DUALITY THEORY: III. TRI-DUALITY IN GLOBAL OPTIMIZATION

Consider the following general nonconvex extremum problem $(\mathcal{P})$

$$P(x) = \Phi(x, \Lambda(x)) \to \text{extremum } \forall x \in \mathcal{X}, \quad (1)$$

where $\mathcal{X}$ is a locally convex topological vector space (l.c.s.), $P : \mathcal{X} \to \bar{\mathbb{R}} := \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$ is a nonconvex and nonsmooth extended function, whose effective domain

$$\mathcal{X}_k = \mathrm{dom} P = \{x \in \mathcal{X} | \ |P(x)| < +\infty\}$$

is a non-empty, convex subset of $\mathcal{X}$; the operator $\Lambda : \mathcal{X} \to \mathcal{Y}$ is a continuous, generally nonlinear, mapping from $\mathcal{X}$ to another l.c.s. $\mathcal{Y}$, and $\Phi : \mathcal{X} \times \mathcal{Y} \to \bar{\mathbb{R}}$ is an associated extended function. Since the cost function $P(x)$ is usually nonconvex, the problem $(\mathcal{P})$ may possess many locally extremum (either minimum or maximum) solutions. The goal of *global optimization* is to find all the local extrema of $P(x)$ over the feasible set $\mathcal{X}_k$. Generally speaking, traditional direct approaches and algorithms for solving nonconvex, nonsmooth global optimization problems are usually very difficult. The classical saddle Lagrange duality methods as well as the well-known Fenchel-Rockafellar duality theory can be used mainly for solving convex problems. For nonconvex problems, there exists a so-called *duality gap* between the primal and the classical dual problems.

Canonical dual transformation method and associated *tri-duality theory* were proposed originally in finite deformation theory [1]. The key idea of this method is to choose a suitable nonlinear operator $\Lambda : \mathcal{X} \to \mathcal{Y}$ such that $\Phi(x, y)$ is either convex or concave in each of its variables. This method can be used to solve many nonconvex, nonsmooth global optimization problems.

Canonical Dual Transformation.

Let $(\mathcal{X}, \mathcal{X}^*)$ be a pair of real linear spaces, placed in duality by a bilinear form $\langle \cdot , \cdot \rangle : \mathcal{X} \times \mathcal{X}^* \to \mathbb{R}$. For a given extended real-valued function $P : \mathcal{X} \to \bar{\mathbb{R}}$, the *sub-differential* of P at $\bar{x} \in \mathcal{X}$ is a convex subset $\partial^- P(\bar{x}) \subset \mathcal{X}^*$ such that for each $\bar{x}^* \in \partial^- P(\bar{x})$, we have

$$\langle \bar{x}^*, x - \bar{x} \rangle \le P(x) - P(\bar{x}) \ \ \forall x \in \mathcal{X}.$$

Dually, the *super-differential* of P at $\bar{x} \in \mathcal{X}$ is a convex subset $\partial^+ P(\bar{x}) \subset \mathcal{X}^*$ such that for each $\bar{x}^* \in \partial^+ P(\bar{x})$, we have

$$\langle \bar{x}^*, x - \bar{x} \rangle \ge P(x) - P(\bar{x}) \ \ \forall x \in \mathcal{X}.$$

Clearly, we always have $\partial^+ P = -\partial^-(-P)$. In convex analysis, it is convention that ∂^- is simply written as ∂. In nonconvex analysis, ∂ stands for either ∂^- or ∂^+, i.e.

$$\partial = \{\partial^-, \ \partial^+\}.$$

If P is smooth, Gâteaux-differentiable at $\bar{x} \in \mathcal{X}_a \subset \mathcal{X}$, then

$$\partial P(\bar{x}) = \partial^- P(\bar{x}) = \partial^+ P(\bar{x}) = \{DP(\bar{x})\},$$

where $DP : \mathcal{X}_a \to \mathcal{X}^*$ denotes the Gâteaux derivative of P at $\bar{x}$.

Definition 1 The set of functions $P : \mathcal{X} \to \bar{\mathbb{R}}$ which are either convex or concave is denoted by $\Gamma(\mathcal{X})$. In particular, let $\check{\Gamma}(\mathcal{X})$ denote the subset of functions $P \in \Gamma(\mathcal{X})$ which are convex and $\hat{\Gamma}(\mathcal{X})$ the subset of $P \in \Gamma(\mathcal{X})$ which are concave.

The *canonical function space* $\Gamma_G(\mathcal{X}_a)$ is a subset of functions $P \in \Gamma(\mathcal{X}_a)$ which are Gâteaux differentiable on $\mathcal{X}_a \subset \mathcal{X}$ and the duality mapping $DP : \mathcal{X}_a \to \mathcal{X}_a^* \subset \mathcal{X}^*$ is invertible.

The *extended canonical function space* $\Gamma_0(\mathcal{X})$ is a subset of functions $P \in \Gamma(\mathcal{X})$ which are either convex, lower semicontinuous or concave, upper semicontinuous, and if P takes the values $\pm\infty$, then P is identically equal to $\pm\infty$. ◇

---

global optimization→ *global optimization*
duality gap→ *duality gap*
Canonical dual transformation method→ *Canonical dual transformation method*
tri-duality theory→ *tri-duality theory*
sub-differential→ *sub-differential*
super-differential→ *super-differential*
canonical function space→ *canonical function space*
extended canonical function space→ *extended canonical function space*

By the Legendre-Fenchel transformation, the *super-conjugate function* of an extended function $P : \mathcal{X} \to \bar{\mathbb{R}}$ is defined by

$$P^\sharp(x^*) = \sup_{x\in\mathcal{X}} \{\langle x, x^*\rangle - P(x)\}.$$

By the theory of convex analysis, $P^\sharp : \mathcal{X}^* \to \vec{\mathbb{R}} := \mathbb{R} \cup \{+\infty\}$ is always convex and lower semicontinuous, i.e. $P^\sharp \in \check{\Gamma}_0(\mathcal{X}^*)$. Dually, the *sub-conjugate function* of P, defined by

$$P^\flat(x^*) = \inf_{x\in\mathcal{X}} \{\langle x, x^*\rangle - P(x)\},$$

is always concave and upper semicontinuous, i.e. $P^\flat \in \hat{\Gamma}_0(\mathcal{X}^*)$, and $P^\flat = -P^\sharp$. Both the super- and sub-conjugates are called Fenchel conjugate functions and we write $P^* = \{P^\flat, P^\sharp\}$. Thus the extended Fenchel transformation can be written as

$$P^*(x^*) = \text{ext}\{\langle x, x^*\rangle - P(x) |\ \ \forall x \in \mathcal{X}\}. \tag{2}$$

Clearly, if $P \in \Gamma_0(\mathcal{X})$, we have the Fenchel equivalent relations, namely,

$$\begin{aligned} & x^* \in \partial P(x) \ \ \Leftrightarrow \\ & x \in \partial P^*(x^*) \ \ \Leftrightarrow \\ & P(x) + P^*(x^*) = \langle x, x^*\rangle. \end{aligned} \tag{3}$$

The pair (x, x^*) is called the *Fenchel duality pair* on $\mathcal{X} \times \mathcal{X}^*$ if and only if equation (3) holds on $\mathcal{X} \times \mathcal{X}^*$.

The conjugate pair (x, x^*) is said to be a *Legendre duality pair* on $\mathcal{X}_a \times \mathcal{X}_a^* \subset \mathcal{X} \times \mathcal{X}^*$ if and only if the equivelant relations

$$\begin{aligned} & x^* = DP(x) \ \ \Leftrightarrow \\ & x = DP^*(x^*) \ \ \Leftrightarrow \\ & P(x) + P^*(x^*) = \langle x, x^*\rangle \end{aligned} \tag{4}$$

hold on $\mathcal{X}_a \times \mathcal{X}_a^*$.

Let $(\mathcal{Y}, \mathcal{Y}^*)$ be an another pair of locally convex topological real linear spaces paired in separating duality by the second bilinear form $\langle \cdot ; \cdot \rangle : \mathcal{Y} \times \mathcal{Y}^* \to \mathbb{R}$. The so-called *geometrical operator* $\Lambda : \mathcal{X} \to \mathcal{Y}$ is a continuous, Gâteaux differentiable operator such that for any given $x \in \mathcal{X}_a \subset \mathcal{X}$, there exists a $y \in \mathcal{Y}_a \subset \mathcal{Y}$ satisfying the *geometrical equation*

$$y = \Lambda(x).$$

The directional derivative of y at $\bar{x}$ in the direction $x \in \mathcal{X}$ is then definded by

$$\delta y(\bar{x}; x) := \lim_{\theta\to 0^+} \frac{y(\bar{x} + \theta x) - y(\bar{x})}{\theta} = \Lambda_t(\bar{x})x,$$

where $\Lambda_t(\bar{x}) = D\Lambda(\bar{x})$ denotes the Gâteaux derivative of the operator Λ at $\bar{x}$. For a given $y^* \in \mathcal{Y}^*$, $\ell(x) = \langle \Lambda(x); y^*\rangle$ is a real-valued function of x on $\mathcal{X}$. Its Gâteaux derivative at $\bar{x} \in \mathcal{X}_a$ in the direction $x \in \mathcal{X}$ is

$$\delta\ell(\bar{x}; x) = \langle \Lambda_t(\bar{x})x; y^*\rangle = \langle x, \Lambda_t^*(\bar{x})y^*\rangle,$$

where $\Lambda_t^*(\bar{x}) : \mathcal{Y}^* \to \mathcal{X}^*$ is the adjoint operator of Λ_t associated with the two bilinear forms.

Let $\Phi : \mathcal{X} \times \mathcal{Y} \to \bar{\mathbb{R}}$ be an extended function such that $P(x) = \Phi(x, \Lambda(x))$. If $\Phi : \mathcal{X} \times \mathcal{Y} \to \bar{\mathbb{R}}$ is an extended canonical function, i.e. $\Phi \in \Gamma_0(\mathcal{X}) \times \Gamma_0(\mathcal{Y})$, the duality relations between the paired spaces $(\mathcal{X}, \mathcal{X}^*)$ and $(\mathcal{Y}, \mathcal{Y}^*)$ can be written as

$$x^* \in \partial_x \Phi(x, y), \ \ y^* \in \partial_y \Phi(x, y). \tag{5}$$

On the product space $\mathcal{X}_a \times \mathcal{Y}_a \subset \mathcal{X} \times \mathcal{Y}$, if the canonical function $\Phi(x, y)$ is finite and Gâteaux differentiable such that the feasible space $\mathcal{X}_k$ can be written as

$$\mathcal{X}_k = \{x \in \mathcal{X}_a |\ \ \Lambda(x) \in \mathcal{Y}_a\}, \tag{6}$$

then on $\mathcal{X}_k$, the critical condition $\delta P(\bar{x}; x) = \langle x, DP(\bar{x})\rangle = 0 \ \ \forall x \in \mathcal{X}_k$ leads to the Euler equation

$$D_x\Phi(\bar{x}, \Lambda(\bar{x})) + \Lambda_t^*(\bar{x}) D_y \Phi(\bar{x}, \Lambda(\bar{x})) = 0, \tag{7}$$

where $D_x\Phi$ and $D_y\Phi$ denote the partial Gâteaux derivatives of Φ with respect to x and y, respectively. Since $\Phi \in \Gamma_G(\mathcal{X}_a) \times \Gamma_G(\mathcal{Y}_a)$ is a canonical function, the Gâteaux derivative $D\Phi : \mathcal{X}_a \times \mathcal{Y}_a \to \mathcal{X}_a^* \times \mathcal{Y}_a^* \subset \mathcal{X}^* \times \mathcal{Y}^*$ is a monotone mapping, i.e. there exists a pair $(\bar{x}^*, \bar{y}^*) \in \mathcal{X}^* \times \mathcal{Y}^*$ such that

$$-\bar{x}^* = D_x\Phi(\bar{x}, \Lambda(\bar{x})), \quad \bar{y}^* = D_y\Phi(\bar{x}, \Lambda(\bar{x})).$$

---

super-conjugate function→ super-conjugate function
sub-conjugate function→ sub-conjugate function
Fenchel duality pair→ Fenchel duality pair
Legendre duality pair→ Legendre duality pair
geometrical operator→ geometrical operator
geometrical equation→ geometrical equation

Thus, in terms of canonical dual variables $\bar{x}^*$ and $\bar{y}^*$, the Euler equation (7) can be written in the so-called *balance (or equilibrium) equation*

$$\bar{x}^* = \Lambda_t^*(\bar{x})\bar{y}^*, \tag{8}$$

which linearly depends on the dual variable $\bar{y}^*$.

Definition 2 Suppose that for a given problem $(\mathcal{P})$, the geometrical operator $\Lambda : \mathcal{X} \to \mathcal{Y}$ can be chosen in such a way that $P(x) = \Phi(x, \Lambda(x))$, $\Phi \in \Gamma_G(\mathcal{X}_a) \times \Gamma_G(\mathcal{Y}_a)$ and $\mathcal{X}_k = \{x \in \mathcal{X}_a | \; \Lambda(x) \in \mathcal{Y}_a\}$. Then

(1) the transformation $\{P; \mathcal{X}_k\} \to \{\Phi; \mathcal{X}_a \times \mathcal{Y}_a\}$ is called the *canonical transformation* , and $\Phi : \mathcal{X}_a \times \mathcal{Y}_a \to \mathbb{R}$ is called the *canonical function associated with* Λ;

(2) the problem $(\mathcal{P})$ is called *geometrically nonlinear (resp. linear)* if $\Lambda : \mathcal{X} \to \mathcal{Y}$ is nonlinear (resp. linear); it is called *physically nonlinear* (resp. linear) if the duality mapping $D\Phi : \mathcal{X}_a \times \mathcal{Y}_a \to \mathcal{X}_a^* \times \mathcal{Y}_a^*$ is nonlinear (resp. linear); it is called *fully nonlinear* if it is both geometrically and physically nonlinear. ◇

The canonical transformation plays a fundamental role in duality theory of global optimization. By this definition, the governing equation (7) for fully nonlinear problems can be written in the *tri-canonical forms*, namely,

(1) geometrical equation: $y = \Lambda(x)$,
(2) physical relations: $(-x^*, y^*) \in \partial\Phi(x, y)$,
(3) balance equation: $x^* = \Lambda_t^*(x)y^*$.

Since $\Lambda : \mathcal{X} \to \mathcal{Y}$ is Gâteaux differentiable, for any given $x \in \mathcal{X}$, we have the *operator decomposition*

$$\Lambda(x) = \Lambda_t(x)x + \Lambda_c(x), \tag{9}$$

where $\Lambda_c = \Lambda - \Lambda_t$ is the *complementary operator* of Λ_t. By this operator decomposition, the relation between the two bilinear forms reads

$$\langle \Lambda(x); y^* \rangle = \langle x, \Lambda_t^*(x) y^* \rangle - G(x, y^*),$$

where $G(x, y^*) = \langle -\Lambda_c(x); y^* \rangle$ is the so-called *complementary gap function*, introduced in [2]. This gap plays an important role in the canonical dual transformation methods. A framework for the fully nonlinear system is shown in Fig. 1. Extensive illustrations of the canonical transformation and the tri-canonical forms in mathematical physics and variational analysis can be found in [1].

$$\begin{tikzcd}
x \in \mathcal{X} \arrow[r, leftrightarrow, "\langle x , x^* \rangle"] \arrow[d, "\Lambda_t + \Lambda_c = \Lambda"'] & \mathcal{X}^* \ni x^* \\
y \in \mathcal{Y} \arrow[r, leftrightarrow, "\langle y ; y^* \rangle"] & \mathcal{Y}^* \ni y^* \arrow[u, "\Lambda_t^* = (\Lambda - \Lambda_c)^*"']
\end{tikzcd}$$

Figure 1. Framework in fully nonlinear systems

Very often, the extended canonical function Φ can be written in the form

$$\Phi(x, y) = W(y) - F(x),$$

where $F \in \Gamma(\mathcal{X})$ and $W \in \Gamma(\mathcal{Y})$ are extended canonical functions. The duality relations (5) in this special case take the forms

$$x^* \in \partial F(x), \quad y^* \in \partial W(y).$$

If $F \in \Gamma_G(\mathcal{X}_a)$ and $W \in \Gamma_G(\mathcal{Y}_a)$ are Gâteaux differentiable, the Euler equation (7) reads

$$\Lambda_t^*(\bar{x}) DW(\Lambda(\bar{x})) - DF(\bar{x}) = 0.$$

If $\Lambda : \mathcal{X} \to \mathcal{Y}$ is linear, and $W : \mathcal{Y} \to \mathbb{R}$ is quadratic such that $DW = Cy$, where $C : \mathcal{Y} \to \mathcal{Y}^*$ is a linear operator, then the governing equations for linear system can be written as

$$\Lambda^* C \Lambda x = Ax = x^*.$$

For conservative systems, the operator $A = \Lambda^* C \Lambda$ is usually symmetric. In static systems, C is usually positive-definite and the associated total potential P is convex. However, in dynamical systems, C is indefinite and P is called the total action, which is usually a d.c. function in

---

balance (or equilibrium) equation→ *balance equation*
canonical transformation→ *canonical transformation*
physically nonlinear→ *physically nonlinear problem*
fully nonlinear→ *fully nonlinear problem*
tri-canonical forms→ *tri-canonical forms*
operator decomposition → *operator decomposition*
complementary operator→ *complementary operator*
complementary gap function→ *complementary gap function*

convex Hamilton systems.

Triality Theory.

We assume that for any given nonconvex extended function $P : \mathcal{X} \to \bar{\mathbb{R}}$, there exists a general nonlinear operator $\Lambda : \mathcal{X} \to \mathcal{Y}$ and a canonical function $W \in \Gamma(\mathcal{Y})$ such that the canonical transformation can be written as

$$P(x) = W(\Lambda(x)) - \langle x, c \rangle, \tag{10}$$

where $c \in \mathcal{X}^*$ is a given source variable. Since $F(x) = \langle x, c \rangle$ is a linear function, the Hamiltonian $H(x, y^*) = W^*(y^*) + \langle x, c \rangle$ is a canonical function on $\mathcal{Z} = \mathcal{X} \times \mathcal{Y}^*$ and the extended Lagrangian reads

$$L(x, y^*) = \langle \Lambda(x); y^* \rangle - W^*(y^*) - \langle x, c \rangle. \tag{11}$$

For a fixed $y^* \in \mathcal{Y}^*$, the convexity of $L(\cdot, y^*) : \mathcal{X} \to \bar{\mathbb{R}}$ depends on $\Lambda(x)$ and $y^* \in \mathcal{Y}^*$.

Let $\mathcal{Z}_a = \mathcal{X}_a \times \mathcal{Y}^*_a \subset \mathcal{Z}$ be the effective domain of L, and let $\mathcal{L}_c \subset \mathcal{Z}_a$ be a critical point set of L, i.e.

$$\mathcal{L}_c = \{(\bar{x}, \bar{y}) \in \mathcal{X}_a \times \mathcal{Y}^*_a | \ DL(\bar{x}, \bar{y}^*) = 0\}.$$

For any given critical point $(\bar{x}, \bar{y}^*) \in \mathcal{L}_c$, we let $\mathcal{X}_r \times \mathcal{Y}^*_r$ be its neighborhood such that on $\mathcal{X}_r \times \mathcal{Y}^*_r$, the pair $(\bar{x}, \bar{y}^*)$ is the only critical point of L. The following result is of fundamental importance in global optimization.

Theorem 1 (Triality Theorem) *Suppose that $W \in \check{\Gamma}(\mathcal{Y}_a)$ is convex, $(\bar{x}, \bar{y}^*) \in \mathcal{L}_c$ is a critical point of L and $\mathcal{X}_r \times \mathcal{Y}^*_r$ is a neighborhood of $(\bar{x}, \bar{y}^*)$.*

If $\langle \Lambda(x); \bar{y}^* \rangle$ is convex on $\mathcal{X}_r$, then

$$\begin{aligned} L(\bar{x}, \bar{y}^*) &= \min_{x \in \mathcal{X}_r} \max_{y^* \in \mathcal{Y}^*_r} L(x, y^*) \\ &= \max_{y^* \in \mathcal{Y}^*_r} \min_{x \in \mathcal{X}_r} L(x, y^*). \end{aligned} \tag{12}$$

However, if $\langle \Lambda(x); \bar{y}^* \rangle$ is concave on $\mathcal{X}_r$, then either

$$\begin{aligned} L(\bar{x}, \bar{y}^*) &= \min_{x \in \mathcal{X}_r} \max_{y^* \in \mathcal{Y}^*_r} L(x, y^*) \\ &= \min_{y^* \in \mathcal{Y}^*_r} \max_{x \in \mathcal{X}_r} L(x, y^*), \end{aligned} \tag{13}$$

or

$$\begin{aligned} L(\bar{x}, \bar{y}^*) &= \max_{x \in \mathcal{X}_r} \max_{y^* \in \mathcal{Y}^*_r} L(x, y^*) \\ &= \max_{y^* \in \mathcal{Y}^*_r} \max_{x \in \mathcal{X}_r} L(x, y^*). \end{aligned} \tag{14}$$

Since $W \in \Gamma(\mathcal{Y}_a)$ is a canonical function, we always have

$$P(x) = \text{ext}\{L(x, y^*) | \ y^* \in \mathcal{Y}^*\} \ \forall x \in \mathcal{X}_k. \tag{15}$$

On the other hand, for a given Gâteaux differentiable geometrical mapping $\Lambda : \mathcal{X}_a \to \mathcal{Y}_a$, the criticality condition $D_x L(\bar{x}, y^*) = 0$ leads to the equilibrium equation

$$\Lambda^*_t(\bar{x}) y^* = c. \tag{16}$$

If there exists a subspace $\mathcal{Y}^*_s \subset \mathcal{Y}^*_a$ such that for any $y^* \in \mathcal{Y}^*_s$ and a given source variable $c \in \mathcal{X}^*$, the equation (16) can be solved for $\bar{x} = \bar{x}(y^*)$, then by the operator decomposition (9), the dual function $P^d : \mathcal{Y}^*_s \to \mathbb{R}$ can be written explicitly in the form

$$\begin{aligned} P^d(y^*) &= \text{sta}\{L(x, y^*) | \ x \in \mathcal{X}\} \\ &= -G^d(y^*) - W^*(y^*) \ \ \forall y^* \in \mathcal{Y}^*_s, \end{aligned}$$

where $G^d : \mathcal{Y}^* \to \mathbb{R}$ is the so-called pure complementary gap function, defined by

$$G^d(y^*) = G(\bar{x}(y^*), y^*) = -\langle \Lambda_c(\bar{x}(y^*)); y^* \rangle.$$

For any given critical point $(\bar{x}, \bar{y}^*) \in \mathcal{L}_c$, we have $G^d(\bar{y}^*) = \langle \bar{x}, c \rangle - \langle \Lambda(\bar{x}(\bar{y}^*)); \bar{y}^* \rangle$. Thus, the Legendre duality relations among the canonical functions W and W^* lead to

$$P(\bar{x}) - P^d(\bar{y}^*) = 0 \ \ \forall (\bar{x}, \bar{y}^*) \in \mathcal{L}_c. \tag{17}$$

This identity shows that there is no duality gap between the nonconvex function P and its canonical dual function P^d. Actually the duality gap, which exists in classical duality theories, is now recovered by the complementary gap function $G(\bar{x}, \bar{y}^*)$.

Theorem 2 (Tri-Duality Theorem) *Suppose that $W \in \check{\Gamma}(\mathcal{Y}_a)$, $(\bar{x}, \bar{y}^*) \in \mathcal{L}_c$ is a critical point of L and $\mathcal{X}_r \times \mathcal{Y}^*_r$ is a neighborhood of $(\bar{x}, \bar{y}^*)$. If $\langle \Lambda(x); \bar{y}^* \rangle$ is convex on $\mathcal{X}_r$, then*

$$P(\bar{x}) = \min_{x \in \mathcal{X}_r} P(x) \ \Leftrightarrow \ P^d(\bar{y}^*) = \max_{y^* \in \mathcal{Y}^*_r} P^d(y^*).$$

However, if $\langle \Lambda(x); \bar{y}^* \rangle$ is concave on $\mathcal{X}_r$, then

$$\begin{aligned} P(\bar{x}) = \min_{x \in \mathcal{X}_r} P(x) \quad &\Leftrightarrow \quad P^d(\bar{y}^*) = \min_{y^* \in \mathcal{Y}^*_r} P^d(y^*); \\ P(\bar{x}) = \max_{x \in \mathcal{X}_r} P(x) \quad &\Leftrightarrow \quad P^d(\bar{y}^*) = \max_{y^* \in \mathcal{Y}^*_r} P^d(y^*). \end{aligned}$$

Example. We now illustrate the application of the interesting tri-duality theory for solving the following nonconvex optimization problem in $\mathcal{X} = \mathbb{R}^n$

$$P(x) = \frac{a}{2}(\frac{1}{2} \parallel Ax \parallel^2 - \mu)^2 - x^T c \to \text{sta} \ \ \forall x,$$

where $a, \mu > 0$ are given parameters, $c \in \mathbb{R}^n$ is a given vector, and $A : \mathbb{R}^n \to \mathbb{R}^m$ is a matrix. The Euler equation associated with this nonconvex stationary problem is a nonlinear algebraic equation in $\mathbb{R}^n$

$$a(\frac{1}{2} \parallel A\bar{x} \parallel^2 - \mu) C\bar{x} = c,$$

where $C = A^T A = C^T \in \mathbb{R}^{nn}$. We are interested in finding all the critical points of P. To set this nonconvex problem in our framework, we let $\mathcal{X} = \mathbb{R}^n = \mathcal{X}^*$, and $\Lambda : \mathbb{R}^n \to \mathcal{Y} = \mathbb{R}$ a quadratic operator

$$y = \Lambda(x) = \frac{1}{2} \parallel Ax \parallel^2 - \mu = \frac{1}{2} x^T C x - \mu.$$

Since $F(x) = \langle x, c \rangle = x^T c$ is a linear function on $\mathbb{R}^n$, the admissible space $\mathcal{X}_a = \mathcal{X} = \mathbb{R}^n$. By the fact that $x^* = DF(x) = c$, the range for the canonical mapping $DF : \mathcal{X} \to \mathcal{X}^* = \mathbb{R}$ is a hyperplane in $\mathbb{R}^n$, i.e.

$$\mathcal{X}_a^* = \{x^* \in \mathbb{R}^n | \ \ x^* = c\}.$$

The feasible set for the primal problem is $\mathcal{X}_k = \{x \in \mathcal{X}_a | \ \ \Lambda(x) \in \mathcal{Y}_a\} = \mathbb{R}^n$.

By the fact that $x^T C x \geq 0 \quad \forall x \in \mathcal{X}_a = \mathcal{X} = \mathbb{R}^n$, the range for the geometrical mapping $\Lambda : \mathcal{X}_a \to \mathbb{R}$ is a closed convex set in $\mathbb{R}$

$$\mathcal{Y}_a = \{y \in \mathbb{R} | \ \ y \geq -\mu\} \subset \mathcal{Y} = \mathbb{R}.$$

On the admissible subset $\mathcal{Y}_a \subset \mathcal{Y} = \mathbb{R}$, the canonical function $W(y) = \frac{1}{2} a y^2$ is quadratic. The range for the constitutive mapping $DW : \mathcal{Y}_a \to \mathcal{Y}^* = \mathbb{R}$ is also a closed convex set in $\mathbb{R}$

$$\mathcal{Y}_a^* = \{y^* \in \mathbb{R} | \ \ y^* \geq -a\mu\}.$$

On $\mathcal{Y}_a^*$, the Legendre conjugate of W is also strictly convex

$$W^*(y^*) = \frac{1}{2} a^{-1} y^{*2}, \tag{18}$$

and the Legendre duality relations hold on $\mathcal{Y}_a \times \mathcal{Y}_a^*$.

On $\mathcal{X}_a \times \mathcal{Y}_a^* = \mathbb{R}^n \times \mathbb{R}$, the extended Lagrangian in this case reads

$$L(x, y^*) = \frac{1}{2} y^* x^T C x - \mu y^* - \frac{1}{2} a^{-1} y^{*2} - x^T c.$$

It is easy to check that the dual function associated with L is

$$P^d(y^*) = \frac{1}{2} (y^*)^{-1} c^T C c - \mu y^* - \frac{1}{2a} y^{*2}.$$

The dual Euler-Lagrange equation is an algebraic equation in $\mathbb{R}$:

$$(\mu + a^{-1} y^*) y^{*2} = \frac{1}{2} \sigma^2, \tag{19}$$

where $\sigma^2 = c^T C c$ is a constant. Since $C \in \mathbb{R}^{nn}$ is positive-definite, this equation holds only on $\mathcal{Y}_a^*$. In algebraic geometry, the dual Euler-Lagrange equation (19) is the so-called singular algebraic curve in (y^*, σ)-space (see Fig. 2). For a given parameter μ and $c \in \mathbb{R}^n$, this dual equation has at most three real roots $y_k^* \in \mathcal{Y}_a^*, \ \ k = 1, 2, 3$, which leads to the primal solution

$$x_k = y_k^* C^+ c, \ \ k = 1, 2, 3.$$

By Lemma 1 we know that each (x_k, y_k^*) is a critical point of L and

$$P(x_k) = L(x_k, y_k^*) = P^d(y_k^*), \ \ k = 1, 2, 3.$$

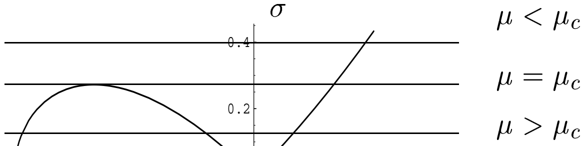
In the case of $n = 1$, the cost function

$$P(x) = \frac{1}{2} a (x^2 - \mu)^2 - cx$$

is a *double-well function* (see Fig. 3 (solid line)), which appears in many physical systems. The graph of the canonical dual function

$$P^d(y^*) = \frac{1}{2} c^2 / y^* - \mu y^* - y^{*2} / 2a$$

has two branches (dashed line). It is easy to prove (see [1]) that if $\mu > \mu_c = 1.5(\sigma/a)^{2/3}$, the dual Euler-Lagrange equation (19) has three roots $y_1^* > 0 > y_2^* > y_3^*$, corresponding to three critical points of P^d (see Fig. 3). Then, y_1^* is a global maximizer of P^d, $x_1 = \sigma / y_1^*$ is a global minimizer of P, P^d takes local minimum and local maximum values at y_2^* and y_3^*, respectively, $x_2 = \sigma / y_2^*$ is a local maximizer of P, while $x_3 = \sigma / y_3^*$ is a local minimizer.



---

double-well function→ double-well function

Figure 2. Singular algebraic curve

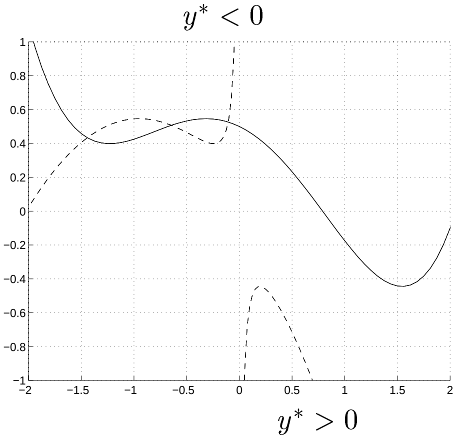


Figure 3. Graphs of $P(u)$ and its dual $P^d(y^*)$

The Lagrangian associated with this double-well energy is

$$L(x,y^*) = \frac{1}{2}x^2y^* - (\frac{1}{2a}y^{*2} + \mu y^*) - y^*x.$$

It is a saddle function for $y^* > 0$. If $y^* < 0$, it is a super-critical point function (see Fig. 4).

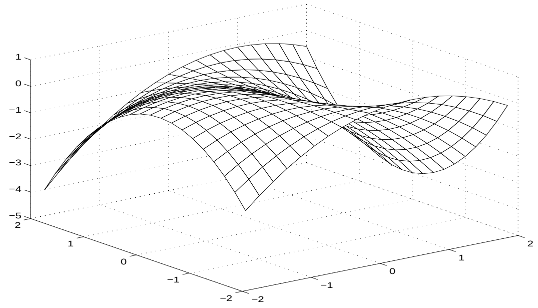


Figure 4. Lagrangian for the double-well energy

References

[1] Gao, DY (1999). *Duality Principles in Nonconvex Systems: Theory, Methods and Applications*, Kluwer Academic Publishers, Dordrecht /Boston / London, 472pp.

[2] Gao, DY and Strang, G. (1989). Geometrical nonlinearity: Potential energy, complementary energy, and the gap functional, *Quart. Appl. Math.*, XLVII(3), 487-504.

David Yang Gao
Virginia Polytechnic Institute and State University
Blacksburg, Virginia, 24061
USA
E-mail address: `gao@math.vt.edu`

AMS 1991 Subject Classification: 49, 90, 93.

Key words and phrases: duality, nonconvexity, global minimization, super-Lagrangian, triduality, triality.