[
Y_{\text{col}} := \{ \sigma \mid b(\sigma) \in \text{Im} A(\sigma) \}, \quad Y^+_{\text{col}} := \{ \sigma \in Y_{\text{col}} \mid A(\sigma) \geq 0 \}, \quad Y^-_{\text{col}} := \{ \sigma \in Y_{\text{col}} \mid A(\sigma) \leq 0 \},
\]

\[
S_0 := Y_0 \cap \text{dom} V^*, \quad S^+ := Y^+ \cap \text{dom} V^*, \quad S^- := Y^- \cap \text{dom} V^*,
\]

\[
S_{\text{col}} := Y_{\text{col}} \cap \text{dom} V^*, \quad S^+_{\text{col}} := Y^+_{\text{col}} \cap \text{dom} V^*, \quad S^-_{\text{col}} := Y^-_{\text{col}} \cap \text{dom} V^*.
\]

Of course, any of the preceding sets might be empty, \( Y_0, Y^+, Y^- \) being always open, and \( Y^+, Y^-, Y_{\text{col}}, Y_{\text{col}} \) being convex, the convexity of the last two sets being proved in \([33]\) Cor. 3. It follows that \( S^+, S^-, S^+_{\text{col}}, S^-_{\text{col}} \) are convex, the first two being open if \( \text{dom} V^* \) is so; moreover \( \text{int} S^+_{\text{col}} \subset S^+ \) (resp. \( \text{int} S^-_{\text{col}} \subset S^- \)) whenever \( S^+_{\text{col}} \neq \emptyset \) (resp. \( S^-_{\text{col}} \neq \emptyset \)). Obviously,

\[
Y^+ \cup Y^- \subset Y_{\text{col}}, \quad Y^+_{\text{col}} \cup Y^-_{\text{col}} \subset Y_{\text{col}}, \quad Y^+ = Y_0 \cap Y^+_{\text{col}}, \quad Y^- = Y_0 \cap Y^-_{\text{col}},
\]

\[
S^+ \cup S^- \subset S_{\text{col}}, \quad S^+_{\text{col}} \cup S^-_{\text{col}} \subset S_{\text{col}}, \quad S^+ = S_0 \cap S^+_{\text{col}}, \quad S^- = S_0 \cap S^-_{\text{col}}.
\]

In \([33]\) we considered a dual function associated to the family \((q_k)_{k \in \mathbb{N}}\), which is denoted by \( D_L \) in this work. More precisely,

\[
D_L : Y_{\text{col}} \to \mathbb{R}, \quad D_L(\sigma) := L(x, \sigma) \text{ with } A(\sigma)x = b(\sigma).
\]

In a similar way, we consider the (dual objective) function \( D \) associated to \((q_k)_{k \in \mathbb{N}}\) and \( V \) defined by

\[
D : S_{\text{col}} \to \mathbb{R}, \quad D(\sigma) := \Xi(x, \sigma) \text{ with } A(\sigma)x = b(\sigma);
\]

hence

\[
D(\sigma) = D_L(\sigma) - V^*(\sigma) \quad \forall \sigma \in S_{\text{col}}. \tag{10}
\]

Setting

\[
x(\sigma) := A(\sigma)^{-1}b(\sigma) := [A(\sigma)]^{-1} \cdot b(\sigma)
\]
for $\sigma \in Y_0$, we obtain that
\[
D(\sigma) = \Xi(x(\sigma), \sigma) = \frac{1}{2} \langle b(\sigma), A(\sigma)^{-1} b(\sigma) \rangle + c(\sigma) - V^*(\sigma) \quad \forall \sigma \in S_0.
\]

From [33, Prop. 4 (i)] we have that $D_L$ is concave and upper semicontinuous (usc) on $Y^+_\col$, and convex and lower semicontinuous (lsc) on $Y^-\col$, and [33, Eq. (9)] holds; moreover $D_L(\sigma)$ is attained at any $x \in \mathbb{R}^n$ such that $A(\sigma)x = b(\sigma)$ whenever $\lambda \in Y^+_\col \cup Y^-\col$, being attained uniquely at $x := x(\sigma)$ for $\sigma \in Y^+ \cup Y^-$. Taking into account (10) we have that
\[
D(\sigma) = \begin{cases} 
\min_{x \in \mathbb{R}^n} \Xi(x, \sigma) & \text{if } \sigma \in S^+_\col, \\
\max_{x \in \mathbb{R}^n} \Xi(x, \sigma) & \text{if } \sigma \in S^-\col,
\end{cases}
\]

the value of $D(\sigma)$ being attained uniquely at $x := x(\sigma)$ when $\sigma \in S^+ \cup S^- (\subset S_0)$; moreover, we have that $D$ is concave and usc on $S^+_\col$ as the sum of two concave and usc functions, while $D$ is a d.c. function (difference of convex functions) on $S^-\col$. In general, $D$ is neither convex nor concave on (the convex set) $S^-\col$. Having in view [33, Eq. (11)] (or by direct calculations), we have that
\[
\frac{\partial D}{\partial \sigma_i}(\sigma) = \frac{1}{2} \langle A(\sigma)^{-1} b(\sigma), A_i A(\sigma)^{-1} b(\sigma) \rangle - \langle b_i, A(\sigma)^{-1} b(\sigma) \rangle + c_i - \frac{\partial V^*}{\partial \sigma_i}(\sigma)
\]
\[
= \frac{1}{2} \langle x(\sigma), A_i x(\sigma) \rangle - \langle b_i, x(\sigma) \rangle + c_i - \frac{\partial V^*}{\partial \sigma_i}(\sigma) = q_i (x(\sigma)) - \frac{\partial V^*}{\partial \sigma_i}(\sigma)
\]
for those $\sigma \in \text{int}S_0$ and $i \in \Gamma m$ for which $\frac{\partial V^*}{\partial \sigma_i}(\sigma)$ exists.

**Proposition 1.** Assume that $V \in \Gamma(\mathbb{R}^m)$ is sublinear. Then $D|_{S^-\col}$ is convex; moreover, $\nabla D(\sigma) = q(x(\sigma))$ for every $\sigma \in S_0 \cap \text{int}(\text{dom } V^*)$.

Proof. Assume that $V$ is sublinear; it follows that $V^* = 0_{\partial V(0)}$. Then $\Xi(x, \sigma) = L(x, \sigma) \in \mathbb{R}$ for $\sigma \in \text{dom } V^*$, and so $\Xi(x, \cdot)|_{\text{dom } V^*}$ is convex because $L(x, \cdot)$ is linear and $\text{dom } V^*$ is a convex set; in particular, $\Xi(x, \cdot)|_{S^-\col}$ is convex because $S^-\col (\subset \text{dom } V^*)$ is convex. Using (11) we obtain that $D|_{S^-\col}$ is convex, too.

Of course, $V^*$ being constant on $\text{dom } V^*$, $\nabla V^*(\sigma) = 0$ for every $\sigma \in \text{int}(\text{dom } V^*)$. Taking into account (12), we obtain that $\nabla D(\sigma) = q(x(\sigma))$ for $\sigma \in S_0 \cap \text{int}(\text{dom } V^*)$ (\subset int $S^-\col$). \qed

Let us denote by $\Gamma_{sc} := \Gamma_{sc}(\mathbb{R}^m)$ the class of those $g \in \Gamma(\mathbb{R}^m)$ which are essentially strictly convex and essentially smooth, that is the class of proper lsc convex functions of Legendre type (see [26, Section 26]). Note that any differentiable and strictly convex function $g : \mathbb{R}^m \to \mathbb{R}$ belongs to $\Gamma_{sc}(\mathbb{R}^m)$; moreover, $\Gamma_{sc}(\mathbb{R})$ consists of those $g \in \Gamma(\mathbb{R})$ which are derivable and strictly convex on $\text{int}(\text{dom } g)$, assumed to be nonempty.

Assume that $g \in \Gamma_{sc}$. Then: $g^* \in \Gamma_{sc}$, dom $\partial g = \text{int}(\text{dom } g)$, and $g$ is differentiable on $\text{int}(\text{dom } g)$; moreover, $\nabla g : \text{int}(\text{dom } g) \to \text{int}(\text{dom } g^*)$ is bijective and continuous with $(\nabla g)^{-1} = \nabla g^*$. In the rest of this section we assume that $V \in \Gamma_{sc}$, and so $V^* \in \Gamma_{sc}$, too. Then, because $V$ is differentiable on $\text{int}(\text{dom } V)$ and $V^*$ is differentiable on $\text{int}(\text{dom } V^*)$, clearly
\[
\nabla f(x) = A_0 x - b_0 + \sum_{i=1}^m \frac{\partial V}{\partial y_i} (q(x)) \cdot (A_i x - b_i) \quad \forall x \in X_0,
\]
\[
\nabla \sigma \Xi(x, \sigma) = q(x) - \nabla V^*(\sigma) \quad \forall (x, \sigma) \in \mathbb{R}^n \times \text{int}(\text{dom } V^*),
\]
where
\[ X_0 := \{ x \in \mathbb{R}^n \mid q(x) \in \text{int}(\text{dom } V) \} \subset \text{dom } f; \] (15)

moreover, it follows that (12) holds for \( \sigma \in \text{int } S_0 \) \([= S_0 \cap \text{int}(\text{dom } V^*)]\) and \( i \in \overline{1,m} \), and so
\[ \nabla D(\sigma') = q(x(\sigma')) - \nabla V^*(\sigma') = \nabla_\sigma \Xi(x(\sigma'), \sigma') \ \forall \sigma' \in S_0 \cap \text{int}(\text{dom } V^*). \] (16)

From (14) and (12) we get
\[ \nabla_\sigma \Xi(x, \sigma) = 0 \iff [\sigma \in \text{int } S_0 \land q(x) = \nabla V^*(\sigma)] \iff [x \in X_0 \land \sigma = \nabla V(q(x))]. \] (17)

From the concavity of \( \Xi(x, \cdot) \) for \( x \in \mathbb{R}^n \) and (17) we obtain the next variant of (3):
\[ f(x) = \sup_{\sigma \in \text{dom } V^*} \Xi(x, \sigma) = \sup_{\sigma \in \text{int}(\text{dom } V^*)} \Xi(x, \sigma) = \Xi(x, \nabla V(q(x))) \ \forall x \in X_0; \] (18)

moreover, using (13) and (14) we obtain that
\[ [x \in X_0 \land \sigma = \nabla V(q(x))] \implies [\nabla f(x) = \nabla_\sigma \Xi(x, \sigma) \land f(x) = \Xi(x, \sigma)]. \] (19)

Furthermore, using (6) and (14), for \( (x, \sigma) \in \mathbb{R}^n \times \text{int}(\text{dom } V^*) \) we have that
\[ \nabla \Xi(x, \sigma) = 0 \iff [x \in X_0 \land \sigma = \nabla V(q(x)) \land A(\sigma)x = b(\sigma)]. \] (20)

3 The case \( \bar{\sigma} \in S_{\text{col}}^+ \)

The preceding considerations yield directly the next result.

Proposition 2 Let \( V \in \Gamma(\mathbb{R}^m) \) and \( (\bar{x}, \bar{\sigma}) \in \mathbb{R}^n \times \text{dom } V^* \).

(i) Assume that \( \nabla_\sigma \Xi(\bar{x}, \bar{\sigma}) = 0 \) and \( q(\bar{x}) \in \partial V^*(\bar{\sigma}) \). Then \( (\bar{x}, \bar{\sigma}) \in \text{dom } f \times S_{\text{col}}, \bar{\sigma} \in \partial V(q(\bar{x})), \) and
\[ f(\bar{x}) = \Xi(\bar{x}, \bar{\sigma}) = D(\bar{\sigma}). \] (21)

(ii) Moreover, assume that \( A(\bar{\sigma}) \succeq 0 \). Then \( \bar{\sigma} \in S_{\text{col}}^+ \) and
\[ f(\bar{x}) = \inf_{x \in \text{dom } f} f(x) = \Xi(\bar{x}, \bar{\sigma}) = \sup_{\sigma \in S_{\text{col}}^+} D(\sigma) = D(\bar{\sigma}); \] (22)

furthermore, if \( \bar{\sigma} \in S^+ \), then \( \bar{x} \) is the unique global solution of problem (P).

Proof. (i) Because \( q(\bar{x}) \in \partial V^*(\bar{\sigma}) \), from (11) and (1) we obtain that
\[ \bar{\sigma} \in \partial V(q(\bar{x})) \land \nabla V(q(\bar{x}))+V^*(\bar{\sigma}) = \langle q(\bar{x}), \bar{\sigma} \rangle, \]
whence \( \bar{x} \in \text{dom } f \) and
\[ f(\bar{x}) = q_0(\bar{x}) + V(q(\bar{x})) = q_0(\bar{x}) + [(q(\bar{x}), \bar{\sigma}) - V^*(\bar{\sigma})] = \Xi(\bar{x}, \bar{\sigma}); \]
hence the first equality in (21) holds. Because \( A(\bar{\sigma})\bar{x} - b(\bar{\sigma}) = \nabla_\sigma \Xi(\bar{x}, \bar{\sigma}) = 0 \), we have that \( \bar{\sigma} \in S_{\text{col}} \), and the second equality in (21) holds by the definition of \( D \). (ii) By (i) we have that (21) holds and \( \bar{\sigma} \in S_{\text{col}}^+ \). Because \( A(\bar{\sigma}) \succeq 0 \) we have that \( \bar{\sigma} \in S_{\text{col}}^+ \) and \( \Xi(\cdot, \bar{\sigma}) \) is convex, while
because $\nabla_x \Xi(x, \sigma) = 0$ we have that $(f(\sigma) = \Xi(x, \sigma) \leq \Xi(x, \sigma) \leq f(x)$ for $x \in \text{dom } f \setminus \{\sigma\}$, the latter inequality being equivalent to

$$q_0(x) + \langle q(x), \sigma \rangle - V^*(\sigma) \leq q_0(x) + V(q(x)),$$

which is true by the Fenchel–Young inequality [that is the inequality in (11)]; furthermore, $\Xi(x, \sigma) < \Xi(x, \sigma)$ when $A(\sigma) > 0$. In particular, $f(\sigma) = \min_{x \in \text{dom } f} f(x)$. Using (13), the inclusion $S^+_{\text{col}} \subset \text{dom } V^*$, obvious inequalities, and (11), we get the following sequence of inequalities:

$$f(\sigma) = \inf_{x \in \text{dom } f} f(x) = \inf_{x \in \text{dom } f} \sup_{\sigma \in \text{dom } V^*} \Xi(x, \sigma) \geq \inf_{x \in \text{dom } f} \sup_{\sigma \in S^+_{\text{col}}} \Xi(x, \sigma) \geq \sup_{\sigma \in S^+_{\text{col}}} \inf_{x \in \text{dom } f} \Xi(x, \sigma) = \sup_{\sigma \in S^+_{\text{col}}} D(\sigma) \geq D(\sigma).$$

The inequalities above and (21) show that (22) holds.

**Proposition 3** Let $V \in \Gamma_{\text{sc}}$ and $(\sigma, \sigma) \in \mathbb{R}^n \times \text{int}(\text{dom } V^*)$.

(i) Assume that $(\sigma, \sigma)$ is a critical point of $\Xi$. Then $(\sigma, \sigma) \in X_0 \times S^+_{\text{col}}, \sigma$ is a critical point of $f$, and (21) holds; moreover, if $\sigma \in S_0$ then $\sigma$ is a critical point of $D$.

(ii) Assume that $(\sigma, \sigma)$ is a critical point of $\Xi$ such that $A(\sigma) > 0$. Then $\sigma \in S^+_{\text{col}}$ and (22) holds; moreover, if $A(\sigma) > 0$ then $\sigma$ is the unique global solution of problem (P).

(iii) Assume that $\sigma \in S_0$ and $\sigma$ is a critical point of $D$. Then $(\sigma, \sigma)$ is a critical point of $\Xi$, where $\sigma := A(\sigma)^{-1}b(\sigma)$; therefore, (i) and (ii) apply.

Proof. Observe first that $\partial V^*(\sigma) = \{\nabla V^*(\sigma)\}$ because $V^*$ is differentiable on $\text{int}(\text{dom } V^*)$.

(i) Since $\nabla \Xi(\sigma, \sigma) = 0$, from (17) and (19) we have that $q(\sigma) = \nabla V^*(\sigma) \in \partial V^*(\sigma), \sigma \in X_0$, and $V(\sigma) = \nabla_z \Xi(\sigma, \sigma) = 0$. Applying Proposition 2 (i) we get the first conclusion of (i). Using (16) we obtain that $\nabla D(\sigma) = 0$ when $\sigma \in S_0$. (ii) As seen in the proof of (i), $q(\sigma) \in \partial V^*(\sigma)$. The conclusion follows using Proposition 2 (ii).

(iii) Using (16) we have that $\nabla_z \Xi(\sigma, \sigma) = \nabla D(\sigma) = 0$. The choice of $\sigma$ implies $\nabla_z \Xi(\sigma, \sigma) = 0$, and so $(\sigma, \sigma)$ is a critical point of $\Xi$.

In the rest of this section we consider the important particular case in which $V := V_J := \iota_{C_J}$ for $J \subseteq \overline{1, m}$, $J^c := \overline{1, m} \setminus J$, and

$$C_J := \{y \in \mathbb{R}^m \mid \forall j \in J : y_j = 0 \land \forall j \in J^c : y_j \leq 0\} \subset \mathbb{R}^m.$$ 

Of course, $C_J$ is a closed convex cone, $V_J \in \Gamma(\mathbb{R}^m)$ is sublinear, and $V_J^* := (V_J)^* = \iota_{\Gamma_J}$, where

$$\Gamma_J := \{\sigma \in \mathbb{R}^m \mid \forall j \in J^c : \sigma_j \geq 0\};$$

hence,

$$\text{int } \Gamma_J := \{\sigma \in \mathbb{R}^m \mid \forall j \in J^c : \sigma_j > 0\} \neq \emptyset.$$ 

For $y, \sigma \in \mathbb{R}^m$ we have that

$$\sigma \in \partial V_J(y) \iff y \in \partial V_J^*(\sigma) \iff [y \in C_J \land \sigma \in \Gamma_J \land \langle y, \sigma \rangle = 0] \iff [\forall j \in J : y_j = 0] \land [\forall j \in J^c : y_j \leq 0, \sigma_j \geq 0, y_j \sigma_j = 0].$$ 

For $y, \sigma \in \mathbb{R}^m$ we have that

$$\sigma \in \partial V_J(y) \iff y \in \partial V_J^*(\sigma) \iff [y \in C_J \land \sigma \in \Gamma_J \land \langle y, \sigma \rangle = 0] \iff [\forall j \in J : y_j = 0] \land [\forall j \in J^c : y_j \leq 0, \sigma_j \geq 0, y_j \sigma_j = 0].$$

(23)

(24)
Note that
\[ C_{1,m} = \{0\}, \quad \Gamma_{1,m} = \mathbb{R}^m, \quad C_0 = \mathbb{R}_-^m := -\mathbb{R}_+^m, \quad \Gamma_0 = \mathbb{R}_+^m, \]
while for \(y, \sigma \in \mathbb{R}^m\) \cite{23}, becomes, respectively,
\[ \sigma \in \partial V_{1,m}(y) \iff y \in \partial V_{1,m}^*(\sigma) \iff y = 0, \]
\[ \sigma \in \partial V_0(y) \iff y \in \partial V_0^*(\sigma) \iff \left[ y \in \mathbb{R}_-^m \land \sigma \in \mathbb{R}_+^m \land \langle y, \sigma \rangle = 0 \right]. \]

For \(J \subset \overline{1,m}\) we get \(f_J := q_0 + V_J \circ q = q_0 + \iota_{X_J}\) and \(\Xi_J(x, \sigma) := L(x, \sigma) - \iota_{\Gamma_J}(\sigma)\), where
\[ X_J := \{x \in \mathbb{R}^n \mid \forall j \in J : q_j(x) = 0 \land [\forall j \in J^c : q_j(x) \leq 0]\} = \text{dom } f_J. \]

So, for \(V := V_J\) the problem \((P)\) becomes the problem \((P_J)\) of minimizing \(q_0\) on \(X_J\); \((P_{1,m})\) is the quadratic problem \((P_e)\) of minimizing \(q_0\) on \(X_e := X_{\overline{1,m}}\), while \((P_0)\) is the quadratic problem \((P_i)\) of minimizing \(q_0\) on \(X_i := \Gamma_0\). These problems are considered in \cite{33}.

The dual function corresponding to \(V_J\) is denoted by \(D_J := D_L|_{Y_{\text{col}}^+}\), where \(Y_{\text{col}}^+ := \Gamma_J \cap Y_{\text{col}}^+\). As observed immediately after getting the formula of \(D\) in \cite{11}, \(D_J\) is concave on \(Y_{\text{col}}^+\), while from Proposition \cite{1} we have that \(D_J\) is convex on \(Y_{-\text{col}}^+ := \Gamma_J \cap Y_{-\text{col}}^+\) because \(V_J\) is sublinear.

**Corollary 4** Let \((\overline{x}, \overline{\sigma}) \in \mathbb{R}^n \times \mathbb{R}^m\) be a \(J\)-LKKT point of \(L\), that is \(\nabla_x L(\overline{x}, \overline{\sigma}) = 0\), and
\[ \left[ \forall j \in J^c : \sigma_j \geq 0 \land \frac{\partial L}{\partial \sigma_j}(\overline{x}, \overline{\sigma}) \leq 0 \land \sigma_j \cdot \frac{\partial L}{\partial \sigma_j}(\overline{x}, \overline{\sigma}) = 0 \right] \land \left[ \forall j \in J : \frac{\partial L}{\partial \sigma_j}(\overline{x}, \overline{\sigma}) = 0 \right]. \tag{25} \]
Then \((\overline{x}, \overline{\sigma}) \in X_J \times Y_{\text{col}}^+\) and \(q_0(\overline{x}) = L(\overline{x}, \overline{\sigma}) = D_L(\overline{\sigma})\). Moreover, assume that \(A(\overline{\sigma}) \geq 0\). Then \(\overline{x} \in Y_{\text{col}}^+\) and
\[ q_0(\overline{x}) = \inf_{x \in X_J} q_0(x) = L(\overline{x}, \overline{\sigma}) = \sup_{\sigma \in Y_{\text{col}}^+} D_L(\sigma) = D_L(\overline{\sigma}); \]
furthermore, if \(A(\overline{\sigma}) > 0\), then \(\overline{x}\) is the unique global minimizer of \(q_0\) on \(X_J\).

**Proof.** From \cite{25} we have that \(\overline{\sigma} \in \Gamma_J\), and so \(\nabla_x \Xi_J(\overline{x}, \overline{\sigma}) = \nabla_x L(\overline{x}, \overline{\sigma}) = 0\). Since
\[ \frac{\partial L}{\partial \sigma_j}(\overline{x}, \overline{\sigma}) = q_j(\overline{x}) \quad \text{for } j \in \overline{1,m}, \]
using again \cite{25} we obtain that \(q(\overline{x}) \in C_J\), whence \(\overline{x} \in X_J\), and \(\langle q(\overline{x}), \overline{\sigma} \rangle = 0\). Using now \cite{23} we obtain that \(q(\overline{x}) \in \partial V^*(\overline{\sigma})\). The conclusion follows now using Proposition \cite{2} for \(V := V_J\). \(\square\)

The variant for maximizing \(q_0\) on \(X_J\) is the following result; it can be obtained from the preceding corollary replacing \(q_0\) by \(-q_0\) and \(\sigma\) by \(-\sigma\) in the definition of \(L\).

**Corollary 5** Let \((\overline{x}, \overline{\sigma}) \in \mathbb{R}^n \times \mathbb{R}^m\) be such that \(\nabla_x L(\overline{x}, \overline{\sigma}) = 0\), and
\[ \left[ \forall j \in J^c : \sigma_j \leq 0 \land \frac{\partial L}{\partial \sigma_j}(\overline{x}, \overline{\sigma}) \leq 0 \land \sigma_j \cdot \frac{\partial L}{\partial \sigma_j}(\overline{x}, \overline{\sigma}) = 0 \right] \land \left[ \forall j \in J : \frac{\partial L}{\partial \sigma_j}(\overline{x}, \overline{\sigma}) = 0 \right]. \tag{26} \]
Then \((\overline{x}, \overline{\sigma}) \in X_J \times (Y_{\text{col}} \cap (-\Gamma_J))\) and \(q_0(\overline{x}) = L(\overline{x}, \overline{\sigma}) = D_L(\overline{\sigma})\). Moreover, assume that \(A(\overline{\sigma}) \leq 0\). Then \(\overline{\sigma} \in Y_{\text{col}}^- := Y_{\text{col}}^- \cap (-\Gamma_J)\) and
\[ q_0(\overline{x}) = \sup_{x \in X_J} q_0(x) = L(\overline{x}, \overline{\sigma}) = \inf_{\sigma \in Y_{\text{col}}^-} D_L(\sigma) = D_L(\overline{\sigma}); \]
furthermore, if \(A(\overline{\sigma}) < 0\), then \(\overline{x}\) is the unique global maximizer of \(q_0\) on \(X_J\).
Corollary 6 Let $(\sigma, \tilde{\sigma}) \in \mathbb{R}^n \times \mathbb{R}^m$ be a critical point of $L$. Then $\sigma \in X_e$, $\tilde{\sigma} \in Y_{\text{col}}$ and $q_0(\sigma) = L(\sigma, \tilde{\sigma}) = D_L(\tilde{\sigma})$. Moreover, if $A(\tilde{\sigma}) \geq 0$, then $\tilde{\sigma} \in Y_{\text{col}}^+$ and $q_0(\sigma) = \inf_{x \in X_e} q_0(x) = L(\sigma, \tilde{\sigma}) = \sup_{\sigma \in Y_{\text{col}}^+} D_L(\sigma) = D_L(\tilde{\sigma})$; if $A(\tilde{\sigma}) \leq 0$, then $\tilde{\sigma} \in Y_{\text{col}}^-$ and $q_0(\sigma) = \sup_{x \in X_e} q_0(x) = L(\sigma, \tilde{\sigma}) = \inf_{\sigma \in Y_{\text{col}}^-} D_L(\sigma) = D_L(\tilde{\sigma})$.

Proof. For the first two assertions one applies Corollary 4 for $J := \mathbb{1}, \mathbb{m}$, while for the third assertion one applies Corollary 5. \qed

Notice that Corollaries 4 and 5 are parts of [33, Prop. 9] and [33, Prop. 12], respectively, while Corollary 6 is [33, Prop. 5 (i)].

In many papers by DY Gao and his collaborators one speaks about “triality theorems” in which, besides the minimax result established for the case $A(\tilde{\sigma}) \geq 0$ (see Proposition 2), one obtains also “bi-duality” results (“double-min duality” and “double-max duality”) established for $A(\tilde{\sigma}) < 0$, that is $\sigma$ and $\tilde{\sigma}$ are simultaneously local minimizers (maximizers) for $f$ on $\text{dom} f$ and for $D$ on $S^-$, respectively.

The next example shows that such triality results are not valid for general $V \in \Gamma(\mathbb{R}^m)$, even for $n = m = 1$. We concentrate on the case $\sigma \in S^-$ of Proposition 2 (i), that is $(\sigma, \tilde{\sigma}) \in \mathbb{R}^n \times \mathbb{R}^m$ is such that $A(\tilde{\sigma})\sigma = b(\tilde{\sigma})$ and $\tilde{\sigma} \in S^- \cap \partial V(q(\tilde{\sigma}))$, and so $\sigma \in q^{-1}(\text{dom} V) = \text{dom} f$.

Example 7 Consider $n := m := 1$, $V := v_{\mathbb{R}_+}$, and $q_0(x) := -\frac{1}{2}x^2 + x$, $q(x) = q_1(x) := \frac{1}{2} (x^2 - 1)$ for $x \in \mathbb{R}$. Then $f := f_0 = q_0 + \iota_{[-1,1]}$, $A(\sigma) = \sigma - 1$, $b(\sigma) = -1$, $c(\sigma) = -\frac{1}{2} \sigma$, whence $L(x, \sigma) = \frac{1}{2}(\sigma - 1)x^2 + x - \frac{1}{2} \sigma$, $Y_{\text{col}} = \mathbb{R} \setminus \{1\}$, $x(\sigma) = 1/(1 - \sigma)$, and so $D_L(\sigma) = \frac{1}{2} \left( \frac{1}{1 - \sigma} - \sigma \right)$, for $\sigma \in \mathbb{R}$; moreover, $D := D_\emptyset = D_L$ on $S_{\text{col}} = [0, 1] \cup (1, \infty)$. For $\sigma = 0 \in S^- \cap S^- = [0, 1]$ we get $\sigma := x(0) = 1$. Clearly, $0 \in \partial V(q(1)) = \partial V(0) (= \mathbb{R}_+)$. Hence the pair $(1, 0)$ verifies the hypothesis of Proposition 2 (i), even more, $(1, 0)$ is a critical point of $L$. However, by direct verification, or applying Corollary 5 we obtain that $\sigma = 1$ is the unique global maximizer of $f$ on $\text{dom} f = [-1, 1]$, while applying [33, Prop. 4 (iv)] we obtain that $\sigma = 0$ is the unique global minimizer of $D_L$ on $Y_{\text{col}}^+ = (-\infty, 1)$, whence $0$ is the unique global minimizer of $D$ on $S^-$. These facts show that “double-min duality” and “double-max duality” are not verified in the present case.

In DY Gao’s works published after 2011 the “triality theorems” are established for $V$ a twice differentiable strictly convex function. Our aim in the sequel is to study the problems of “double-min duality” and “double-max duality” for a special class of functions $V$. First, in the next section, we establish a result on positive semidefinite operators in Euclidean spaces needed for getting our “bi-duality” results.

4 An auxiliary result

In order to study the case when $\sigma \in S^-$, we need the following result which is probably known, but we have not a reference for it.
Proposition 8 Let $X$, $Y$ be nontrivial Euclidean spaces and $H : Y \to X$ be a linear operator with $H^* : X \to Y$ its adjoint. Consider $Q := HH^* := H \circ H^*$, $R := H^*H$, and

$$
\varphi : X \to \mathbb{R}, \quad \varphi(x) := \|H^*x\|^2, \quad \psi : Y \to \mathbb{R}, \quad \psi(y) := \|Hy\|^2.
$$

Then the following assertions hold:

(a) $Q$ and $R$ are self-adjoint positive semi-definite operators, $\ker Q = \ker H^*$, $\text{Im} Q = \text{Im} H$, $\ker R = \ker H$, $\text{Im} R = \text{Im} H^*$; consequently, $H = 0 \iff Q = 0 \iff R = 0$.

(b) Setting $S_X := \{x \in X \mid \|x\| = 1\}$, one has $\alpha = \beta$, where

$$
\alpha := \max_{x \in S_X} \varphi(x) = \max\{\lambda \in \mathbb{R} \mid \exists x \in X \setminus \{0\} : Qx = \lambda x\}, \quad (27)
$$

$$
\beta := \max_{y \in S_Y} \psi(y) = \max\{\lambda \in \mathbb{R} \mid \exists y \in Y \setminus \{0\} : Ry = \lambda y\}. \quad (28)
$$

(c) If $H \neq 0$, then $\text{Im} Q \neq \{0\}$, $\text{Im} R \neq \{0\}$, and $\gamma = \delta > 0$, where

$$
\gamma := \min_{x \in S_X \cap \text{Im} Q} \varphi(x) = \min\{\lambda > 0 \mid \exists x \in X \setminus \{0\} : Qx = \lambda x\}, \quad (29)
$$

$$
\delta := \min_{x \in S_Y \cap \text{Im} R} \psi(y) = \min\{\lambda > 0 \mid \exists y \in Y \setminus \{0\} : Ry = \lambda y\}. \quad (30)
$$

(d) The following implications hold:

$$
\min_{x \in S_X} \varphi(x) = 0 \iff \ker Q \neq \{0\} \iff \text{Im} Q \neq X \iff \text{Im} H \neq X,
$$

$$
\min_{y \in S_Y} \psi(y) = 0 \iff \ker R \neq \{0\} \iff \text{Im} R \neq Y \iff \ker H \neq \{0\}.
$$

Proof. Observe that any result obtained for $Q$ is valid for $R$ because $(H^*)^* = H$.

(a) It is obvious that $Q$ is self-adjoint; moreover, $\langle Qx, x \rangle = \langle HH^*x, x \rangle = \langle H^*x, H^*x \rangle = \varphi(x) \geq 0$ for every $x \in X$. The inclusions $\ker H^* \subset \ker Q$ and $\text{Im} Q \subset \text{Im} H$ are obvious.

Take $x \in \ker Q$, that is $Qx = 0$; then $0 = \langle x, Qx \rangle = \|H^*x\|^2$, and so $x \in \ker H^*$.

Because $Q$ is self-adjoint, we have that $\text{Im} Q = \{\ker Q\}^\perp$, and so $X = \ker Q + \text{Im} Q$. Let $x \in \text{Im} H$, that is $x = Hy$ for some $y \in Y$; then $x = Qu + z$ for some $u \in X$ and $z \in \ker Q = \ker H^*$, and so $\|z\|^2 = \langle z, Hy - H^*u \rangle = \langle H^*z, y - H^*u \rangle = 0$. It follows that $x = Qu \in \text{Im} Q$.

Because $\ker H = X \iff H = 0 \iff H^* = 0 \iff \ker H^* = Y$, the mentioned equivalences follow from the first part.

(b) The conclusion is obvious if $H = 0$ (in which case $Q = 0$ and $\varphi = 0$). So, let $H \neq 0$, and so $Q \neq 0$, whence $\alpha > 0$. Even if the equalities in (27) and (28) are well known, they will be recovered below. In fact, the inequalities $\geq$ are almost obvious. Because $\varphi$ is continuous and $S_X$ is compact, there exists $\pi \in S_X$ such that $\alpha = \varphi(\pi)$, and so

$$
\alpha = \|H^*\pi\|^2 = \langle \pi, Q\pi \rangle \geq \varphi(x) = \langle x, Qx \rangle \quad \forall x \in S_X,
$$

whence $\alpha \|x\|^2 \geq \langle x, Qx \rangle$, or equivalently $\langle (\alpha I - Q)x, x \rangle \geq 0$, for $x \in X$. Using Schwarz inequality for positive semi-definite operators and the fact that $\langle (\alpha I - Q)\pi, \pi \rangle = 0$, we get

$$
\langle (\alpha I - Q)x, x \rangle \leq \sqrt{\langle (\alpha I - Q)\pi, \pi \rangle} \sqrt{\langle (\alpha I - Q)x, x \rangle} = 0 \quad \forall x \in X;
$$
hence \((\alpha I - Q)x = 0\), that is \(Qx = \alpha x\). Hence the inequality \(\leq\) holds in \([27]\). Since \(Q = HH^*\), setting \(\overline{y} := \alpha^{-1/2}H^*x \in Y\), we have that \(\|\overline{y}\| = \alpha^{-1/2} \|H^*x\| = 1\), and so \(\overline{y} \in SY\). It follows that
\[
\beta \geq \psi(\overline{y}) = \|H\overline{y}\|^2 = \alpha^{-1} \|HH^*x\|^2 = \alpha^{-1} \|Qx\|^2 = \alpha^{-1} \|Qx\|^2 = \alpha.
\]
Applying the argument above for \(Q\) replaced by \(R\), we obtain that \(\alpha \geq \beta\), and so \(\alpha = \beta\).

(c) First observe that \(S_X \cap \text{Im} Q\) is a nonempty compact set, and so there exists \(x \in S_X \cap \text{Im} Q\) such that \(\gamma = \varphi(x)\), and so
\[
\gamma = \|H^*x\|^2 = \langle x, Qx \rangle \leq \varphi(x) = \langle x, Qx \rangle \quad \forall x \in S_X \cap \text{Im} Q.
\]
Assuming that \((\overline{x}, Q\overline{x}) = 0\), as above, we obtain that \(Q\overline{x} = 0\), that is \(\overline{x} \in \ker Q\). Since \(\ker Q \cap \text{Im} Q = \{0\}\), we get the contradiction \(0 \in S_X\). Therefore, \(\gamma > 0\). From \((31)\) we obtain that \(\gamma \|x\|^2 \leq \langle x, Qx \rangle\), or equivalently \(\langle (Q - \gamma I)x, x \rangle \geq 0\), for \(x \in \text{Im} Q\). Using Schwarz inequality for the positive semi-definite operator \(\Phi := (Q - \gamma I)|\text{Im} Q : \text{Im} Q \to \text{Im} Q\) and the fact that \(\langle \Phi x, x \rangle = 0\), we get \(\|Q - \gamma I\| \leq \sqrt{\langle \Phi x, x \rangle} = 0\) for all \(x \in \text{Im} Q\), whence \(\Phi = 0\), that is \(Q\overline{x} = \gamma \overline{x}\). As in the proof of (b) we take \(\overline{y} := \gamma^{-1/2}H^*\overline{x} \in \text{Im} H^*\); it follows that \(\overline{y} \in SY \cap \text{Im} H^* = SY \cap \text{Im} R\), and so \(\delta \leq \psi(\overline{y}) = \gamma\). The converse inequality follows similarly.

(d) These equivalences are immediate consequences of the equalities in (a) and the arguments at the beginning of the proof of (c).

\[\square\]

5 The case \(\overline{σ} \in S^−\)

Throughout this section we assume that \(V \in \Gamma^2_sc\), where \(\Gamma^2_sc := \Gamma^2_sc(\mathbb{R}^m)\) is the class of those \(g \in \Gamma sc\) which are twice differentiable on \(\text{int}(\text{dom} g)\) with \(\nabla^2 g(y) > 0\) for \(y \in \text{int}(\text{dom} g)\).\footnote{Note that the function \(V\) considered in \([30]\) belongs to \(\Gamma^2_sc(\mathbb{R})\).}

Observe that for \(g \in \Gamma^2_sc\) one has \(g^* \in \Gamma^2_sc\) and
\[
\nabla^2 g^*(\sigma) = (\nabla^2 g((\nabla g)^{-1}(\sigma)))^{-1} \quad \forall \sigma \in \text{int}(\text{dom} g^*).
\]
In the sequel \(V \in \Gamma^2_sc\). It follows that
\[
\langle u, \nabla^2 f(x)u \rangle = \left\langle u, \left[ A_0 + \sum_{i=1}^m \frac{\partial V}{\partial y_i}(q(x)) \cdot A_i \right] u \right\rangle + \langle v_u, \nabla^2 V(q(x))v_u \rangle
\]
for all \(x \in X_0\) and \(u \in \mathbb{R}^n\), where \(v_u := (\langle u, A_1 x - b_1 \rangle, \ldots, \langle u, A_m x - b_m \rangle)^T\), and
\[
\frac{\partial^2 D}{\partial \sigma_i \partial \sigma_k}(\sigma) = -\langle A_i A(\sigma)^{-1} b(\sigma) - b_i, A^{-1} (A_k A(\sigma)^{-1} b(\sigma) - b_k) \rangle - \frac{\partial^2 V^*}{\partial \sigma_i \partial \sigma_k}(\sigma)
\]
\[
= -\langle A_i x(\sigma) - b_i, A^{-1} (A_k x(\sigma) - b_k) \rangle - \frac{\partial^2 V^*}{\partial \sigma_i \partial \sigma_k}(\sigma)
\]
for all \(\sigma \in \text{int} S_0\) and \(i, k \in 1, m\). It follows that
\[
\langle v, \nabla^2 D(\sigma)v \rangle = -\langle A_v x(\sigma) - b_v, A(\sigma)^{-1} (A_v x(\sigma) - b_v) \rangle - \langle v, \nabla^2 V^*(\sigma)v \rangle
\]
for all \(v \in \mathbb{R}^m\) and \(\sigma \in S_0\), where
\[
A_v := \sum_{i=1}^m v_i A_i, \quad b_v := \sum_{j=1}^m v_j b_j \quad (v \in \mathbb{R}^m).
\]
The expression above shows that $D$ is strictly concave on $S^+$, confirming the remark done after getting the formulas for $D$ in (11).

Assume that $(\overline{x}, \sigma) \in X_0 \times S^-$ is a critical point of $\Xi$; by (20) we have that $\sigma = \nabla V (q(\overline{x}))$. Because $A(\overline{x}) < 0$ and $\nabla^2 V (q(\overline{x})) > 0$ there exist non-singular matrices $E \in M_n$ and $F \in M_m$ such that $-A(\overline{x}) = E^* E$ and $\nabla^2 V (q(\overline{x})) = F^* F$, where $E^*$ and $F^*$ are the transposed matrices of $E$ and $F$, respectively; hence $A(\overline{x})^{-1} = -E^{-1}(E^{-1})^*$ and $\nabla^2 V^* (\overline{x}) = F^{-1}(F^{-1})^*$.

Let us set

$$d_i := (E^{-1})^* (A \overline{x} - b_i) \in \mathbb{R}^n \quad (i \in \overline{1, m}).$$

Because for $A_i$ defined in (33) one has

$$A_i \overline{x} - b_i = \sum_{i=1}^{m} v_i (A_i \overline{x} - b_i) = \sum_{i=1}^{m} v_i E^* d_i = E^* \sum_{i=1}^{m} v_i d_i,$$

from (34) we obtain that

$$\langle v, \nabla^2 D(\overline{x}) v \rangle = \left\langle E^* \sum_{i=1}^{m} v_i d_i, E^{-1}(E^{-1})^* E^* \sum_{i=1}^{m} v_i d_i \right\rangle - \langle v, F^{-1}(F^{-1})^* v \rangle$$

$$= \left\| \sum_{i=1}^{m} v_i d_i \right\|^2 - \left\| (F^{-1})^* v \right\|^2 \quad \forall v \in \mathbb{R}^m.$$  

Taking into account (33), we have that

$$\langle u, \nabla^2 f(\overline{x}) u \rangle = \langle u, A(\overline{x}) u \rangle + \langle v, \nabla^2 V (q(\overline{x})) v \rangle = \| F v \|^2 - \| E u \|^2 \quad \forall u \in \mathbb{R}^n,$$

where

$$v_u = (\langle u, A_i \overline{x} - b_i \rangle)_{i \in \overline{1, m}} = (\langle E u, (E^{-1})^* (A_i \overline{x} - b_i) \rangle)_{i \in \overline{1, m}} = ((E u, d_i))_{i \in \overline{1, m}}.$$

Let us set

$$J : \mathbb{R}^m \to \mathbb{R}^n, \quad Jv := \sum_{i=1}^{m} v_i d_i \quad (v \in \mathbb{R}^m);$$

then

$$J^* : \mathbb{R}^n \to \mathbb{R}^m, \quad J^* u = (\langle u, d_1 \rangle, \ldots, \langle u, d_m \rangle)^T =: (\langle u, d_i \rangle)_{i \in \overline{1, m}} \quad (u \in \mathbb{R}^n).$$

Take $H : \mathbb{R}^m \to \mathbb{R}^n$ defined by $H := J \circ F^*$. Then $H^* = F \circ J^* : \mathbb{R}^n \to \mathbb{R}^m$. Because denoting $u' := Eu$ for $u \in \mathbb{R}^n$ and $v' := (F^{-1})^* v$ for $v \in \mathbb{R}^m$, from (38) and (37) we obtain that

$$\langle u, \nabla^2 f(\overline{x}) u \rangle = \| H^* u' \|^2 - \| u' \|^2, \quad \langle v, \nabla^2 D(\overline{x}) v \rangle = \| H v' \|^2 - \| v' \|^2.$$

Because $E$ and $F$ are non-singular, for $\rho \in \{>, \ge, <, \le\}$ and $\rho' \in \{>, \ge, <, \le\}$ with the natural correspondence, we have

$$\nabla^2 f(\overline{x}) \rho' 0 \iff \left[ \| H^* u' \|^2 \rho 1 \forall u' \in S_n \right] \iff [\varphi(u) \rho 1 \forall u \in S_n],$$

$$\nabla^2 D(\overline{x}) \rho' 0 \iff \left[ \| H v' \|^2 \rho 1 \forall v' \in S_n \right] \iff [\psi(v) \rho 1 \forall v \in S_n],$$

where $S_m := \{y \in \mathbb{R}^m \mid \| y \| = 1 \} = S_{\mathbb{R}^m}$, and $\varphi$, $\psi$ are defined in Proposition 8 with

$$H := J \circ F^* : \mathbb{R}^m \to \mathbb{R}^n, \quad H^* = F \circ J^* : \mathbb{R}^n \to \mathbb{R}^m.$$  

Recall that $E \in M_n$ and $F \in M_m$ are such that $-A(\overline{x}) = E^* E$ and $\nabla^2 V (q(\overline{x})) = F^* F$, $(d_i)_{i \in \overline{1, m}}$ are defined in (36), $J$ is defined in (40), and $H$ is defined in (44).

In the next result we shall use Proposition 8 for the operator $H$ defined in (44); therefore, $X = \mathbb{R}^n$ and $Y = \mathbb{R}^m$. Using Proposition 8 if necessary, and setting $\dim \{0\} := 0$, the following assertions hold:
• \( \dim(\text{Im} \, H) = \dim(\text{Im} \, H^*) \leq \min\{n, m\} \),

• \( \dim(\ker H) + \dim(\text{Im} \, H) = m, \ \dim(\ker H^*) + \dim(\text{Im} \, H^*) = n \),

• \( \dim(\ker H^*) \) is equal to the multiplicity of the eigenvalue 0 of \( Q := H \circ H^* \), while \( \dim(\ker H) \) is equal to the multiplicity of the eigenvalue 0 of \( R := H^* \circ H \).

From the above considerations we obtain the following result.

**Proposition 9** Let \((\tau, \sigma) \in X_0 \times S^-\) be a critical point of \( \Xi \). Consider \( E \in \mathcal{M}_n \) such that \( E^*E = -A(\tau), \ d_i \in \mathbb{R}^n \ (i \in \overline{1, m}) \) defined in \((\mathbb{Z})\), and \( H \) defined in \((\mathbb{A})\).

(i) If \( \tau \) (resp. \( \sigma \)) is a local maximizer of \( f \) (resp. \( D \)), then \( \|Hv\| \leq 1 \) for all \( v \in S_m \), or, equivalently, \( (\alpha =) \beta \leq 1 \). Conversely, if \( \|Hv\| < 1 \) for all \( v \in S_m \), then \( \tau \) (resp. \( \sigma \)) is a local strict maximizer of \( f \) (resp. \( D \)). In particular, if \( A_i \tau = b_i \) (or equivalently \( d_i = 0 \)) for all \( i \in \overline{1, m} \), then \( \tau \) and \( \sigma \) are local strict maximizers of \( f \) and \( D \), respectively.

(ii) If \( \tau \) is a local minimizer of \( f \), then \( \|H^*u\| \geq 1 \) for all \( u \in S_n \); in particular \( H \) is surjective, \( m \geq n \), and every positive eigenvalue of \( H^* \circ H \) is greater than or equal to 1. Conversely, if \( \|H^*u\| > 1 \) for all \( u \in S_n \), then \( \tau \) is a local strict minimizer of \( f \); moreover, if \( m > n \) then \( \tau \) is not a local extremum for \( D \).

(iii) If \( \sigma \) is a local minimizer of \( D \), then \( \|Hv\| \geq 1 \) for all \( v \in S_m \); in particular \( H \) is injective, \( m \leq n \), and every eigenvalue of \( H \circ H^* \) is greater than or equal to 1. Moreover, if \( m < n \) then \( \sigma \) is not a local extremum for \( f \). Conversely, if \( \|Hv\| > 1 \) for all \( v \in S_n \), then \( \sigma \) is a local strict minimizer of \( D \).

(iv) Assume that \( m = n \) and \( \{A_i \tau - b_i \mid i \in \overline{1, m}\} \) is a basis of \( \mathbb{R}^m \). If \( \|Hv\| > 1 \) for all \( v \in S_n \), then \( \tau \) and \( \sigma \) are local strict minimizers of \( f \) and \( D \), respectively.

Proof. Taking into account the well known second order necessary or sufficient conditions for local extrema of unconstrained problems, the assertions are immediate consequences of \((\mathbb{H}2)\), \((\mathbb{H}3)\) and Proposition \((\mathbb{S})\).

Note that Proposition \((\mathbb{S})\) (iii) gives a positive answer to the question formulated on the sixth line from below of \((\mathbb{Z})\) p. 234 because in that case \( \tau \) is a strict local minimum of \( P^d \) (= \( D \)) and \( \tau \) is not a local extremum of \( P \) (= \( f \)) since \( m = 1 < 2 \leq n \). In \((\mathbb{Z})\) one uses the following assumption: “(A3) The critical points of problem \((P)\) are non-singular, i.e., if \( \nabla^2 \Pi(\tau) = 0 \), then \( \det \nabla^2 \Pi(\tau) \neq 0 \).” Under such a condition we have the following result.

**Corollary 10** Let \((\tau, \sigma) \in X_0 \times S^-\) be a critical point of \( \Xi \) such that \( \det \nabla^2 f(\tau) \neq 0 \) [that is 0 is not an eigenvalue of \( \nabla^2 f(\tau) \)]. The following assertions hold:

(a) \( \tau \) is a local maximizer of \( f \) if and only if \( \|Hv\| < 1 \) for all \( v \in S_m \), if and only if \( \sigma \) is a local maximizer of \( D \).

(b) Assume that \( m = n \). Then \( \tau \) is a local minimizer of \( f \) if and only if \( \|Hv\| > 1 \) for all \( v \in S_m \), if and only if \( \sigma \) is a local minimizer of \( D \).

Proof. First observe that for \( A \in \mathcal{G}_n \) one has \( A > 0 \) if and only if \( [A \geq 0 \text{ and } \det A \neq 0] \). Recall that \( \alpha := \max_{u \in S_m} \|H^*u\|^2 = \max_{u \in S_n} \|Hv\|^2 =: \beta \).

(a) Assume that \( \tau \) is a local maximizer of \( f \). Then \( A := \nabla^2 f(\tau) \leq 0 \) and so, \( A < 0 \). By \((\mathbb{H}2)\) we have that \( 1 > \alpha = \beta \) (that is \( \|Hv\| < 1 \) for all \( v \in S_m \)), which at its turn implies that \( \sigma \) is a local maximizer of \( D \) by Proposition \((\mathbb{S})\) (i).
Assume that \( \sigma \) is a local maximizer of \( D \). Then \( \alpha \leq 1 \) by Proposition 9 (i), and so \( A \preceq 0 \) by (42), whence \( A < 0 \). Using again Proposition 9 (i), we have that \( \sigma \) is a local maximizer of \( f \).

The proof of (b) is similar to that of (a). \( \square \)

6 Relations with previous results

In this section we analyze results obtained by DY Gao and his collaborators in papers dedicated to unconstrained optimization problems, related to “triality theorems”. The main tool to identify the papers where this class of problems are considered was to look in the survey papers [4] (which practically includes [5]), [7] (which is almost the same as [6]), [15] (which is very similar to [8]), [13] (which is the same as [12]), as well as in the recent book [10].

Though, in order to understand the chronology of the development of this topic let us quote first the following texts from [13, p. 40] (see also [12, p. NP30]) and [20, p. 136] (see also [19, p. 5]), respectively:

Q1 – “the triality was proposed originally from post-buckling analysis [42] in “either-or” format since the double-max duality is always true but the double-min duality was proved only in one-dimensional nonconvex analysis [49].”

Q2 – “the triality theorem was formed by these three pairs of dualities and has been used extensively in nonconvex mechanics [10, 17] and global optimization [3, 21, 34]. However, it was realized in 2003 [12, 13] that if the dimensions of the primal problem and its canonical dual are different, the double-min duality (30) needs “certain additional conditions”. For the sake of mathematical rigor, the double-min duality was not included in the triality theory and these additional constraints were left as an open problem (see Remark 1 in [12], also Theorem 3 and its Remark in a review article by Gao [13]). By the facts that the double-max duality (29) is always true and the double-min duality plays a key role in real-life applications, it was still included in the triality theory in the either-or form in many applications for the purposes of perfection in esthesis and some other reasons in reality.”

Having in view Q1 and Q2, it seems that the main steps in the development of the “triality theory” are marked by [2] (where the triality theorem was proved for the one-dimensional case), [5] (where it is mentioned that “certain additional conditions” are needed for the

---

2The reference “[42]” is “Gao, D.Y.: Dual extremum principles in finite deformation theory with applications to post-buckling analysis of extended nonlinear beam theory. Appl. Mech. Rev. 50(11), S64–S71 (1997)”, while “[49]” is DY Gao’s book [2] from our bibliography. Unfortunately, it is not given the precise place (e.g. the page) where the “double-min duality” was proved for \( n = m = 1 \).

Among other results, in “[42]” there is “Theorem 7 (Triality Theorem)”, stated without proof; immediately after it is said: “The proofs of these theorems are given elsewhere”.

3The references “[8]”, “[17]” and “[34]” are: “Fang, S.C., Gao, D.Y., Sheu, R.L., Wu, S.Y.: Canonical dual approach for solving 0–1 quadratic programming problems. J. Ind. Manag. Optim. 4, 125–142 (2008)”, “Gao, D.Y., Ogden, R.W.: Multiple solutions to non-convex variational problems with implications for phase transitions and numerical computation. Q. J. Mech. Appl. Math 61(4), 497–522 (2008)” and “Ruan, N., Gao, D.Y., Jiao, Y.: Canonical dual least square method for solving general nonlinear systems of quadratic equations. Comput. Optim. Appl. 47, 335–347 (2010)”, respectively; the references “[10]”, “[12]”, “[13]”, “[21]” are our items [2], [3], [4], and [15], respectively.

4On the web-page http://www.isogop.org/organization/david-y-gao/super-duality-triality (accessible at least until September 1st, 2014), DY Gao confessed: “Actually, I even forgot my this problem left in 2003 [1,2] due to busy life during those years”. So, which is the truth about continuing to formulate the “triality theorems” in the “either-or form” in the papers published in the period 2003-2011?
“double-min duality” to be valid), [20] and its preprint version [18] (where “this double-min duality has been proved for ... general global optimization problems”, as mentioned in [13, p. 40]).

Let us compare first our results with those from the most recently published paper on this topic for general $V$, that is [20].

Putting together Assumptions (A1) and (A2) of Gao and Wu’s paper [20] (see also [18]), the function $V$ considered there is real-valued, strictly convex, and twice continuously differentiable on $\text{Im} q$ (see also [20, p. 134]). Hence $V$ from [20] is more general than being in $\Gamma^2_s$, when $\text{dom} V = \mathbb{R}^m$. Of course, the strict convexity of $V$ implies $\nabla^2 V(y) \succeq 0$ for $y \in \mathbb{R}^m$, but this property does not imply $(\nabla^2 V)(q(\overline{\sigma})) > 0$, which is used for example in [20, Eq. (36)].

In “Theorem 2 (Tri-duality Theorem)” (the case $n = m$) and “Theorem 3. (Triality Theorem)” (the case $n \neq m$), $\overline{\sigma} \in S_{\text{col}}$ is a “critical point of the canonical problem ($\mathcal{P}^d$)” and $\overline{\sigma} := [A(\overline{\sigma})]^{-1} b(\overline{\sigma})$ and Assumption (A3) holds, that is $[\nabla f(x) = 0 \Rightarrow \det \nabla^2 f(x) \neq 0]$. Our result in the case $A(\overline{\sigma}) \succeq 0$ is more general than those in [20] Ths. 2, 3 not only because the hypothesis on $V$ in Proposition 2 is weaker and Assumption (A3) is not present, but also because the conclusion in [20] Ths. 2, 3 is weaker, more precisely $f(\overline{\sigma}) = \inf_{x \in \mathbb{R}^m} f(x) \Leftrightarrow \sup_{\sigma \in S^+_{\text{col}}} D(\sigma) = D(\overline{\sigma})$. In what concerns the case $A(\overline{\sigma}) < 0$ and $V \in \Gamma^2_{sc}$, Corollary 10 is much more precise than the corresponding results in [20] Ths. 2, 3 because it is mentioned when $\overline{\sigma}$ and $\overline{\sigma}$ are local minimizers (maximizers). Moreover, our proofs are very different from those of [20], and follow the lines of the proof of [30, Prop. 1].

Similar results to those in [20] Ths. 2, 3 for particular $V$ can be found in several papers co-authored by DY Gao after he became acquainted with the content of our paper [30] (see also [16], [17], [21], [24], [13], [27], [4], [21], [22]).

Gao and Wu in [16], [17] and [19] (which are essentially the same) prove [20] Ths. 2, 3 for $V(y) := \frac{1}{2} \sum_{k=1}^m \beta_k y_k^2$ with $\beta_k > 0$ and $b_k := 0$ ($k \in \overline{1, m}$) [under Assumption (A3)], using similar arguments. Note that $\overline{\sigma}$ is taken to be a critical point of $D$ in “Theorem 4.3 (Refined Triality Theorem)” (the case $n \neq m$) instead of being a “critical point of Problem ($\mathcal{P}^d$)”, as in “Theorem 3.1 (Tri-Duality Theorem)”.

Morales-Silva and Gao in [23] discuss the problem from [16] with $A_0 := 0$ and $m := 1$. Morales-Silva and Gao in [24] (and [23]) consider $V(y) := \sum_{k=1}^p \exp(y_k) + \frac{1}{2} \sum_{k=p+1}^m \beta_k y_k^2$ for $0 \leq p \leq m$ (setting $\sum_{k=1}^j \gamma_k := 0$ when $j < i$) with $\beta_k > 0$ for $k \in \overline{p+1, m}$; moreover, $b_k := 0$ and $A_k \geq 0$ for $k \in \overline{1, m}$ are such that there exists $(\alpha_k)_{k \in \overline{1, m}} \subset \mathbb{R}^+_m$ with $\sum_{k=1}^m \alpha_k A_k > 0$. Under Assumption (A3) and using similar arguments to those in [19] Th. 2, they prove [20] Ths. 2, 3 for $\overline{\sigma}$ “a stationary point of $D$.

Chen and Gao in [1] consider $V(y) := \frac{1}{2} \log (1 + \sum_{k=1}^p \exp(\beta y_k)) + \frac{1}{2} \sum_{k=p+1}^m \beta_k y_k^2$ with $\beta, \beta_k > 0$ ($k \in \overline{p+1, m}$) and $b_k := 0$ ($k \in \overline{1, m}$). By an elementary computation (and using

\[ \text{Note that dom } V = (0, \infty)^m \text{ in [20, Eq. (61)].} \]

\[ \text{It is not defined what is meant by critical point of the problem } \neg ((\mathcal{P}^d) : \text{ext}(\Pi^d(\varsigma)) = -\frac{1}{2} \langle (G(\varsigma))^{-1} F(\varsigma), F(\varsigma) \rangle - V^*(\varsigma) | \varsigma \in S_a \rangle, \text{ where } S_a \text{ is our } S_{\text{col}} \text{ and } G^{-1} \text{ should be understood as a generalized inverse if } \det G = 0 [11]^\#, \text{ “[11]” being item } 3 \text{ from our bibliography. Moreover, the formula for } \nabla^2 \Pi^d(\varsigma) \text{ in Eq. (34) is not justified, having in view that } \varsigma \in S_a (= S_{\text{col}}). \]

\[ \text{The paper [20] was submitted to MMOR on 11/09/2009 (manuscript MMOR-D-09-00165), rejected on DY Gao’s report on 15/04/2011, and re-submitted, without any modification, on 27/04/2011 (manuscript MMOR-D-11-00075); see the submission date of [19] to arxiv.} \]

14
Cauchy's inequality) one obtains \( \nabla^2 h(z) \succ 0 \) for \( z \in \mathbb{R}^p \), where \( h(z) := \ln \left( 1 + \sum_{k=1}^{p} \exp(z_k) \right) \) for \( z \in \mathbb{R}^p \); therefore, \( \nabla x \in \Gamma^2_{sc} \). In “Theorem 4 (Triality Theorem)”, for \( \sigma \in S_0 \) a critical point of \( D \), one obtains the “min-max duality” for \( \sigma \in S^+ \), while for \( \sigma \in S^- \) one obtains the “double-max duality” and “double-min duality” (this one for \( m = n \)) without using Assumption (A3). However, the proof in the case \( \sigma \in S^- (= S^-_a \mbox{ with the notation in [1]} \) is not convincing.

Let us quote Note 1 from [1] p. 421: “We use the same definition of the neighborhood as defined in [15] (Note 1 on page 306), i.e., a subset \( X_0 \) is said to be the neighborhood of the critical point \( \bar{x} \) if \( \bar{x} \) is the only critical point in \( X_0 \).”

Jin and Gao in [21] and [22] (which are essentially the same) consider practically the same V as in [25], that is \( V(y) := \sum_{k=1}^{p} \exp(y_k) + \frac{1}{2} \sum_{k=p+1}^{m} y_k^2 \) with \( 0 \leq p \leq m \); moreover, \( b_k := 0 \) for \( k \in 1, m \), \( \alpha_0 > 0 \) and \( A_k \succ 0 \) for \( k \in \overline{p+1,m} \). Note that for this \( V \), the statements of “Theorem 2. (Triality theorem)” in [21] and [22] and their proofs are almost the same as those in [1] Th. 4. The differences are: a) in the case of “min-max duality”, Assumption 2 in [21] (resp. Assumption 1 in [22]) implies \( \sigma \in S^+ \), b) the case \( n \neq m \) for the “double-min duality” is missing in [21] and [22], and c) “for some neighborhood \( \lambda_0 \times S_0 \subset \mathbb{R}^n \times S^-_a \) of \((\sigma,\bar{\sigma})\)” from [1] Th. 4 is replaced by \((\sigma \in \lambda_0 \subset \mathbb{R}^n \mbox{ and } \bar{\sigma} \in S_0 \subset S^-_a \)” in [21] Th. 2 and [22] Th. 2 [1]. Of course, the drawbacks in the proof of [1] Th. 4 mentioned in Note 8 remain valid for the proofs of [21] Th. 2] and [22] Th. 2 [1].

A special place among DY Gao’s papers published after 2010 is occupied by [27].

Gao, Ruan and Pardalos in [14] take the same V as in [19] but Assumption (A3) is not considered. Putting together Theorems 2 and 3 from [19] for “\( \sigma \) a critical point of the canonical dual function \( P(\sigma) \)”, with the mention “If \( n \neq m \), the double-min duality (25) holds conditionally”, one gets “Theorem 2 (Triality Theorem)” of [14]. A detailed proof is provided in the case \( \sigma \in S^+_a (= S^+_{col}) \). The proof for the case \( \sigma \in S^-_a (= S^-) \) is the following:

“If \( \sigma \in S^-_a \), the matrix \( G(\sigma) \) is a negative definite. In this case, the Gao–Strang complementary function \( \Xi(\sigma,\xi) \) is a so-called super-Lagrangian [14], i.e., it is locally concave in both \( x \in X_0 \subset X_\sigma \) and \( \sigma \in S_0 \subset S^-_a \). By the fact that

\[
\max_{x \in X_0} \max_{\sigma \in S_0} \Xi(x,\sigma) = \max_{\sigma \in S_0} \max_{x \in X_0} \Xi(x,\sigma)
\]

(26) holds on the neighborhood \( X_0 \times S_0 \) of \((\sigma,\xi)\), we have the double-max duality statement (24). If \( n = m \), we have [33]:

\[
\min_{x \in X_0} \max_{\sigma \in S_0} \Xi(x,\sigma) = \min_{\sigma \in S_0} \max_{x \in X_0} \Xi(x,\sigma)
\]

(27) which leads to the double-min duality statement (25). This proves the theorem. [12]

[8] In the proof of [1] Th. 4.2] one says: “Suppose \( \sigma \) is a local maximizer of \( \Pi^d(\xi) \) in \( S^-_a \). Then we have \( \nabla^2 \Pi^d(\sigma) = -F^T G_x^a F - D^{-1} \preceq 0 \) and there exists a neighborhood \( S_0 \subset S^-_a \) such that for all \( \xi \in \sigma \), \( \nabla^2 \Pi^d(\xi) \preceq 0 \). Since the map \( x = G_x^a f \) is continuous over \( S_0 \), the image of the map over \( S_0 \) is a neighborhood of \( \sigma \), which we denote as \( X_0 \). Next we are going to prove that for any \( x \in X_0 \), \( \nabla^2 \Pi^d(x) \preceq 0 \), which plus the fact that \( \sigma \) is a critical point of \( \Pi(x) \) implies \( \sigma \) is a maximizer of \( \Pi(x) \) over \( X_0 \).” A similar argument is used for proving [1] Th. 4.3], too.

The drawbacks in the quoted text are the following: (a) the fact that \( \nabla^2 \Pi^d(\sigma) \preceq 0 \) implies that \( \nabla^2 \Pi^d(\xi) \preceq 0 \) for \( \xi \in \sigma \) in a neighborhood of \( \sigma \) is not motivated; (b) generally, a continuous function is not open.

[9] The reference “[15]” is item [14] from our bibliography. With this definition, in Example 1 of [1] p. 426], \( \sigma_1 \) is a minimizer of \( \Pi \) on \( X_0 := (-1.5,1.5) \setminus \{(\sigma_1,\sigma_2)\} \); of course, this is false.

[10] Why not take \( X_0 := \mathbb{R}^n \) and \( S_0 := S^-_a \)?

[11] Setting \( G := G(\sigma) \), [21] Eq. (38) and [22] Eq. (25) assert that \( \eta := \inf_{x \in X_0} \left( \frac{1}{2} \langle x, G x \rangle + \langle x, f \rangle \right) = -\frac{1}{2} \langle f, G^{-1} f \rangle \) if \( G \succ 0 \) and \( \eta = -\infty \) otherwise. In fact \( \eta \in \mathbb{R} \) if and only if \( \|G \| \vee \langle f, G \rangle \) if \( G \succ 0 \) and \( f = G x_0 \) then \( \eta = -\frac{1}{2} \langle x_0, G x_0 \rangle \) (see e.g., [31], Prop. 2.1 (i)).

[12] The reference “[14]” is Gao’s book [2], while “[33]” is “Gao, D.Y. and Wu, C-Z. (2010). On the Triality
Ruan and Gao in [27] take the same $V$ as in [14] [$(a_k)$ and $(b_k)$ being different] and similarly, Assumption (A3) is not considered. The differences in [27, Th. 2] with respect to [14, Th. 2] are: (a) $S^+_a$ is $S^+_d$ instead of $S^+_{col}$, (b) “on the neighborhood” is replaced by “on its neighborhood”, (c) $m = n$ is replaced by $\dim X_a = \dim S_a$, and (d) $n \neq m$ in the case $\tau \in S_a^-$ is missing. In [27, Rem. 1] one mentions: “The double-max duality statement (24) can be proved easily by the fact that $\max_{x \in X_0} \max_{\varsigma \in S_\alpha} \Xi(x, \varsigma) = \max_{\varsigma \in S_\alpha} \max_{x \in X_0} \Xi(x, \varsigma)$ $\forall (x, \varsigma) \in X_0 \times S_0 \subset X_a \times S_a^-$.

The reference [32] is item [5] from our bibliography. Thus, the equality (23) leads to the statement (21), while (24) leads to the statement (22). This proves the theorem.”

There are very few differences between the proof of [14, Th. 2] in the case $\tau \in S_a^-$ and that of “Theorem 2 (Triality Theorem)” from [11], where $m = 1$, for the same case:

“If $\tau \in S_a^-$, the matrix $A_d(\tau)$ is negative definite. In this case, the Gao-Strang complementary function $\Xi(\overline{x}, \tau)$ is a so-called super-Lagrangian (see Gao (2000a)), i.e., it is locally concave in both $x \in X_0 \subset \mathbb{R}^n$ and $\varsigma \in S_a^0 \subset S_a$. Thus, by the triality theory developed in Gao (2000a), we have that either

$$\min_{x \in X_0} \max_{\varsigma \in S_a} \Xi(x, \varsigma) = \min_{\varsigma \in S_0} \max_{x \in \mathbb{R}^n} \Xi(x, \varsigma)$$

(23) or

$$\max_{x \in X_0} \min_{\varsigma \in S_a} \Xi(x, \varsigma) = \max_{\varsigma \in S_0} \max_{x \in \mathbb{R}^n} \Xi(x, \varsigma)$$

(24) holds on the neighborhood $X_0 \times S_0$ of $(\overline{x}, \tau)$. Thus, the equality (23) leads to the statement (21), while (24) leads to the statement (22). This proves the theorem.”

It is worth comparing the two proofs above with that (for the same case) of “THEOREM 2 (Triality Theorem)” from [27] (and of “Theorem 3 (Global Minimizer and Maximizer)” from [31]), where $m = 1$:

“If $\Sigma \in \mathbb{Y}_a^*$, then $(\overline{x}, \Sigma)$ is a so-called super-critical point of the extended Lagrangian $\Xi(x, y^*)$, i.e. $\Xi(\overline{x}, \Sigma)$ is locally concave in each of its variables $x$ and $y^*$ on the neighborhood $X_a \times \mathbb{Y}_a^*$.

In this case, we have

$$P(\overline{x}) = \max_{x \in X_a} \max_{y^* \in \mathbb{Y}_a^*} \Xi(x, y^*) = \max_{y^* \in \mathbb{Y}_a^*} \max_{x \in X_a} \Xi(x, y^*) = P_d(y^*)$$

by the fact that the maxima of the super-Lagrangian $\Xi(x, y^*)$ can be taken in either order on the open set $X_a \times \mathbb{Y}_a^*$ (see [17]). This proves the rest part of the theorem and (38).”

The presentation above shows that the papers [14] and [27] make the transition from the proofs of “triality theorems” published before 2010 (in which one observed that for “$\tau \in S_a^-$”, $\Xi$ is a “so-called super-Lagrangian”, and so “the triality theory developed in Gao (2000a)” applies), and the proofs of the other “triality theorems” published after 2011 with detailed and complicated (but not completely convincing) proofs for twice differentiable strictly convex functions $V$.

Coming back to Q1, we did not succeed to identify the place in [2] where “the double-min duality was proved only in one-dimensional nonconvex analysis”. We may consider the following text from [20, p. 131] as a hint for the above assertion:

Q3 “Therefore, instead of the mono-duality in static systems, convex Hamiltonian systems are controlled by the so-called bi-duality theory.”
Bi-Duality Theorem [10]: If \((\mathbf{x}, \mathbf{y}^*)\) is a critical point of the Lagrangian \(L(x, y^*)\), then \(\mathbf{x}\) is a critical point of \(\Pi(x)\), \(\mathbf{y}^*\) is a critical point of \(\Pi^*(y^*)\) and \(\Pi(\mathbf{x}) = L(\mathbf{x}, \mathbf{y}^*) = \Pi^*(\mathbf{y}^*)\). Moreover, if \(n = m\), we have

\[
\Pi(\mathbf{x}) = \max_{x \in X} \Pi(x) \iff \max_{y^* \in Y^*} \Pi^*(y^*) = \Pi^*(\mathbf{y}^*) \quad (10)
\]

\[
\Pi(\mathbf{x}) = \min_{x \in X} \Pi(x) \iff \min_{y^* \in Y^*} \Pi^*(y^*) = \Pi^*(\mathbf{y}^*) \quad (11)
\]

This bi-duality is actually a special case of the triality theory in geometrically linear systems, which was originally presented in Chap. 2 [10] for one-dimensional dynamical systems with a simple proof.\(^{15}\)

Denoting assertions (10) and (11) above by (65) and (66), respectively, and putting “or” between them, one gets the statement of the “Bi-Duality Theorem” from [20, p. 148]. Notice that only the “Bi-Duality Theorem” from [20, p. 148] is present in the preprint version of [20], that is \(^{16}\)

The “Bi-Duality Theory” is presented in [2, Sect. 2.6.2]. Apparently the result above is related to “Theorem 2.6.5 (Double-Min and Double-Max Duality)” from [2, p. 86] and to “Theorem 4 (Bi-Duality Theory [38])” \(^{17}\) from [15]; in these two theorems there are not references to the dimensions. Example 4.5 from [28] provides a counterexample for both [2, Th. 2.6.5] and [15, Th. 4], as well as for the bi-duality theorem from [20, p. 131]; however, that example is not a counterexample for the bi-duality theorem from [20, p. 148].

Another hint should be [2, Sect. 3.5] which is called “Tri-Extremum Principles and Triality Theory”, with its subsections 3.5.2 and 3.5.3 which are called “Triality Theorems” and “Tri-Duality Theory”, respectively. At the beginning of [2, Sect. 3.5] it is said:

“In this section we present the so-called triality theory under the following assumption. Assumption 3.5.1 Let \(\{U, U^*\}; \langle \ast, \ast \rangle\) and \(\{E, T\}; \langle \ast, \ast \rangle\) be two inner product spaces. ...

\(\text{(A1)}\) \(\Lambda: I \times U \rightarrow E\) is a quadratic operator \(\Lambda(u) = \frac{1}{2}a(x)u'(x)^2 + b(x)u'(x) + c(x), a(x) > 0\) \(\forall x \in I\), where \(a, b, c \in C^1(I)\) are given real-valued functions;

\(\text{(A2)}\) \(F: U_a \subset U \rightarrow \mathbb{R}\) is a linear, Gâteaux differentiable functional and, on \(U_a \times U^*_a \subset U \times U^*\), \(u^* = DF(u) \iff u = DF^*(u^*) \iff \langle u, u^* \rangle = F(u) + F^*(u^*)\);

\(\text{(A3)}\) \(W: E_a \subset E \rightarrow \mathbb{R}\) is either convex or concave and on \(E_a \times T_a \subset E \times T\), the Legendre duality relations \(c = DW(\xi) \iff \xi = DW^c(\xi) \iff \langle \xi; c \rangle = W(\xi) + W^c(\xi)\) hold.”

From [2, (3.107)], [2, (3.108)] and [2, (3.113)] we learn that \(\Pi(u) = W(\Lambda(u)) - F(u)\) for \(u \in U_k\) with \(U_k = \{u \in U_a \mid \Lambda(u) \in E_a\}\), \(L(u, \xi) = \langle \Lambda(u); \xi \rangle - W^c(\xi) - F(u)\), and \(\Pi^d(\xi) = F^c(u^*(\sigma)) - W^c(\xi) - G^c(\xi)\), respectively, in which \(F^c(u^*(\sigma))\) is the Legendre conjugate of \(F(u)\), and \(G^c: T_a \rightarrow \mathbb{R}\) is a pure complementary gap functional.

The above text shows that, at least in [2, Sect. 3.5], \(U\) is a function space like \(H^1(I)\). Of course, \(F\) being a linear function on \(U_a \subset U\), \(U_a\) has to be a linear subspace endowed we the trace topology. A linear functional \(f\) defined on a topological vector space \(U\) is Gâteaux differentiable if and only if \(f\) is continuous, in which case \(Df(u) = f\) for every \(u \in U\); moreover, it is not possible to speak about “the Legendre conjugate of \(F\)”. So, (A2) has not a mathematical meaning. Moreover, in order to speak about \(DW(\xi)\) and \(DW^c(\xi)\) in (A3),

\(^{15}\)The reference “[10]” is Gao’s book [2].

\(^{16}\)Notice the following (easy to be verified) false assertion from [20, Acknowledgements]: “The paper was posted online on April 15, 2011 at https://arxiv.org/abs/1104.2970”; just compare the submission dates (and Acknowledgements) of [16] and [18].

\(^{17}\)The reference “[38]” is Gao’s book [2].
one needs $E_a$ and $T_a$ be at least algebraically open (convex) subsets of $E$ and $T$, respectively. It is clear that the concerned spaces are not one-dimensional.

Because “Theorem 3.5.2 (Triality Theorem)” from [2] does not refer to primal and dual functions as in the usual formulations of “triality theorems” we quote such a result from [7] (which is maybe the last one) attributed to (Gao, 2000a), that is our reference [2].

“Theorem 3 (Triality theory (Gao, 2000a)). Suppose that $x$ is a critical point of $P^d$ and $x = G^d(\tau)\tau(\tau)$. If $G(\tau) \geq 0$, then $x$ is a global minimizer of ($P$), $x$ is a global maximizer of ($P^d$), and $\min_{x \in X_o} P(x) = \Xi(x, S) = \max_{\varepsilon \in S^*} P^d(\varepsilon)$. If $G(\tau) < 0$, then on a neighborhood $X_o \times S_o \subset X_o \times S_o^*$ of $(\tau, \tau)$, we have either $\min_{x \in X_o} P(x) = \Xi(\tau, S) = \min_{\varepsilon \in S_o} P^d(\varepsilon)$, or $\max_{x \in X_o} P(x) = \Xi(\tau, \tau) = \min_{\varepsilon \in S_o} P^d(\varepsilon)$.”

We consider that there is a misprint in the last $\min_{\varepsilon \in S_o} P^d(\varepsilon)$ of [7, Th. 3]; it has to be replaced by $\max_{\varepsilon \in S_o} P^d(\varepsilon)$, as in [6, Th. 2] (and all the other Gao’s papers containing a “triality theorem”).

In [7, Th. 3], “$X_o \subset \mathbb{R}^n$ is a given feasible space”, and “without losing much generality”, $V : E_a \rightarrow \mathbb{R}$ “is convex and lower semicontinuous”. Moreover “$G^d$ is the Moore–Penrose generalized inverse of $G$”. Without looking to details, [7, Th. 3] is similar to “Theorem 3.5.3 (Tri-Duality Theorem)” from [2]; note that the Moore–Penrose generalized inverse is not considered in [2].

It is worth quoting the most recent version of the general “triality theorem”, that is [13, Th. 3] (the same as [12, Th. 3]):

“Theorem 3 (Triality theorem) Suppose $\xi^*$ is a stationary point of $\Pi^d(\xi^*)$ and $\tau = G(\xi^*)^{-1}\tau$. If $\xi^* \in S^*_c$, we have
$$\Pi(\tau) = \min_{x \in X_o} \Pi(x) \Leftrightarrow \max_{\varepsilon^* \in S^*_c} \Pi^d(\varepsilon^*) = \Pi^d(\tau);$$
If $\xi^* \in S^*_c$, then on a neighborhood $X_o \times S_o \subset X_c \times S_c^*$ of $(\tau, \tau^*)$, we have either
$$\Pi(\tau) = \max_{x \in X_o} \Pi(x) \Leftrightarrow \max_{\varepsilon^* \in S_o} \Pi^d(\varepsilon^*) = \Pi^d(\tau),$$
or (only if $\dim \tau = \dim \tau^*$)
$$\Pi(\tau) = \min_{x \in X_o} \Pi(x) \Leftrightarrow \min_{\varepsilon^* \in S_o} \Pi^d(\varepsilon^*) = \Pi^d(\tau).$$

Note 5 in [13, Th. 3] (quoted above) is “The neighborhood $X_o$ of $\tau$ means that on which, $\tau$ is the only stationary point” (see also our Note 9). Related to this theorem, in [13, p. 14, 15] (and [12, p. NP13]) it is said:

“The triality theory was first discovered by Gao 1996 in post-buckling analysis of a large deformed beam [42, 52]. The generalization to global optimization was made in 2000 [51]. It was realized in 2003 that the double-min duality (32) holds under certain additional condition [57, 58]. Recently, it is proved that this additional condition is simply $\dim \tau = \dim \tau^*$ to have the strong canonical double-min duality (32), otherwise, this double-min duality holds weakly in subspaces of $X_o \times S_o$ [79, 80, 112, 113].”

Coming back to [13, Th. 3], we have to know which are the conditions on the function corresponding to our $V$, that is $\Phi$. At the beginning of Section “2.4 Triality Theory” of...

\footnote{Compare this text with Q2. The references “[52]” and “[112]” are “Gao, D.Y.: Finite deformation beam models and triality theory in dynamical post-buckling analysis. Int. J. Non-Linear Mech. 5, 103–131 (2000)” and “Morales-Silva, D.M., Gao, D.Y.: Complete solutions and triality theory to a nonconvex optimization problem with double-well potential in $\mathbb{R}^n$. Numer. Algebra Contr. Optim. 3(2), 271–282 (2013)”, for “[42]” see Note [2] while “[51]”, “[57]”, “[58]”, “[79]”, “[80]” and “[113]” are the items [3], [4], [5], [19], [20] and [25] from our bibliography, respectively. Reference “[112]” seems to be the published version of [24].}
it is said “we need to assume that the canonical function $\Phi : \mathcal{E}_a \to \mathbb{R}$ is convex”. In [13, Def. 2] it is said: “A real-valued function $\Phi : \mathcal{E}_a \to \mathbb{R}$ is called canonical if the duality mapping $\partial \Phi : \mathcal{E}_a \to \mathcal{E}_a^*$ is one-to-one and onto”, while on [13, p. 10] it is said: “A canonical function $\Phi(\xi)$ can also be nonsmooth but should be convex such that its conjugate can be well-defined by Fenchel transformation $\Phi^*(\xi^*) = \sup\{\langle \xi, \xi^* \rangle - \Phi(\xi) \mid \xi \in \mathcal{E}_a\}$. This means that $V := \Phi \in \Gamma(\mathbb{R}^m)$, that is $V$ is the same as in [7, Th. 3]. However, the hypotheses of [13, Th. 3] are stronger than those of [7, Th. 3] because in the latter one asks $\sigma \in S_0$ (instead of $S_{col}$) and, for having the “double-min duality”, one assumes that $\dim \tilde{\nabla} = \dim \tilde{\nabla}$.

So, the framework of [7, Th. 3] is that of Proposition 2; however, applying the latter we obtain only the first assertions of [7, Th. 3] and [13, Th. 3]. Example 19 from [33] shows that the “double-max duality” and “double-min duality” are not true for $n = 2$ and $m \in \{1, 2\}$, taking $V := t_0(0) \in \Gamma(\mathbb{R})$ for $m := 1$ and $V := t_{\mathbb{R}^2} \in \Gamma(\mathbb{R}^2)$ for $m := 2$. Moreover, for $V$ from Example 7 in which $D$ is differentiable on its domain $[0, 1) \cup (1, \infty)$, one has that $D'(0) = 0$ and $0 \succ G(0) = A(0)$, but $\xi := x(0)$ and $\xi := 0$ are not simultaneously local minimizers (maximizers) for $P (= f)$ and $P_d$ on $\mathcal{X}_c (= [-1, 1])$ and $\{0, 1\} (= S_c^+)$, respectively.

In particular, Example 7 shows that the assertion “double-max duality is always true” from Q1 is false. In particular, even Theorems 3 in [5] and [4] are false because “double-max duality” is false, as mentioned above.

Because the proofs for the “bi-duality” given after 2011, less that in [14], are sufficiently involved and refer to $V$ in a restricted class of convex functions, the natural question is what is happening with the “bi-duality” results when $V \in \Gamma_{sc}(\mathbb{R}^m)$. So, we formulate the following open problem:

**Open problem.** Is the next statement true? Let $V \in \Gamma_{sc}(\mathbb{R}^m)$, $\sigma \in S^- \cap \text{int}(\text{dom} V^*)$ be a critical point of $D$, and $\bar{\sigma} := A(\sigma)^{-1}b(\sigma)$. Then $\bar{\sigma}$ is a local maximizer of $f$ on dom $f$ if and only if $\sigma$ is a local maximizer of $D$ on $S^-$; moreover, if $m = n$ then $\bar{\sigma}$ is a local minimizer of $f$ on dom $f$ if and only if $\sigma$ is a local minimizer of $D$ on $S^-$.

In this context let us quote from [13, p. 40] (or [12, NP 30]) and [9, p. 19], respectively:

Q4 “Six papers are in this group on the triality theory. By listing simple counterexamples (cf. e.g., [137]), Voisei and Zalinescu claimed: “a correction of this theory is impossible without falling into trivial”. However, even some of these counterexamples are correct, they are not new. This type of counterexamples was first discovered by Gao in 2003 [57, 58], i.e., the double-min duality holds under certain additional constraints (see Remark on page 288 [57] and Remark 1 on page 481 [58]). But neither [57] nor [58] was cited by Voisei and Zalinescu in their papers. ...”}

19 Which is the meaning of “$\dim \tilde{\nabla} = \dim \tilde{\nabla}$”? Why is not [13, Th. 3] attributed to [2] at least for $n = m = 1$?
20 Recall Q2 where it is said: “these additional constraints were left as an open problem (see Remark 1 in [12], also Theorem 3 and its Remark in a review article by Gao [13])”. In fact in [22] and [23] there are not open problems related to CDT. In Mathematical Economics there is an interesting axiom denoted NFL, and coming from “no free lunch”; this could be translated by ‘one gets nothing from nothing’, so for getting even the “double-max duality” one needs “certain additional conditions”.
21 The references “[137]”, “[57]” and “[58]” are [22], [5] and [4] from our bibliography, respectively.
22 Quite detailed answers to this kind of assertions can be found in [52] Sect. 2.
23 From the text “even some of these counterexamples are correct, they are not new” we have to understand
Regarding the so-called “not convincing proof”, serious researcher should provide either a convincing proof or a disproof, rather than a complaint.

Paraphrasing the text in Q5, we could say: Regarding the text in Q4, as a serious and honest researcher, DY Gao should have mentioned either his results which are not true, or even more, he should have written down those “additional constraints” under which the conclusions of those results become true, rather than the complaint that Voisei and Zalinescu never cited either [2] or [4].

7 Conclusions

– In Proposition 2 we showed, with a simple proof, that the “min-max duality” from the “triality theorem” for problem \((P)\) is true for \(V\) a proper lower semicontinuous convex function on \(\mathbb{R}^m\). Moreover, we showed that the “min-max duality” from quadratic minimization problems with quadratic constraints can be obtained using Proposition 2.

– We pointed out which are the relationships between the facts that \(\sigma\) and \(\pi := x(\sigma)\) are (strict) local maximizers (minimizers) of \(D\) on \(S^{-}\) and of \(f\) on \(dom f\), respectively, in the case in which \(V \in \Gamma_{sc}^2\) (see Proposition 9). In particular, in Corollary 10, we recovered Theorems 2 and 3 of [20] under [20, Assumption 3] for \(V \in \Gamma_{sc}^2\), our result being less precise that of [20, Th. 3] for \(m \neq n\); however, see the discussion about [20] in Section 6.

– In Section 6 we compared our results with those on “triality theorems” published by DY Gao and his collaborators after 2010, mentioning several drawbacks in proofs and inconsistencies in statements and presentations.

– We showed that the “double-min duality” and “double-max duality” of the general “triality theorem” from [13] are false even for \(m = n = 1\). – We formulated an open problem concerning the “double-min duality” and “double-max duality” when \(V \in \Gamma_{sc}\), problem related to that mentioned in Q2.

Acknowledgement We thank Prof. Marius Durea for reading a previous version of the paper and for his useful remarks.

References

[1] Y Chen, DY Gao, Global solutions to nonconvex optimization of 4th-order polynomial and log-sum-exp functions, J. Global Optim. 64 (2016), 417–431 (see also arXiv:1308.4732, 3 versions).

[2] DY Gao, Duality principles in nonconvex systems: theory, methods and applications, Dordrecht: Kluwer Academic Publishers (2000).

[3] DY Gao, Canonical dual transformation method and generalized triality theory in nonsmooth global optimization, J. Global Optim. 17 (2000), 127–160.

[4] DY Gao, Nonconvex semi-linear problems and canonical dual solutions, in: DY Gao, RW Ogden (eds) Advances in Mechanics and Mathematics, vol. II, pp. 261–312. Springer (2003).
[5] DY Gao, Perfect duality theory and complete solutions to a class of global optimization problems, Optimization 52 (2003), 467–493.

[6] DY Gao, Advantages in canonical duality theory with applications to global optimization, Proceedings of the Fifth International Conference on Foundations of Computer-Aided Process Operations, pp. 73–82. Omni Press, Cambridge, MA (2008) (almost the same as [7]).

[7] DY Gao, Canonical duality theory: Unified understanding and generalized solution for global optimization problems, Comput. Chem. Eng. 13 (2009), 1964–1972 (almost the same as [6]).

[8] DY Gao, Introduction to canonical duality theory (2009) (found with Google Scholar).

[9] D.Y. Gao, On unified modeling, canonical duality-triality theory, challenges and breakthrough in optimization, arXiv:1605.05534v3 (2016).

[10] DY Gao, V Latorre, N Ruan (eds), Canonical Duality Theory. Unified Methodology for Multidisciplinary Study, Advances in Mechanics and Mathematics 37. Cham: Springer (2017).

[11] DY Gao, N Ruan, Solutions and optimality criteria for nonconvex quadratic-exponential minimization problem, Math. Meth. Oper. Res. 67 (2008), 479–491.

[12] D.Y. Gao, N. Ruan, V. Latorre, RETRACTED: Canonical duality-triality theory: bridge between nonconvex analysis/mechanics and global optimization in complex system, Mathematics and Mechanics of Solids, 21(3) (2016), NP5–NP36 (see also arXiv:1410.2665).

[13] D.Y. Gao, N. Ruan, V. Latorre, Canonical duality-triality theory: bridge between nonconvex analysis/mechanics and global optimization in complex system, in: DY Gao, V Latorre, N Ruan (eds) Canonical Duality Theory. Advances in Mechanics and Mathematics 37, pp. 1–47. Springer, Cham (2017).

[14] DY Gao, N Ruan, PM Pardalos, Canonical dual solutions to sum of fourth-order polynomials minimization problems with applications to sensor network localization, In: V Boginski, C Commander, P Pardalos, Y Ye (eds) Sensors: Theory, Algorithms, and Applications. Springer Optimization and Its Applications 61, pp. 37–54. Springer, New York, NY (2012).

[15] DY Gao, H Sherali, Canonical duality theory: Connections between nonconvex mechanics and global optimization, in: DY Gao, H Sherali (eds) Advances in Applied Mathematics and Global Optimization, pp. 257–326. Springer, USA (2009).

[16] DY Gao, C Wu, On the triality theory in global optimization. (I) Unconstrained problems, arXiv:1104.2970v1 [math.OC] 15 Apr 2011 (to be compared with [17] and [19]).

[17] DY Gao, C Wu, On the triality theory for a quartic polynomial optimization problem, arXiv:1110.0293v1 [math.OC] 3 Oct 2011 (to be compared with [16] and [19]).

[18] DY Gao, C Wu, Triality theory for general unconstrained global optimization problems, arXiv:1104.2970v2 [math.OC] 20 Feb 2012 (to be compared with [20]).
[19] DY Gao, C Wu, *On the triality theory for a quartic polynomial optimization problem*, J. Ind. Manag. Optim. 8 (2012), 229–242 (to be compared with [16] and [17]).

[20] DY Gao, C Wu, *Triality theory for general unconstrained global optimization problems*, in: D Gao, V Latorre, N Ruan (eds) Canonical Duality Theory. Advances in Mechanics and Mathematics 37, pp. 127–153. Springer, Cham (2017) (to be compared with [18]).

[21] Z Jin, DY Gao, *On modeling and global solutions for d.c. optimization problems by canonical duality theory*, Applied Mathematics and Computation 296 (2017), 168–181 (see also arXiv:1607.03426).

[22] Z Jin, DY Gao, *On d.c. optimization problems*, in: D Gao, V Latorre, N Ruan (eds) Canonical Duality Theory. Advances in Mechanics and Mathematics 37, pp. 203–221. Springer, Cham (2017).

[23] D Morales-Silva, DY Gao, *Complete solutions and triality theory to a nonconvex optimization problem with double-well potential in $\mathbb{R}^n$*, arXiv:1110.0285v1 (2011).

[24] D Morales-Silva, DY Gao, *Canonical duality theory and triality for solving general non-constrained global optimization problems*, arXiv:1210.0180v1 (2012) (the same as [25]).

[25] D Morales-Silva, DY Gao, *Canonical duality theory and triality for solving general global optimization problems in complex systems*, Math. Mech. Compl. Sys. 3 (2015), 139–161.

[26] RT Rockafellar, *Convex Analysis*, Princeton University Press, N.J. (1972).

[27] N Ruan, DY Gao, *Global optimal solutions to a general sensor network localization problem*, Perform. Eval. 75–76 (2014), 1–16 (see also arXiv:1302.3673v1).

[28] R Strugariu, MD Voisei, C Zalinescu, *Counter-examples in bi-duality, triality and tri-duality*, Discrete and Continuous Dynamical Systems - Series A (DCDS-A) 31 (2011), 1453–1468

[29] MD Voisei, C Zalinescu, *Counterexamples to some triality and tri-duality results*, J. Glob. Optim. 49 (2011), 173–183.

[30] MD Voisei, C Zalinescu, *Counterexamples to a triality theorem for quadratic-exponential minimization problems*, Math. Meth. Oper. Res. 77 (2013), 227–237.

[31] C Zalinescu, *On second-order generalized convexity*, J. Optim. Theory Appl. 168 (2016), 802–829.

[32] C Zalinescu, *On V. Latorre and D.Y. Gao’s paper “Canonical duality for solving general nonconvex constrained problems”*, Optim. Lett. 10 (8) (2016), 1781–1787.

[33] C Zalinescu, *On quadratic optimization problems and canonical duality theory*, arXiv:1809.09032v1 (2018).