Control and Cybernetics

vol. **30** (2001) No. 1

On some connection between invex and convex problems in nonlinear programming

by

Marek Galewski

Faculty of Mathematics, University of Łódź
Banacha 22, 90-238 Łódź, Poland

Abstract: A method of solving problems involving invex functions via certain convex problems is presented. Nonsmooth problems are also considered. A definition of such a Fenchel-Young type duality for an invex function f that its second dual is equal to f is provided.

Keywords: invexity, Kuhn-Tucker conditions, Fenchel-Young duality.

1. Introduction

Relationships between convexity and invexity and consequently between invex and convex nonlinear programming problems have been the subject of thorough research, see for example Pini (1994), Craven (1995), Hanson, Mond (1987). In the present paper some other connections between invexity and convexity are shown. In Section 2 we consider invex and convex nonlinear programming problems. Relations between Karush-Kuhn-Tucker points and solutions to both problems are given. A formula for a convex function which is related to a given invex one is derived. In Section 3 some results from Section 2 are generalized to locally Lipschitz functions. In Section 4 we provide the definition of such a Fenchel-Young type transform, Rockafellar, Wets (1998), operating on invex functions, that the second dual gives the original function. These results, which are based on the relationships between the subclass of invex functions determined by (1.1) and differentiable convex functions seem to be new.

We recall, Hanson (1981), that a differentiable function $f : R^n \to R$ is

a) invex on R^n, if there exists a vector function $\eta : R^n \times R^n \to R^n$ such that for any $x, \bar{x} \in R^n$ we have
$f(x) - f(\bar{x}) \geq \langle \nabla f(\bar{x}), \eta(x, \bar{x}) \rangle$,

b) quasi-invex on R^n, if there exists a vector function $\eta : R^n \times R^n \to R^n$ such that for any $x, \bar{x} \in X$ we have

We will call a function f incave or quasi-incave if $-f$ is invex or quasi-invex. The fundamental theorem says, Ben-Israel, Mond (1986), that a function $f : R^n \to R$ is invex, whenever all its stationary points are global minimizers.

Pini (1994) proved that a real differentiable function defined on a manifold M is invex iff a certain function of a real variable is convex. Hanson, Mond (1987) consider the nonlinear programming problems which are transformable into convex problems, i.e. for which there exists such a regular C^1 diffeomorphism φ that the functions involved become actually convex when composed with φ. Such functions, which are necessarily invex, are called convexifiable. It is shown that not all invex functions possess this property. Craven (1995) investigates further the connections between invex functions with respect to a certain vector function η and convex functions. The subclass of invex functions that are convexifiable is determined and some Lagrangian duality connections with invexity are investigated. Let a differentiable function $f : R^n \to R$ be given which is invex at $\overline{x} \in R^n$ with respect to $\eta : R^n \to R^n$, called scale function. Put $\omega(x) = \eta(x - \overline{x}, \overline{x})$ and assume that $\omega(0) = 0$ and $\nabla\omega(0) = 1$, where 1 denotes the identity mapping. Function $q = f \circ \omega^{-1}$ appears to be convex at 0.

Our results are connected to those of Craven (1995), although we make use of the diffeomorphism that makes a given problem convex transformable while in Craven (1995) the scale functions is used to obtain a formula for a convex function related to given invex one. Our results seems to be easier to apply, compare Example 2.2.

The invex problem that we shall study reads: Find an $\overline{x} \in S_I$, if it exists, such that

$$f_0(\overline{x}) = \min_{x \in S_I} f_0(x),$$

where set S_I is defined as below

$$S_I = \{x \mid x \in R^n,\ f_i(x) \leq 0,\ i = 1, 2, \ldots, m,\ h_j(x) = 0,\ j = 1, 2, \ldots, k\},$$

and is called a primal feasible set; numerical functions f_i, $i = 0, 1, \ldots, m$, are differentiable and functions h_j, $j = 1, 2, \ldots, k$ are continuously differentiable on R^n. Throughout the paper, if not stated otherwise, we shall assume that

A There exists a continuously differentiable function $\varphi : R^n \xrightarrow{1\text{-}1} R^n$, such that a matrix $(\nabla_x \varphi(x))^{-1}$ exists for any $x \in R^n$ and that a function f_0 is invex, functions f_i, $i = 1, \ldots, m$ are quasi-invex and functions h_j, $j = 1, 2, \ldots, k$ are both quasi-invex and quasi-incave on R^n with respect to the same function $\eta : R^n \times R^n \to R^n$ given by the formula

$$\eta(x, \bar{x}) = (\nabla_x \varphi(\bar{x}))^{-1}(\varphi(x) - \varphi(\bar{x})). \tag{1.1}$$

In solving the Problem (PI) the Karush-Kuhn-Tucker necessary conditions, K-K-T conditions for short, are used. Hence, in order to solve the Problem (PI) one may apply the following procedure: find all K-K-T points of the Probhas to make sure that certain constraint qualification holds. There are many constraint qualifications when the function involved are convex or standard generalizations of convex function (quasi- or pseudo-convex), see Bazaara, Sherali, Shetty (1991), Mangasarian (1969). But not all constraint qualifications may be applied when the problem is invex, see Hanson, Mond (1984), Ben-Israel, Mond (1986), for certain constraint qualifications in that case. The invex problems are also difficult to solve, see for example the problem in Hanson (1981). Thus, we will present a way of finding the K-K-T points of the above invex problem using certain convex problem, for which there are not only many constraint qualifications but also there are methods for determining either explicitly or numerically K-K-T points, Bazaara, Sherali, Shetty (1991).

2. Solution of (PI)

There exists a convex problem closely related to a given invex problem. Indeed we have the following theorem:

THEOREM 2.1 *There exist a convex function g_0, quasi-convex functions g_i, $i = 1, 2, \ldots, m$, both quasi-convex and quasi-concave functions p_j, $j = 1, 2, \ldots, k$ such that*

$$g_i(\varphi(x)) = f_i(x) \text{ for any } x \in R^n, \tag{2.1}$$

$$p_j(\varphi(x)) = h_j(x) \text{ for any } x \in R^n, \tag{2.2}$$

Proof. We prove only that a function g_0 given by the formula $g_0 = f_0 \circ \varphi^{-1}$ is convex and satisfies (2.1). Let us take arbitrary $\xi, \bar{\xi} \in R^n$. By invexity of f_0 we than have the following inequality satisfied

$$f_0(\varphi^{-1}(\xi)) - f_0(\varphi^{-1}(\bar{\xi})) \geq \langle \nabla_x f_0(\varphi^{-1}(\bar{\xi})), \eta(\varphi^{-1}(\xi), \varphi^{-1}(\bar{\xi}))\rangle.$$

Using definition (1.1) we obtain that

$$\begin{aligned} & f_0(\varphi^{-1}(\xi)) - f_0(\varphi^{-1}(\bar{\xi})) \\ & \geq \langle ((\nabla_x \varphi(\varphi^{-1}(\bar{\xi})))^{-1})^T \nabla_x f_0(\varphi^{-1}(\bar{\xi})), \varphi(\varphi^{-1}(\xi)) - \varphi(\varphi^{-1}(\bar{\xi}))\rangle, \end{aligned}$$

which gives in turn that

$$g_0(\xi) - g_0(\bar{\xi}) \geq \langle \nabla_\xi g_0(\bar{\xi}), \xi - \bar{\xi}\rangle, \tag{2.3}$$

since $\nabla_\xi g_0(\bar{\xi}) = ((\nabla_x \varphi(\varphi^{-1}(\bar{\xi})))^{-1})^T \nabla_x f_0(\varphi^{-1}(\bar{\xi}))$. Since the inequality (2.3) holds for every $\xi, \bar{\xi} \in R^n$ the convexity of g_0 is proven.

The same argument leads to the conclusion that functions g_i given by the formulas $g_i = f_i \circ \varphi^{-1}$ are quasi-invex for $i = 1, 2, \ldots, m$ and functions $p_j = h_j \circ \varphi^{-1}$ are both quasi-invex and quasi-incave for $j = 1, 2, \ldots, k$. ■

EXAMPLE 2.2 *Consider the function $f : R \to R$ given by the formula $f(x) = x^3 + x$. It is of course invex with respect to $\widetilde{\eta}(x, \overline{x}) = \frac{x^3+x-\overline{x}^3-\overline{x}}{3\overline{x}^2+1}$ and a diffeomorphism φ that makes g convex reads $\varphi(x) = x^3 + x$. Thus, $g(\xi) = \xi$. If we try to determine the convex function as considered in Craven (1995) we have to solve with respect to x the equation $\frac{x^3+x-\overline{x}^3-\overline{x}}{3\overline{x}^2+1} = \xi$.*

In consequence to a given invex problem (PI) we associate a convex problem in which the functions involved are determined by the above theorem and which is called the Problem (PC):

Find an $\overline{\xi} \in S_C$, if it exists, such that

$$g_0(\overline{\xi}) = \min_{\xi \in S_C} g_0(\xi),$$

where the set S_C is defined as below

$$S_C = \{\xi \mid \xi \in R^n,\ g_i(\xi) \le 0,\ i = 1, 2, \ldots, m,\ p_j(\xi) = 0,\ j = 1, 2, \ldots, k\}.$$

The following theorem provides the announced relation between the K-K-T points of the Problems (PC) and (PI) respectively.

THEOREM 2.3 *Let $\overline{\xi} \in S_C$ be a K-K-T point for the Problem (PC). Then there exists a point $\overline{x} \in S_I$, which is a K-K-T point for the Problem (PI) and conversely, if a point $\overline{x} \in S_I$ is a K-K-T point for the Problem (PI), then there exists a point $\overline{\xi} \in S_C$ which is a K-K-T point for the Problem (PC) with the same vectors of Lagrange multipliers.*

Proof. Let $\overline{\xi}$ satisfy the Karush-Kuhn-Tucker conditions for the Problem (PI), i.e. there exist a vector $\overline{\lambda} \in R^m$, a vector $\overline{\mu} \in R^k$ such that, see Bazaara, Sherali, Shetty (1991),

$$\nabla_\xi g_0(\overline{\xi}) + \sum_{i=1}^{m} \overline{\lambda}_i \nabla_\xi g_i(\overline{\xi}) + \sum_{j=1}^{k} \overline{\mu}_j \nabla_\xi p_j(\overline{\xi}) = 0, \tag{2.4}$$

$$\sum_{i=1}^{m} \overline{\lambda}_i g_i(\overline{\xi}) = 0, \tag{2.5}$$

$$\overline{\lambda}_i \ge 0, \text{ for } i = 1, 2, \ldots, m. \tag{2.6}$$

By the assumption on the function φ we obtain the existence of such an $\overline{x} \in S_I$ that $\overline{\xi} = \varphi(\overline{x})$. From (2.1) and (2.2) we obtain for $i = 0, 1, \ldots, m$ and $j = 1, \ldots, k$ that

$$\nabla_x f_i(\overline{x}) = (\nabla_x \varphi(\overline{x}))^T \nabla_\xi g_i(\varphi(\overline{x})), \tag{2.7}$$

Now let us replace in (2.5) $g_i(\overline{\xi})$ by $f_i(\overline{x})$ and multiply (2.4) by $\nabla_x\varphi(\overline{x})$. Taking into account (2.1), (2.2) and (2.7), (2.8) we come to the following relations

$$\nabla_x f_0(\overline{x}) + \sum_{i=1}^{m} \overline{\lambda}_i \nabla_x f_i(\overline{x}) + \sum_{j=1}^{k} \overline{\mu}_j \nabla_x h_j(\overline{x}) = 0, \tag{2.9}$$

$$\sum_{i=1}^{m} \overline{\lambda}_i f_i(\overline{x}) = 0, \tag{2.10}$$

$$\overline{\lambda}_i \geq 0, \text{ for } i = 1, 2, \ldots, m. \tag{2.11}$$

Hence, $\overline{x}$ is a K-K-T point for (PI) with $(\overline{\lambda}, \overline{\mu})$ being vector of Lagrange multipliers.

Conversely, let us suppose that conditions (2.9)-(2.11) are satisfied for some point $\overline{x} \in S_I$ and vectors $\overline{\lambda} \in R^m, \overline{\mu} \in R^k$. Put $\overline{\xi} = \varphi(\overline{x})$. Again $\overline{\xi}$ is feasible for the Problem (PC) and taking into account (2.7), (2.8) we obtain consequently from (2.9)

$$(\nabla_x\varphi(\overline{x}))^T \Big(\nabla_\xi g_0(\bar{\xi}) + \sum_{i=1}^{m} \overline{\lambda}_i \nabla_\xi g_i(\bar{\xi}) + \sum_{j=1}^{k} \overline{\mu}_j \nabla_\xi p_j(\bar{\xi})\Big) = 0,$$

$$\nabla_\xi g_0(\bar{\xi}) + \sum_{i=1}^{m} \overline{\lambda}_i \nabla_\xi g_i(\bar{\xi}) + \sum_{j=1}^{k} \overline{\mu}_j \nabla_\xi p_j(\bar{\xi}) = 0,$$

where the last relation holds because $\nabla_x\varphi(\overline{x})$ is invertible. Thus, condition (2.4) is satisfied. Conditions (2.5)-(2.6) are satisfied because of (2.1) and conditions (2.10)-(2.11). ∎

COROLLARY 2.4 *The Problem (PC) has a solution if and only if the Problem (PI) has a solution.*

Proof. The proof bases on the previous theorem and on the fact that for both an invex, Hanson (1981) and a convex, Bazaara, Sherali, Shetty(1991) problems all the K-K-T points are global minimizers. ∎

Now we provide two examples: the first of an invex problem solved via a convex one and the second of a convex problem solved via an easier convex problem.

EXAMPLE 2.5 *Consider the Problem (PI)*

$$f_0(x_1, x_2) = x_1^2 + 2x_1^4 + x_1^6 + x_2^2 + 2x_2^4 + x_2^6 \to \inf,$$
$$f_1(x_1, x_2) = x_1^2 + 2x_1^4 + x_1^6 + x_2^2 + 2x_2^4 + x_2^6 - 5 \leq 0,$$
$$f_2(x_1, x_2) = -x_1 - x_1^3 \leq 0,$$
$$f_3(x_1, x_2) = x_2 + x_2^3 \leq 0,$$

Here the functions f_0 and f_1 are convex, the functions f_2, f_3 are quasi-convex on the feasible set, but the function h_1 is not quasi-convex, which can be seen by taking $x = (1,0)$ and $u = (0,-1)$. Both points are feasible to the problem and $h_1(x) - h_1(u) = 0$ while $(x-u)^T \nabla h_1(u) > 0$. Let us put

$$\varphi(x_1, x_2) = (x_1 + x_1^3, x_2 + x_2^3).$$

Then, using Theorem 2.1, we solve the following Problem (PC)

$$\begin{aligned}
&g_0(\xi_1,\xi_2) = \xi_1^2 + \xi_2^2 \to \inf,\\
&g_1(\xi_1,\xi_2) = \xi_1^2 + \xi_2^2 - 5 \le 0,\\
&g_2(\xi_1,\xi_2) = -\xi_1 \le 0,\\
&g_3(\xi_1,\xi_2) = \xi_2 \le 0,\\
&h_1(\xi_1,\xi_2) = -\xi_1 + \xi_2 + 2 = 0,
\end{aligned}$$

associated with a given invex problem. The functions considered satisfy a constraint qualification at the point $(1,-1)$ since gradients of active functions at that point are linearly independent and they are continuously differentiable. This point being a K-K-T point is the solution to the Problem (PC). Hence, by Theorem 2.3 a point $(\overline{x}_1, \overline{x}_2)$ such that $\overline{x}_1 = \frac{1}{6}\sqrt[3]{(108+12\sqrt{93})} - \frac{2}{\sqrt[3]{(108+12\sqrt{93})}}$, $\overline{x}_2 = -\frac{1}{6}\sqrt[3]{(108+12\sqrt{93})} + \frac{2}{\sqrt[3]{(108+12\sqrt{93})}}$ is a K-K-T point for the Problem (PI) and by Corollary 2.4 its global solution.

EXAMPLE 2.6 *It appears that the procedure described above has applications also when the Problem (PI) is actually convex. In that case we can find an easier way to solve the convex problem. Indeed, consider the Problem (PI)*

$$\begin{aligned}
&f_0(x_1,x_2) = x_1^2 + 2x_1^4 - 6x_1 + x_1^6 - 6x_1^3\\
&\quad + x_2^2 + 2x_2^4 - 4x_2 + x_2^6 - 4x_2^3 + 13 \to \inf,\\
&f_1(x_1,x_2) = x_1^2 + 2x_1^4 + x_1^6 + x_2^2 + 2x_2^4 + x_2^6 - 5 \le 0,\\
&f_2(x_1,x_2) = -x_1 - x_1^3 \le 0,\\
&f_3(x_1,x_2) = -x_2 - x_2^3 \le 0,
\end{aligned}$$

which is actually convex on the feasible set. Proceeding as above with the same functions η and $\widetilde{\eta}$ we find the Problem (PC) which reads

$$\begin{aligned}
&g_0(\xi_1,\xi_2) = \xi_1^2 - \xi_1 + \xi_2^2 - \xi_2 + 13 \to \inf,\\
&g_1(\xi_1,\xi_2) = \xi_1^2 + \xi_2^2 - 5 \le 0,\\
&g_2(\xi_1,\xi_2) = -\xi_1 \le 0,\\
&g_3(\xi_1,\xi_2) = -\xi_2 \le 0.
\end{aligned}$$

Here the K-K-T point is $(0.5, 0.5)$. *We then compute that the solution to (PI) is* $(0.42385, 0.42385)$.

3. The nondifferentiable case

In this section we shall not assume that the functions involved are differentiable. We still deal with the problem (PI) which now reads: Find an $\overline{x} \in S_I$, if it exists, such that

$$f_0(\overline{x}) = \min_{x \in S_I} f_0(x),$$

where the set S_I is defined as below

$$S_I = \{x \mid x \in R^n,\ f_i(x) \leq 0,\ i = 1, 2, \ldots, m\}.$$

and is called a primal feasible set, where numerical functions f_i, $i = 0, 1, \ldots, m$, are defined and locally Lipschitz on R^n. We shall assume that

A There exists a continuously differentiable function $\varphi : R^n \xrightarrow{1\text{-}1} R^n$, such that a matrix $(\nabla_x \varphi(x))^{-1}$ exists for any $x \in R^n$ and a function f_0 is invex on R^n, functions f_i $i = 1, 2, \ldots, m$ are quasi-invex on R^n with respect to the same function $\eta : R^n \times R^n \to R^n$ given by the formula
$\eta(x, \bar{x}) = (\nabla_x \varphi(\bar{x}))^{-1}(\varphi(x) - \varphi(\bar{x}))$.

Here the definition of η is the same as in the differentiable case.

To this end $\partial f(\bar{x})$ denotes the Clarke subdifferential of a function f at a point $\bar{x} \in R^n$ and $f^0(\bar{x}, d)$ a generalized directional derivative at a point $\bar{x} \in X$ in the direction $d \in R^n$, see Clarke (1984). We recall, Reiland (1991), that a locally Lipschitz function $f : R^n \to R$ is,

a) invex on R^n, if there exists a vector function $\widetilde{\eta} : R^n \times R^n \to R^n$ such that for any $x, \bar{x} \in R^n$ we have
$f(x) - f(\bar{x}) \geq f^0(\bar{x}, \widetilde{\eta}(x, \bar{x}))$,
or equivalently
$f(x) - f(\bar{x}) \geq \langle \zeta, \widetilde{\eta}(x, \bar{x}) \rangle$ for every $\zeta \in \partial f(\bar{x})$,

b) quasi-invex on R^n, if there exists a vector function $\widetilde{\eta} : R^n \times R^n \to R^n$ such that for any $x, \bar{x} \in R^n$ we have
$f(x) - f(\bar{x}) \leq 0 \Longrightarrow f^0(\bar{x}, \widetilde{\varphi}(x, \bar{x})) \leq 0$,
or equivalently
$f(x) - f(\bar{x}) \leq 0 \Longrightarrow \langle \zeta, \widetilde{\varphi}(x, \bar{x}) \rangle \leq 0$ for every $\zeta \in \partial f(\bar{x})$.

The fundamental theorem again says, Reiland (1991), that a function $f : R^n \to R$ is invex if all its stationary points, i.e. points $x \in R^n$ satisfying the relation $0 \in \partial f(x)$, are global minimizers.

Again with the given invex problem we can associate a certain convex problem. Indeed, we have the following

THEOREM 3.1 *There exists a convex function g_0 and quasi-convex functions g_i for $i = 1, 2, \ldots, m$ such that*

$$g_i(\varphi(x)) = f_i(x) \text{ for any } x \in R^n, \text{ for } i = 0, 1, 2, \ldots, m. \tag{3.1}$$

Proof. We prove only that a function g_0 given by $g_0 = f_0 \circ \varphi^{-1}$ is convex and satisfies (3.1). Let $\xi, \bar{\xi} \in R^n$. By invexity of f_0, using definition (1.1) we obtain that

$$f_0(\varphi^{-1}(\xi)) - f_0(\varphi^{-1}(\bar{\xi})) \geq \zeta^T(\nabla_x \varphi(\varphi^{-1}(\bar{\xi})))^{-1}(\varphi(\varphi^{-1}(\xi)) - \varphi(\varphi^{-1}(\bar{\xi})))$$

which, in view of $(\nabla_x \varphi(\varphi^{-1}(\bar{\xi})))^{-1} = \nabla_\xi \varphi^{-1}(\bar{\xi})$ yields

$$g_0(\xi) - g_0(\bar{\xi}) \geq ((\nabla_\xi \varphi^{-1}(\bar{\xi}))^T \zeta)^T (\xi - \bar{\xi}).$$

Now $\partial_\xi g_0(\bar{\xi}) = (\nabla_\xi \varphi^{-1}(\bar{\xi}))^T \partial_x f_0(\varphi^{-1}(\bar{\xi}))$ by the chain formula, Rockafellar, Wets (1998). So

$$g_0(\xi) - g_0(\bar{\xi}) \geq \varsigma^T(\xi - \bar{\xi}), \text{ for every } \varsigma \in \partial_\xi g_0(\bar{\xi}).$$

This proves the convexity of g_0.

The same argument leads to the conclusion that functions g_i, $i = 1, 2, \ldots, m$, given by the formulas $g_i = f_i \circ \varphi^{-1}$ are quasi-invex. ■

To a given invex problem (PI) we associate a convex problem (PC) which reads: Find an $\overline{\xi} \in S_C$, if it exists, such that

$$g_0(\overline{\xi}) = \min_{\xi \in S_C} g_0(\xi),$$

where the set S_C is defined as below

$$S_C = \{\xi \mid \xi \in R^n, \ g_i(\xi) \leq 0, \ i = 1, 2, \ldots, m\}.$$

Following the results obtained in Section 2 we have

THEOREM 3.2 *Let $\overline{\xi} \in S_C$ be a K-K-T point for the Problem (PC). Then there exists a point $\overline{x} \in S_I$ which is a K-K-T point for the Problem (PI) and conversely, if a point $\overline{x} \in S_I$ is a K-K-T point for the Problem (PI), then there exists a point $\overline{\xi} \in S_C$ which is a K-K-T point for the Problem (PC) with the same vector of Lagrange multipliers.*

Proof. Let $\overline{\xi}$ satisfy the Karush-Kuhn-Tucker conditions for the Problem (PI), i.e. there exists a vector $\overline{\lambda} \in R^m$, such that, Hiriart-Urruty (1978),

$$0 \in \left(\partial g_0(\overline{\xi}) + \sum_{i=1}^{m} \overline{\lambda}_i \partial g_i(\overline{\xi})\right), \tag{3.2}$$

$$\sum_{i=1}^{m} \overline{\lambda}_i g_i(\overline{\xi}) = 0, \tag{3.3}$$

$$\overline{\lambda}_i \geq 0, \text{ for } i = 1, 2, \ldots, m. \tag{3.4}$$

Put $\overline{x} = \varphi^{-1}(\overline{\xi}) \in S_I$. From (3.2) it follows that there exist ς_i, $i = 0, 1, \ldots m$, such that $\varsigma_i \in \partial g_i(\overline{\xi})$ and the following relation holds

$$\varsigma_0 + \overline{\lambda}_m \varsigma_1 + \ldots + \overline{\lambda}_m \varsigma_m = 0. \tag{3.5}$$

From (3.1) we obtain by chain rule differentiation, Rockafellar, Wets (1998), that

$$(\nabla_x \varphi(\overline{x}))^T \varsigma_i = \tau_i, \tag{3.6}$$

where and $\tau_i \in \partial f_i(\overline{x})$ for $i = 0, 1, \ldots, m$. Thus, multiplying (3.5) by $(\nabla_x \varphi(\overline{x}))^T$ and taking into account (3.1), (3.6) we come to the following relations

$$0 \in \Big(\partial f_0(\overline{x}) + \sum_{i=1}^{m} \overline{\lambda}_i \partial f_i(\overline{x})\Big), \tag{3.7}$$

$$\sum_{i=1}^{m} \overline{\lambda}_i f_i(\overline{x}) = 0, \tag{3.8}$$

$$\overline{\lambda}_i \geq 0, \text{ for } i = 1, 2, \ldots, m, \tag{3.9}$$

from which we obtain that $\overline{x}$ is a K-K-T point for (PI).

Conversely, let us suppose that conditions (3.7)-(3.9) are satisfied for some point $\overline{x} \in S_I$ and a vector $\overline{\lambda} \in R^m$. Put $\overline{\xi} = \varphi(\overline{x}) \in S_C$. By (3.7), we obtain the existence of τ_i, $i = 0, 1, \ldots, m$, such that $\tau_i \in \partial f_i(\overline{x})$, satisfying the following relation

$$\tau_0 + \overline{\lambda}_1 \tau_1 + \overline{\lambda}_m \tau_m = 0. \tag{3.10}$$

Since (3.10) holds we obtain (3.5) by (3.6), that gives (3.2). Conditions (3.3) and (3.4) follow from (3.8) and (3.9) if we take into account the assumptions on the functions f_i and g_i, $i = 1, 2, \ldots, m$. ■

COROLLARY 3.3 *The Problem (PC) has a solution if and only if the Problem (PI) has a solution.*

Proof. The proof bases on Theorem 3.2 and the fact that for both an invex, Reiland (1991), and a convex problems all the K-K-T points are global minimizers. ■

REMARK 1 *It is now easy to derive relationships between Mond-Weir duals to both invex and convex problems in the differentiable and nonsmooth cases.*

4. Fenchel-Young duality for an invex function

In this section we follow Rockafellar, Wets (1998), Section 11.L* entitled "Generalized conjugacy" and provide a kind of generalized conjugacy operating on invex functions. The general framework for such generalized duality is to be found in the book mentioned above but the results given in the present paper are not contained there. This concept appears to be of use for example when a variational method is applied to show the existence of a solution to a certain differential inclusion and the action functional happens to lack convexity property, does not satisfy suitable growth conditions and its second Fenchel-Young dual is equal to $-\infty$, see Nowakowski (1992), Orpel (1997), and references therein. We shall require the following assumption to hold:

A Let $f : R^n \to R$ be locally Lipschitz. Let there exist a continuously differentiable function $\varphi : R^n \xrightarrow{1\text{-}1} R^n$ such that the matrix $\nabla_x \varphi(x)$ is invertible for any $x \in R^n$ and that the function f is invex on R^n with respect to the function $\eta : R^n \times R^n \to R^n$ given by the formula (1.1), namely

$$\eta(x, \bar{x}) = (\nabla_x \varphi(\bar{x}))^{-1}(\varphi(x) - \varphi(\bar{x}))$$

If f is differentiable then it suffices to assume that φ is differentiable. By Theorem 3.1 or 2.1 in the differentiable case, we obtain that there exists a function $g : R^n \to R$ such that g is convex and $g(\varphi(x)) = f(x)$ for all $x \in R^n$.

For a convex function g the Fenchel-Young dual g^* and the second dual g^{**} are defined as follows, see Rockafellar, Wets (1998),

$$g^*(\xi^*) = \sup_{\xi \in R^n} \{\langle \xi^*, \xi \rangle - g(\xi)\} \tag{4.1}$$

$$g^{**}(\xi) = \sup_{\xi^* \in R^n} \{\langle \xi, \xi^* \rangle - g^*(\xi^*)\} \tag{4.2}$$

We define $f^\varphi(x^*)$, where $x^* \in R^n$ is arbitrary, to be

$$f^\varphi(x^*) = \sup_{x \in R^n} \{\langle \varphi(x^*), \varphi(x) \rangle - f(x))\} \tag{4.3}$$

THEOREM 4.1 *Let f satisfy assumption A and let $g = f \circ \varphi^{-1}$. We obtain that*

$$f^\varphi = g^* \circ \varphi.$$

Proof. Since the function $R^n \ni x \longmapsto \langle \varphi(x^*), \varphi(x) \rangle - f(x) \in R$ is incave as a sum of incave functions with respect to the function $\widetilde{\eta}(x, \bar{x}) = (\nabla_x \varphi(\bar{x}))^{-1}(\varphi(x) - \varphi(\bar{x}))$ we have by Theorem 3.2 or 2.3 in differentiable case, if convexity is replaced by concavity, that

$$f^\varphi(x^*) = \sup_{\xi \in R^n} \{\langle \xi^*, \xi \rangle - g(\xi)\},$$

where we put $\varphi(x^*) = \xi^*$. This is the definition of the Fenchel-Young conjugate of the convex function g. ■

We define the second dual $f^{\varphi\varphi}$ of a function f as follows

$$f^{\varphi\varphi}(x) = \sup_{x^* \in R^n} \{\langle \varphi(x^*), \varphi(x) \rangle - f^\varphi(x^*)\} \tag{4.4}$$

We have the following theorem

THEOREM 4.2 *Let f satisfy assumptions A and let $g = f \circ \varphi^{-1}$. Let $f^{\varphi\varphi}$ and f^φ be calculated as in (4.3) and (4.4), respectively. Then*

Proof. By (4.4) for any $x \in R^n$ we have $f^{\varphi\varphi}(x) = \sup_{x^* \in R^n}\{\langle\varphi(x^*), \varphi(x)\rangle - f^{\varphi}(x^*)\} = \sup_{\xi^* \in R^n}\{\langle\xi^*, \xi\rangle - g^*(\xi^*)\}$. The last relation holds by Corollary 3.3 or 2.4 in differentiable case, because the function $g^* \circ \varphi$ is invex with respect to η given by (1.1). By the theorem on the Fenchel-Young transform for a convex function, we obtain consequently $f^{\varphi\varphi}(x) = g^{**}(\xi) = g(\xi)$. Finally by the assumptions on f and g we have for any $x \in R^n$, $f^{\varphi\varphi}(x) = f(x)$. ■

The dual and second dual introduced above have the following properties:

PROPOSITION 4.3 *Let f, f_1, f_2 satisfy assumption A and let $g = f \circ \varphi^{-1}$, $g_1 = f_1 \circ \varphi^{-1}$, $g_2 = f_2 \circ \varphi^{-1}$. Then we have*

1. *(the Fenchel-Young inequality)*
   $f^{\varphi}(x^*) + f(x) \geq \langle\varphi(x^*), \varphi(x)\rangle$
2. *if $f_1 \leq f_2$ then $f_1^{\varphi} \geq f_2^{\varphi}$.*

Proof. Put $\xi = \varphi(x)$. Since $g(\varphi(x)) = f(x)$ and $f^{\varphi}(x^*) = g^*(\varphi(x^*))$ we obtain

$$f^{\varphi}(x^*) + f(x) = g^*(\varphi(x^*)) + g(\varphi(x)) \geq \langle\varphi(x^*), \varphi(x)\rangle$$

Let now $f_1 \leq f_2$.Then $g_i(\varphi_i(x)) = f_i(x)$, $i = 1, 2$, and of course $g_1 \leq g_2$. Thus, by the properties of the Fenchel-Young dual we obtain that $g_1^*(\xi^*) \geq g_2^*(\xi^*)$, which by definition of f_1^{φ} and f_2^{φ} means that $f_1^{\varphi}(x^*) \geq f_2^{\varphi}(x^*)$, where $\xi^* = \varphi(x^*)$. ■

To conclude the section we provide easy examples to back our theory up.

EXAMPLE 4.4 *Let us consider the function $f(x) = x^3 + x$ from Example 2.2. $f^{**} = -\infty$ if it is calculated by (4.1). Here $\varphi(x) = x + x^3$ and a convex function g is linear. Hence we obtain:*

$$f^{\varphi}(x^*) = \sup_{x \in R}\{\varphi(x^*)\varphi(x) - f(x)\}$$
$$= \sup_{\xi \in R}\{\varphi(x^*)\xi - \xi\} = \begin{cases} 0 & \text{if } \varphi(x^*) = 1, \\ \infty & \text{otherwise.} \end{cases}$$

By (4.4) we obtain

$$f^{\varphi\varphi}(x) = \sup_{\xi \in R^n}\{\varphi(x^*)\varphi(x) - f^{\varphi}(x^*)\} = \varphi(x) = f(x).$$

REMARK 2 *For a convex function f the two concepts of duality give the same second dual. If one takes $\varphi(x) = x$ then first duals are also equal, take for*

References

BEN-ISRAEL, A. anf MOND, B. (1986) What is Invexity. *J. Austral. Math. Soc. Ser. B*, **28**, 1–9.

CLARKE, F.H. (1984) *Nonsmooth Optimization.* Nauka, Moscow (in Russian).

CRAVEN, B.D. (1995) Relationships Between Invex Properties. In: *Recent Trends in Optimization Theory*, ed. Agrawal, R.P., World Scientific, Singapore.

HANSON, M.A. (1981) On the Sufficiency of the Kuhn-Tucker Conditions. *J. Math. Anal. Appl.*, **80**, 545–550.

HANSON, M.A. and MOND, B. (1984) Necessary and Sufficient Conditions in Constraint Optimization. In: *FSU Statistic Report M 683*, Florida State University, Dept. of Statistics, Tallahassee, Florida, U.S.A., 51–58.

HANSON, M.A. and MOND, B. (1987) Convex Transformable Programming Problems and Invexity. *J. Inf. Optim. Sci.*, **8**, 201–207.

HIRIART-URRUTY, J.B. (1978) On Optimality Conditions in Nondifferentiable Programming. *Math. Prog.*, **14**, 73–86.

MANGASARIAN, O.L. (1969) *Nonlinear Programming.* McGraw-Hill, New York.

MARTIN, D.H. (1985) The Essence of Invexity. *J. Optim. Theory Appl.*, **47**, 65–76.

MOND, B. and WEIR, T. (1986) Concave Functions and Duality. In: Schaible, T., *Concavity and Economics.*

NOWAKOWSKI, A. (1992) A New Variational Principle and Duality for Periodic Solutions of Hamilton's Equations. *J. Differential Equations*, **97**, 1, 174–188.

ORPEL, A. (1997) Periodic Solutions of Differential Equations and Nonconvex Functionals in Calculus of Variations. *Optimization*, **40**, 313–333.

PINI, R. (1991) Invexity and Generalized Convexity. *Optimization*, **22**, 513–525.

PINI, R. (1994) Convexity Along Curves and Invexity. *Optimization*, **29**, 301–309.

ROCKAFELLAR, T.R. and WETS, R.J.-B. (1998) *Variational Analysis.* Springer-Verlag.