Acta Math. Univ. Comenianae
Vol. LXV, 2(1996), pp. 229–245

DUALITY OF TRANSFORMATION FUNCTIONS IN THE INTERIOR POINT METHODS

M. HALICKÁ AND M. HAMALA

ABSTRACT. In this paper a duality of transformation functions in the interior point method is treated. A dual pair of convex or linear programming problems is considered and the primal problem is transformed by the parametrized transformation function of a more general form than logarithmic is. The construction of the parametrized transformation function for the dual problem is carried out so that both transformation functions were dual. The result obtained explains the unlucid construction of dual transformation functions so far known as a special case of a simple general principle of constructing dual transformation functions.

1. INTRODUCTION

In the framework of the interior point methods (IPM) the linear programming problem

$$\text{(LP)} \qquad \min\left\{c^T x \mid Ax = b,\ x \geq 0\right\}, \quad x, c \in \mathbb{R}^n,\ b \in \mathbb{R}^m,\ A \in \mathbb{R}^{m\times n}$$

is solved using logarithmic transformation function

$$T(x;\mu) = c^T x - \mu \sum_{i=1}^{n} \ln x_i \tag{1}$$

where $x \in \mathcal{P}^o = \{x \mid Ax = b,\ x > 0\}$ and $\mu > 0$ is a parameter.

The standard assumptions on (LP) in this context are: $\operatorname{rank}(A) = m \leq n$, $\mathcal{P}^o \neq \emptyset$ and $\mathcal{P}^* \neq \emptyset$, $\mathcal{P}^*$ bounded, $\mathcal{P}^*$ being the set of optimal solutions of (LP). These assumptions together with excellent properties of a logarithm in the function T guarantee the existence of the unique minimum x^μ of $T(\cdot;\mu)$ for any $\mu > 0$ as well as the convergence of x^μ to an optimal solution of (LP) for $\mu \to 0$. These theoretical results offer the possibility to solve the original problem (LP) by means of approximate minima of $T(\cdot;\mu)$ for decreasing values of μ.

Received May 14, 1996.
1980 *Mathematics Subject Classification* (1991 *Revision*). Primary 90C25; Secondary 90C05, 90C30.
Key words and phrases. Linear programming, convex programming, interior point methods, transformation function, dual problem.

In the last ten years many algorithms have been developed, analyzed and implemented which more or less clearly follow the idea mentioned above. A most interesting property of these algorithms is that they, unlike the simplex ones, are polynomial. In order to prove the polynomiality it is necessary to estimate the number of iterations needed and to keep checking the quality of running approximations. For this purpose it is of advantage to exploit the information on dual variable. Namely, along with the primal problem (LP) we can also consider the associated dual problem

$$\text{(LD)} \qquad \max\left\{b^T y \mid A^T y + z = c,\ z \geq 0\right\}, \quad z \in \mathbb{R}^n$$

and the corresponding transformation function

$$Q(y, z; \mu) = b^T y + \mu \sum_{i=1}^{n} \ln z_i, \quad (y, z) \in \mathcal{D}^o \tag{2}$$

where $\mathcal{D}^o = \{(y, z) \mid A^T y + z = c,\ z > 0\}$. Then between T and Q there exists the following dual relationship (enabling us to call the function Q the dual function to T):
(i) there exists a one-to-one correspondence (given by an explicit rule) between the minimum x^μ of $T(\cdot; \mu)$ on $\mathcal{P}^o$ and the maximum (y^μ, z^μ) of $Q(\cdot, \cdot; \mu)$ on $\mathcal{D}^o$;
(ii) in the extremal solutions x^μ, (y^μ, z^μ) one has

$$T(x^\mu; \mu) - Q(y^\mu, z^\mu; \mu) = n\mu(1 - \ln \mu). \tag{3}$$

These duality properties and their possible consequences are more or less clearly used in most polynomiality proofs for algorithms with logarithmic transformation function.

Actually, the logarithmic transformation functions are not the only ones to be used in the IPM approach. The theory of IPM was originally developed as a method for solving convex programming problems in [**3**], where the transformation function has more general form

$$T(x; \mu) = c^T x - \mu \sum_{i=1}^{n} \psi(x_i), \tag{4}$$

ψ is a suitable function. One of the most important properties of the transformation functions is that their minima have to be from the interior of the feasible set. This condition can be fulfilled by well-known barrier functions ψ for which $\lim_{\xi \to 0^+} \psi(\xi) = -\infty$ or, by so called quasi-barrier ones (from [**4**]), for which $\lim_{\xi \to 0^+} \psi(\xi)$ is finite and $\lim_{\xi \to 0^+} \psi'(\xi) = \infty$ (ψ' is the derivative of ψ). Examples of barrier and quasi-barrier functions are $\psi(\xi) = \ln \xi$, $\psi(\xi) = \frac{1}{p}\xi^p$, $p < 0$ and $\psi(\xi) = \frac{1}{p}\xi^p$, $0 < p < 1$, resp.

In the current state of development of IPM attention is mostly given to the logarithmic-like approach which allows to state some polynomiality properties of proposed algorithms. However, there are exceptions, e.g. [**1**], [**2**]. An algorithm based on the transformation (4) where $\psi(\xi) = \frac{1}{p}\xi^p$, $p < 0$, is proposed in [**1**]. Here the iterations number estimate is of exponential type (turning to be polynomial in the limit logarithmic case $p \to 0$). In [**2**] some duality results for a more general convex programming case are given. In our LP notations these duality results show that transformation functions T, Q for (LP), (LD), resp., where

$$(5) \qquad T(x;\mu) = c^T x - \mu \sum_{i=1}^{n} \frac{1}{p}\, x_i^p, \quad p < 1, \quad p \neq 0, \quad x \in \mathcal{P}^o, \qquad \text{and}$$

$$(6) \qquad Q(y,z;\mu) = b^T y + \mu^{\frac{1}{1-p}} \sum_{i=1}^{n} \frac{p-1}{p}\, z_i^{\frac{p}{p-1}}, \quad (y,z) \in \mathcal{D}^o,$$

are dual (in the sense similar to the above logarithmic duality, but with a zero duality gap, i.e., with zero on the right side of equality (3)).

Notice that whereas the dual functions (1), (2) are both constructed by the same function $\psi(\xi) = \ln \xi$, the dual functions (5), (6) are constructed by two different functions ψ_1, ψ_2 one of which is barrier and the other is quasi-barrier. Moreover, whereas dual functions (1), (2) have the parameter in the first power, this is not the case of functions (5), (6), where (6) has μ in the power $\frac{1}{1-p}$.

How to explain these differences? Is it possible to account for the structure of the dual functions (1), (2) and (5), (6) by some universal scheme? And, having a transformation function T for (LP), how to construct a transformation function Q for (LD) so that T and Q were dual?

In this paper we give answers to these questions in terms of the Legendre transformation for the concave function with parameter. To obtain a general rule for constructing the dual transformation function it will be necessary to consider the transformation function T of a more general type than the function (4) is. The general properties of such kind of functions will be briefly summarized in Section 2 for the convex programming case. Section 3 has also an auxiliary character. It gives some properties of the Legendre transformation for convex functions, adapted to our needs. In Sections 4 and 5 we formulate and solve the main problem of our paper for the linear and convex programming case. The last section gives some consequences of the proven duality of transformation functions.

2. Parametric Interior Point Methods

Consider the convex programming problem

$$\text{(CP)} \qquad \min\{f(x) \mid g_i(x) \geq 0 \ (i = 1, \ldots, m)\}$$

where f, $-g$ are convex and C^1 on $\mathbb{R}^n$,

$$K^o = \{x \in \mathbb{R}^n \mid g_i(x) > 0 \ (i = 1, \dots, m)\} \neq \emptyset$$

and the set of optimal solutions K^* is nonempty and bounded.

In the general theory of the IPM (i.e., the barrier and quasi-barrier ones) the problem (CP) is transformed to the parametrized unconstrained problem of the type

$$(\mathrm{CP}_\mu) \qquad \min\{T(x;\mu) \mid x \in K^o\}$$

where

$$T(x;\mu) = f(x) - \sum_{i=1}^{m} \Lambda(g_i(x);\mu). \tag{7}$$

Here given any positive value of the parameter μ, the function $\Lambda(\cdot;\mu)$ is considered to be defined on $(0,\infty)$ only with the first derivative continuous.

To be able to build up this theory some additional assumptions to the function Λ have been given. From [**5**] it follows that it is possible to separate the assumptions sufficient to prove the existence of an optimal solution of (CP_μ) from those which (once the existence is guaranteed) ensure some kind of convergence. We recall these results in the next two propositions. Note that in this paper we will often deal with the following assumptions of a function $\varphi \in C^1(0,\infty)$:

$$\text{(A)} \qquad \lim_{\xi\to\infty} \varphi'(\xi) = 0 \qquad \text{(asymptotic property)}$$

$$\text{(B)} \qquad \lim_{\xi\to 0^+} \varphi'(\xi) = \infty \qquad \text{(barrier property)}$$

$$\text{(C)} \qquad \varphi' \ \text{ is decreasing} \qquad \text{(strict concavity).}$$

Proposition 1. *Let $\mu > 0$ be given. Assuming that $\Lambda(\cdot;\mu)$ has the properties (A), (B), (C), the problem (CP_μ) has an optimal solution x^μ.*

Remark 1. (a) Note that the assumption (B) is a generalization of properties of the barrier and quasi-barrier functions mentioned in Section 1.
(b) The asymptotic assumption (A) seems to be only a technical one used for the proof of Proposition 1 but we will see a close dual relationship between (A) and (B) later.
(c) Using (C) we see that $\Lambda(\cdot;\mu)$ is strictly concave . Combining (A), (B), (C) we have $\Lambda'(\xi;\mu) > 0$ for all $\xi > 0$, i.e. $\Lambda(\cdot;\mu)$ is increasing. These two properties imply that the function $T(\cdot;\mu)$ defined by (7) is convex (under our assumption, that f, $-g$ are convex).

For a proof of Proposition 1 see [**3**, Theorem 25] in the case of barrier functions and [**5**] in the case of quasi-barrier ones.

Proposition 2. *Suppose*

$$\forall\, \xi > 0 \;:\; \lim_{\mu \to 0^+} \Lambda(\xi\ ;\mu) = 0 \tag{8}$$

and let at least one of the following two assumptions hold:

$$\forall\, \xi > 0,\ \forall\, \mu > 0 \;:\; \Lambda(\xi\ ;\mu) = \mu\ \psi(\xi), \tag{9}$$

$$\forall \xi > 0,\ \forall \mu > 0 \;:\; \Lambda(\xi\ ;\mu) \leq 0. \tag{10}$$

Let $\{\mu_k\}_{k=1}^{\infty}$ *be a sequence satisfying* $\mu_k > 0$, $\lim_{k\to\infty} \mu_k = 0$ *and suppose for each* μ_k *there exists an optimal solution* x^k *of* (CP_{μ_k})*. Then we have:*

$$\lim_{k\to\infty} T(x^k;\mu_k) = f^*, \tag{11}$$

and

$$\lim_{k\to\infty} f(x^k) = f^* \tag{12}$$

where f^* *is the optimal objective value of (CP). Moreover, if* (9) *holds and* $\{\mu_k\}$ *is monotonic, then we have the monotonic convergence in* (12) *as well.*

For a proof see [**3**, Theorems 25, 27].

Remark 2. Note that the proof of the last proposition can be performed even in the case when (9) is replaced by:

$$\forall\, \xi > 0 \;:\; \Lambda(\xi\ ;.) \text{ is convex} \tag{9a}$$

3. Legendre Transformation of Concave Functions

In this section we recall some basic properties of the classical Legendre transformation for concave functions of one variable. We also give some properties of the Legendre transformation for some special concave functions used in this paper.

Definition 1. Let (a,b), (c,d) be open intervals in $\mathbb{R}$. Let $\psi \in \mathcal{C}^1(a,b)$ be such that its derivative ψ' is a strictly monotone function mapping (a,b) onto (c,d). Then the function $\psi_L : (c,d) \mapsto \mathbb{R}$ defined by

$$\psi_L(\eta) = \eta\ \xi - \psi(\xi), \tag{13}$$

where

$$\xi = (\psi')^{-1}(\eta), \tag{14}$$

is called the Legendre transformation for ψ.

It is well known that the Legendre transformation for concave functions has the following properties (see e.g. [**9**]).

Lemma 1. *Let $\psi \in \mathcal{C}^1(a,b)$ be strictly concave with ψ' mapping (a,b) onto (c,d). Then*

(i) $\forall\, \eta \in (c,d)\ :\ \psi_L(\eta) = \min\limits_{a<\xi<b}[\eta\xi - \psi(\xi)]$,
(ii) $\forall\, \xi \in (a,b),\ \eta \in (c,d)\ :\ \psi(\xi) + \psi_L(\eta) \le \xi\ \eta$,
(iii) ψ_L *is a strictly concave function on* (c,d),
(iv) $\psi_L \in \mathcal{C}^1(c,d)$ *and* $\psi'_L = (\psi')^{-1}$,
(v) $(\psi_L)_L = \psi$.

Note that (i) illustrates the close relationship between the Legendre transformations and the Fenchel-Rockafellar conjugate functions in our case.

Lemma 2. *Let $\psi \in \mathcal{C}^1(0,\infty)$ have properties (A), (B), (C). Then ψ_L exists, $\psi_L \in \mathcal{C}^1(0,\infty)$ and ψ_L has properties (A), (B), (C).*

Proof. From (A), (B), (C) it follows that ψ' maps $(0,\infty)$ onto $(0,\infty)$, so Definition 1 can be applied and $\psi_L\colon (0,\infty) \to \mathbb{R}$. By Lemma 1, ψ_L is strictly concave on $(0,\infty)$, so ψ'_L is strictly decreasing . This is (C) for ψ_L. From Lemma 1(iv) we have $\psi'_L[\psi'(\xi)] = \xi$; thus (A) for ψ implies (B) for ψ_L, and also (B) for ψ implies (A) for ψ_L. □

Two examples of functions with properties (A), (B), (C) and the corresponding Legendre transformations are :

$$\psi(\xi) = \ln \xi + c, \qquad \psi_L(\eta) = \ln \eta + (1-c) \tag{15}$$

and

$$\psi(\xi) = \frac{1}{p}\xi^p, \qquad \psi_L(\eta) = \frac{1}{q}\eta^q,\quad p<1,\ p\neq 0,\ \frac{1}{p}+\frac{1}{q}=1. \tag{16}$$

As a particular case of (15) we have $\psi(\xi) = \ln \xi + \frac{1}{2}$ which can be considered as "self-Legendre".

We turn now to parametrized functions $\Lambda(\xi;\mu)\colon (0,\infty) \mapsto \mathbb{R}$ where $\mu > 0$ is a parameter. Then by $\Lambda_L(\xi;\mu)$ we denote the Legendre transformation for $\Lambda(\xi;\mu)$ as a function of ξ for fixed $\mu > 0$.

Definition 2. The parametrized function $\Lambda(\xi;\mu) : (0,\infty) \mapsto \mathbb{R}$ (where $\mu > 0$ is parameter) is called quasilinear in μ if it can be represented in the form

$$\Lambda(\xi\ ;\mu) = a(\mu)\ \psi(\xi) + b(\mu) \tag{17}$$

where $a(\mu) > 0$, $b(\mu)$ are some functions of the parameter $\mu > 0$.

Lemma 3. *Let $\Lambda(\xi;\mu) \in \mathcal{C}^1(0,\infty)$ be such that Λ has property (C) for any fixed $\mu > 0$. Then*

(i) *If $\Lambda(\xi;\mu) = a(\mu)\psi(\xi) + b(\mu)$, $a(\mu) > 0$, then*

$$\Lambda_L(\eta;\mu) = a(\mu)\ \psi_L\left(\frac{\eta}{a(\mu)}\right) - b(\mu). \tag{18}$$

(ii) *If $\forall\, \xi > 0$: $\Lambda(\xi;\cdot)$ is convex (in μ), then $\forall\, \eta > 0$: $\Lambda_L(\eta;\cdot)$ is concave in μ.*

Lemma 4. *Let $\Lambda(\cdot;\mu) \in \mathcal{C}^1(0,\infty)$ have properties (A), (B), (C) for any fixed $\mu > 0$. Then*

(i) *If $\forall\, \xi > 0$: $\lim_{\mu\downarrow 0} \Lambda(\xi\ ;\mu) = 0$, then $\forall\, \eta$: $\lim_{\mu\downarrow 0} \Lambda_L(\eta\ ;\mu) = 0$ as well.*

(ii) *If $\forall\, \xi > 0$, $\forall\, \eta > 0$: $\Lambda(\xi\ ;\mu) \leq 0$, then $\forall\, \eta$, $\forall\, \mu > 0$: $\Lambda_L(\eta\ ;\mu) \geq 0$.*

The proofs follow from Lemma 1(i).

Remarks.

(i) It is easy to see that the Legendre transformation of a quasilinear function need not be quasilinear.

(ii) If $\Lambda(\xi;\mu)$ is quasilinear of the form (17), where $\psi(\xi)$ is of the form (15) or (16), then $\Lambda_L(\eta;\mu)$ is also quasilinear, i.e.,

$$\text{if } \Lambda(\xi\ ;\mu) = \mu(\ln\xi + c), \quad \text{then } \Lambda_L(\eta\ ;\mu) = \mu(\ln\eta + (1+c) - \ln\mu),$$

$$\text{if } \Lambda(\xi\ ;\mu) = \mu\frac{1}{p}\xi^p,\ p<1,\ p\neq 0, \text{ then } \Lambda_L(\eta\ ;\mu) = \mu^{(1-q)}\frac{1}{q}\eta^q,\ \frac{1}{p}+\frac{1}{q}=1.$$

(iii) In fact, the functions from the previous remark are the only quasilinear in μ functions satisfying (A), (B), (C) for which the Legendre transformation is also quasilinear in μ. This is a consequence of properties of Pexider's functional equations: $f(xy) = g(x) + h(y)$, $f(xy) = g(x)\cdot h(y)$ [**8**].

4. Duality of Transformed Problems in Linear Programming

Consider the following pair of dual linear programming problems:

$$\min\left\{c^Tx \mid Ax = b,\ \ x \geq 0\right\}, \tag{LP}$$

$$\max\left\{b^Ty \mid A^Ty + z = c,\ \ z \geq 0\right\}, \tag{LD}$$

where $c,\ x,\ z \in \mathbb{R}^n$; $b, y \in \mathbb{R}^m$; $A \in \mathbb{R}^{m\times n}$; $\operatorname{rank} A = m \leq n$.

Let us denote

$$\mathcal{P} = \{x \in \mathbb{R}^n \mid Ax = b,\ \ x \geq 0\},$$
$$\mathcal{P}^o = \{x \in \mathbb{R}^n \mid Ax = b,\ \ x > 0\},$$
$\mathcal{P}^*$ is the set of optimal solutions of (P),
$$\mathcal{D} = \{(y,z) \in \mathbb{R}^m \times \mathbb{R}^n \mid A^Ty + z = c,\ \ z \geq 0\},$$
$$\mathcal{D}^o = \{(y,z) \in \mathbb{R}^m \times \mathbb{R}^n \mid A^Ty + z = c,\ \ z > 0\},$$
$\mathcal{D}^*$ is the set of optimal solutions of (D) .

Note that the full rank assumption for A implies a one-to-one correspondence between y and z in the pairs $(y,z) \in \mathcal{D}$, which allowed us to refer to any pair $(y,z) \in \mathcal{D}$ simply as $y \in \mathcal{D}$ or $z \in \mathcal{D}$.

In linear programming the following two statements for the dual pair (LP), (LD) are known as the weak and the strong duality results, respectively.

Proposition 4. *(a)* $\forall\, x \in \mathcal{P},\ \forall\, y \in \mathcal{D}:\ \ c^T x \geq b^T y$.
(b) $\mathcal{P}^* \neq \emptyset \Longleftrightarrow \mathcal{D}^* \neq \emptyset$ *and* $\forall\, x^* \in \mathcal{P}^*,\ \forall\, y^* \in \mathcal{D}^*:\ c^T x^* = b^T y^*$.

As we are going to apply the IPM to both (LP) and (LD) problems, it is natural to assume that the "interiors" of the feasible sets for these problems are non-empty. So throughout this section we will assume that

$$\mathcal{P}^o \neq \emptyset, \quad \mathcal{D}^o \neq \emptyset. \tag{19}$$

Remark 3. It is well known that the following statements are equivalent:

(a) $\mathcal{P}^o \neq \emptyset$, $\mathcal{D}^o \neq \emptyset$, ,
(b) $\mathcal{P}^o \neq \emptyset$, $\mathcal{P}^* \neq \emptyset$, $\mathcal{P}^*$ bounded,
(c) $\mathcal{D}^o \neq \emptyset$, $\mathcal{D}^* \neq \emptyset$, $\mathcal{D}^*$ bounded,
(d) $\mathcal{P}^* \neq \emptyset$, $\mathcal{D}^* \neq \emptyset$, $\mathcal{P}^*$, $\mathcal{D}^*$ bounded.

For a proof see e.g. [**7**].

Now, by analogy with (CP),(CP$_\mu$) and (7), given $\mu > 0$ we assign to (LP) and (LD) the following transformed mathematical programming problems:

$$(\mathrm{LP}_\mu,) \qquad \min \{T(x;\mu) \mid x \in \mathcal{P}^o\}$$

$$(\mathrm{LD}_\mu,) \qquad \max \{Q(y,z;\mu) \mid (y,z) \in \mathcal{D}^*\}$$

where

$$T(x;\mu) = c^T x - \sum_{j=1}^{n} \Lambda(x_j;\mu), \tag{20}$$

$$Q(y,z;\mu) = b^T y + \sum_{j=1}^{n} \Gamma(z_j;\mu) \tag{21}$$

and $\Lambda(\cdot;\mu), \Gamma(\cdot;\mu) \in\ C^1(0,\infty)$. We will assume that both $\Lambda(\cdot;\mu)$ and $\Gamma(\cdot;\mu)$ have (A), (B), (C) properties.

Let $x^\mu > 0$ be an optimal solution of (LP$_\mu$). Then by Lagrange theorem there exists $u^\mu \in \mathbb{R}^n$ (the Lagrange multiplier for the constraint $Ax = b$) such that

$$Ax = b, \quad x > 0 \tag{22a}$$

$$A^T u + v = c, \tag{22b}$$

$$\text{where} \quad v_j^\mu = \Lambda'(x_j^\mu\ ;\mu) > 0. \qquad (j = 1, \ldots, n) \tag{22c}$$

Thus $(u^\mu, v^\mu) \in \mathcal{D}^o$ forms a feasible solution of (LD$_\mu$). Similarly, if (y^μ, z^μ) with $z^\mu > 0$ is an optimal solution of (LD$_\mu$), then there exists $w^\mu \in \mathbb{R}^n$ (the Lagrange multiplier for the constraint $A^T y + z = c$) such that

$$A^T y + z = c, \qquad z > 0 \tag{23a}$$

$$Aw = b, \tag{23b}$$

$$\text{where} \quad w_j^\mu = \Gamma'(z_j^\mu\ ; \mu) > 0. \quad (j = 1, \ldots, n) \tag{23c}$$

Thus $w^\mu \in \mathcal{P}^o$ is a feasible solution of (LP$_\mu$). We can now state our problem for the linear programming case.

For a given function Λ find an appropriate function Γ such that for each $\mu > 0$ the following hold:
(i) The solution (x^μ, u^μ, v^μ) of (22) coincides with the solution (w^μ, y^μ, z^μ) of (23).
(ii) The transformed pair (LP$_\mu$), (LD$_\mu$) exhibits duality relations analogous to those for the original pair (LP), (LD) as given in Proposition 3.

The following theorem gives a solution for this problem.

Theorem 1. *Let (LP), (LD) be the linear programming dual pair satisfying* (19) *defined above. Given $\mu > 0$ let (LP$_\mu$), (LD$_\mu$) be the corresponding transformed pair with $\Lambda(\cdot; \mu)$ satisfying (A), (B), (C) and $\Gamma(\cdot; \mu)$ being the Legendre transformation of $\Lambda(\cdot; \mu)$ (i.e. $\Gamma(\cdot; \mu) = \Lambda_L(\cdot; \mu)$). Then*
(a) $\forall\, x^\mu \in \mathcal{P}^o$, $(y, z) \in \mathcal{D}^o$: $T(x; \mu) \geq Q(y, z; \mu)$,
(b) If x^μ is an optimal solution of (LP$_\mu$), then z^μ defined by

$$z_j^\mu = \Lambda'(x_j^\mu\ ; \mu), \qquad (j = 1, \ldots, n) \tag{24}$$

forms an optimal solution of (LD$_\mu$). And vice versa: if z^μ is an optimal solution of (LD$_\mu$), then x^μ defined by

$$x_j^\mu = \Lambda_L'(z_j^\mu\ ; \mu), \qquad (j = 1, \ldots, n)$$

forms an optimal solution of (LP$_\mu$). In both cases we have

$$T(x^\mu; \mu) = Q(y^\mu, z^\mu; \mu).$$

First we note that by Lemma 1 the function $\Lambda_L(\cdot; \mu) \in C^1(0, \infty)$ and has (A), (B), (C) properties so we are entitled to put $\Gamma = \Lambda_L$ in the formulation of the theorem.

Proof. (a) Let $x \in \mathcal{P}^o$, $(y, z) \in \mathcal{D}^o$. Then obviously $c^T x - b^T y = z^T x$ and using Lemma 1(b) we have

$$\begin{aligned} T(x; \mu) - Q(y, z; \mu) &= \left[c^T x - \sum_{j=1}^n \Lambda(x_j; \mu) \right] - \left[b^T y + \sum_{j=1}^n \Lambda_L(z_j; \mu) \right] \\ &= \sum_{j=1}^n [z_j x_j - \Lambda(x_j; \mu) - \Lambda_L(z_j; \mu)] \geq 0. \end{aligned}$$

(b) By Remark 1(c), $\Lambda(\cdot;\mu)$ is strictly concave and thus $T(\cdot;\mu)$ is strictly concave on $\mathcal{P}^o$. By Lemma 1(iii), $\Lambda_L(\cdot;\mu)$ is also strictly concave and thus $Q(y,z;\mu)$ is strictly concave on $\mathcal{D}^o$. This and (19) with Remark 3 imply that problems (LP_μ) and (LD_μ) each have a unique optimal solution and thus the corresponding necessary and sufficient conditions (22), (23) for optimality each have a unique solution. Moreover, by Lemma 1(iv) $\eta = \Lambda'(\xi;\mu)$ is equivalent to $\xi = \Lambda_L'(\eta;\mu)$ which implies that system (22) is equivalent to system (23). This proves the first part of the statement (b).

It now follows that the optimal solutions x^μ of (LP_μ) and z^μ of (LD_μ) satisfy (24) or, equivalently, $(\Lambda')^{-1}(z_j^\mu;\mu) = x_j^\mu$. Then by Definition 1 we have

$$T(x^\mu;\mu) - Q(y^\mu, z^\mu;\mu) = \sum_{j=1}^n \left[z_j^\mu x_j^\mu - \Lambda(x_j^\mu;\mu) - \Lambda_L(z_j^\mu;\mu)\right] = 0$$

and the theorem is proved. □

The previous result allows us to talk about a duality between (LP_μ) and (LD_μ). So, a transformed problems pair (LP_μ), (LD_μ), where $\Gamma = \Lambda_L$, is called a dual transformed pair, or, a dual pair of transformed problems. Similarly, corresponding transformation functions T, Q, where $\Gamma = \Lambda_L$, are called dual transformation functions.

Note that the original linear programming dual pair (LP), (LD) exhibits some symmetric properties (see sign "$\Longleftrightarrow$" in Proposition 4). The same kind of symmetry is shared by the pair (LP_μ), (LD_μ) (see "vice versa" in Theorem 1) although the problems (LP_μ) and (LD_μ) are not linear. Moreover, having the optimal solution of one of the problems we have an explicit rule to obtain the solution of the other.

Two simple examples of dual transformation functions are:

$$T(x;\mu) = c^T x - \mu \sum_{j=1}^n \ln x_j,$$
$$Q(y,z;\mu) = b^T y + \mu \sum_{j=1}^n \ln z_j + n\mu - n\mu \ln \mu$$

and

$$T(x;\mu) = c^T x - \mu \frac{1}{p} \sum_{j=1}^n x_j^p, \qquad p < 1,\ p \neq 0,$$
$$Q(y,z;\mu) = b^T y + \mu^{(1-q)} \frac{1}{q} \sum_{j=1}^n u_j^q, \qquad \frac{1}{p} + \frac{1}{q} = 1.$$

These results are in agreement with dual functions from [**2**] also mentioned in Section 1. The last pair of functions can be rewritten (by $\nu = \mu^{|p|}$) to the more symmetric form:

$$T(x;\nu) = c^T x - \nu^{|p|}\frac{1}{p}\sum_{j=1}^{n} x_j^p,$$
$$Q(y,z;\nu) = b^T y + \nu^{|p|}\frac{1}{q}\sum_{j=1}^{n} z_j^q.$$

5. Duality of Transformed Problems in Convex Programming

Consider the following convex programming problem

$$\text{(CP)} \qquad \min\left\{ f(x) \mid g^i(x) \geq 0, \quad (i = 1,\ldots,m), x \in X \right\}$$

where f, $-g_i$, $(i = 1,\ldots,m)$, are convex functions defined on an open, nonempty, convex set $X \subset \mathbb{R}^n$. Assume that $f, -g_i \in C^1$.

Let L be the Lagrangian function for this problem, i.e.

$$L(x,u) = f(x) - \sum_{i=1}^{m} u_i g_i(x), \qquad x \in X, \quad u_i \geq 0. \tag{25}$$

The Wolfe dual problem associated with (CP) is

$$\text{(CD)} \qquad \max\left\{L(x,u) \mid \nabla_x L(x,u) = 0,\ u \geq 0,\ x \in X\right\}.$$

Similarly as in the LP case we denote:

$\mathcal{P}_c = \{x \in X \mid g_i(x) \geq 0,\ i = 1,\ldots,m\}$,
$\mathcal{P}_c^o = \{x \in X \mid g_i(x) > 0,\ i = 1,\ldots,m\}$,
$\mathcal{P}_c^*$ is the set of optimal solutions of (CP),
$\mathcal{D}_c = \{(x,u) \mid \nabla_x L(x,u) = 0,\ u \geq 0,\ x \in X\}$,
$\mathcal{D}_c^o = \{(x,u) \mid \nabla_x L(x,u) = 0,\ u > 0,\ x \in X\}$,
$\mathcal{D}_c^*$ is the set of optimal solutions of (CD).

It is well known that the pair (CP), (CD) has the following dual properties:

Proposition 5. *Let (CP), (CD) be the problems defined above. Then*

(a) $\forall\, x \in \mathcal{P}_c,\ \forall\,(y,u) \in \mathcal{D}_c$: $L(y,u) \leq f(x)$.
(b) *If* $\mathcal{P}_c^o \neq \emptyset$ *and* $x \in \mathcal{P}_c^*$, *then there exists* u_x *such that* $(x,u_x) \in \mathcal{D}_c^*$ *and* $L(x,u_x) = f(x)$.

Although the problem (CD) is not convex we can formally transform the problems (CP), (CD) to the unconstrained ones

$$(\mathrm{CP}_\mu) \qquad \min\{T(x;\mu) \mid x \in \mathcal{P}_c^o\},$$

$$(\mathrm{CD}_\mu) \qquad \max\{Q(x,u;\mu) \mid (x,u) \in \mathcal{D}_c^o\},$$

where

$$T(x;\mu) = f(x) - \sum_{i=1}^{m} \Lambda(g_i(x);\mu),$$

$$Q(x,u;\mu) = L(x,u) + \sum_{i=1}^{m} \Gamma(u_i;\mu).$$

Here $\mu > 0$ is a parameter and $\Lambda(\cdot;\mu)$, $\Gamma(\cdot;\mu) \in C^1(0,\infty)$. We assume that for any given $\mu > 0$ the function $\Lambda(\cdot;\mu)$ has the properties (A), (B), (C).

Note that if x^μ is an optimal solution of (CP_μ), then

$$\nabla_x T(x^\mu;\mu) = \nabla_x f(x^\mu) - \sum_{i=1}^{m} \Lambda'(g_i(x^\mu);\mu) \nabla_x g_i(x^\mu) = 0.$$

If we put

$$u_i^\mu = \Lambda'(g_i(x^\mu);\mu), \quad (i = 1,\dots,m) \tag{26}$$

then by Remark 1(c) $u_i^\mu > 0$ and so (x^μ, u^μ) forms the feasible solution not only of (CD) but also of (CD_μ).

The following theorem shows that given any function Λ with properties (A), (B), (C) we can find the function Γ such that the above (x^μ, u^μ) is optimal solution of corresponding problem ȧnd moreover the pair (CP_μ), (CD_μ) has dual properties analogous to those given in Proposition 5.

Theorem 2. *Let (CP), (CD) be the dual pair of convex programming with $\mathcal{P}_c^o \neq \varnothing$ defined above. Given $\mu > 0$, let (CP_μ), (CD_μ) be the corresponding transformed pair with $\Lambda(\cdot;\mu)$ satisfying (A), (B), (C) and $\Gamma(\cdot;\mu)$ being the Legendre transformation of $\Lambda(\cdot;\mu)$. Then*

(a) $\forall\, x \in \mathcal{P}_c^o$, $\forall\,(y,u) \in \mathcal{D}_c^o : T(x;\mu) \geq Q(y,u;\mu)$

(b) *If x^μ is the solution of (CP_μ), then (x^μ, u^μ), where u^μ is given by* (26), *forms an optimal solution of (CD_μ) and*

$$T(x^\mu;\mu) = Q(x^\mu, u^\mu;\mu). \tag{27}$$

Proof. (a) Let $x \in \mathcal{P}_c^o$ and $(y,u) \in \mathcal{D}_c^o$. Due to the convexity of $f(x)$ and $-g_i(x)$ $(i = 1, \ldots, m)$, the Lagrangian function (25) is convex in x. So we have

$$L(x,u) - L(y,u) \geq \nabla_x L(y,u)^T (x-y) = 0, \tag{28}$$

since $(y,u) \in \mathcal{D}_c^o$ and so $\nabla_x L(y,u) = 0$. From (28) and from the inequality of Lemma 1(b) for Legendre transformation we have

$$\begin{aligned} T(x;\mu) - Q(y,u;\mu) &= \left[f(x) - \sum_{i=1}^m \Lambda(g_i(x);\mu) \right] - \left[L(y,u) - \sum_{i=1}^m \Lambda_L(u_i;\mu) \right] \\ &= L(x,u) - L(y,u) + \sum_{i=1}^m \left[g_i(x) u_i - \Lambda(g_i(x);\mu) - \Lambda_L(u_i;\mu) \right] \geq 0 \end{aligned}$$

(b) Now let x^μ be an optimal solution of (CP_μ). As shown above, (x^μ, u^μ) is a feasible solution of (CD_μ). Now because of (26) and the equality (13) in Definition 1 of Legendre function we have

$$\begin{aligned} T(x^\mu;\mu) - Q(x^\mu,u^\mu;\mu) &= f(x^\mu) - \sum_{i=1}^m \Lambda(g_i(x^\mu);\mu) \\ &\quad - \left[f(x^\mu) - \sum_{i=1}^m u_i^\mu g_i(x^\mu) + \sum_{i=1}^m \Lambda_L(u_i^\mu;\mu) \right] \\ &= \sum_{i=1}^m \left[u_i^\mu g_i(x^\mu) - \Lambda(g_i(x^\mu)) - \Lambda_L(u_i^\mu;\mu) \right] = 0. \end{aligned}$$

From this and from statement (a) of this theorem we obtain that (x^μ, u^μ) is an optimal solution of (CD_μ). □

Now we outline a different, more constructive deduction of duality for (CP_μ) and (CD_μ). This proof provides a different view of the duality of transformed problems in convex programming.

Along with the problem (CP_μ), which is unconstrained, we shall consider a connected constrained problem of the larger dimension, i.e.

$$(\mathrm{CP}_\mu^+) \qquad \min \left\{ T^+(x,y;\mu) \mid y > 0,\ x \in X,\ g(x) - y \geq 0 \right\},$$

where

$$T^+(x,y;\mu) = f(x) - \sum_{i=1}^m \Lambda(y_i;\mu). \tag{29}$$

Here Λ is the function from the definition of T for (CP_μ) having properties (A), (B), (C). So, $\Lambda(\cdot;\mu)$ is strictly increasing and thus the problems (CP_μ) and (CP_μ^+) are equivalent in the following sense:

(i) If x^μ is an optimal solution of (CP_μ), then (x^μ, y^μ), where $y^\mu = g(x^\mu)$, is an optimal solution of (CP_μ^+).

(ii) If (x^μ, y^μ) is an optimal solution of (CP_μ^+), then $y^\mu = g(x^\mu)$ and x^μ is an optimal solution of (CP_μ).

(iii) In both cases we have $T(x^\mu;\mu) = T^+(x^\mu, y^\mu;\mu)$.

Now let $\mathcal{L}$ be the Lagrangian function for the problem (CP_μ^+), i.e.,

$$\mathcal{L}(x,y,u) = f(x) - \sum_{i=1}^{m} \Lambda(y_i;\mu) - \sum_{i=1}^{m} u_i(g_i(x) - y_i)$$

for $x \in X$, $y > 0$ and $u \geq 0$.
Obviously

$$\begin{aligned} \nabla_x \mathcal{L}(x,y,u) &= L(x,u), \\ \nabla_{y_i} \mathcal{L}(x,y,u) &= -\Lambda'(y_i;\mu) + u_i, \quad i = 1,\dots,m \end{aligned}$$

Now the Wolfe dual problem associated with (CP_μ^+) is

$$\max\left\{\mathcal{L}(x,y,u) \mid \nabla_x \mathcal{L} = 0,\ \nabla_{y_i}\mathcal{L} = 0\ (i = 1,\dots m),\ u \geq 0,\ y \geq 0\right\}.$$

This is the same as

$$\max\Big\{ L(x,u) + \sum_{i=1}^{m}(u_i y_i - \Lambda(y_i;\mu) \ \Big| \ \nabla_x L(x,u) = 0,\ y > 0,\ u_i = \Lambda'(y_i;\mu),\ u_i \geq 0,\ i = 1,\dots m\Big\}.$$

From the properties of $\Lambda'(\cdot;\mu)$ it follows that the values of the function $\Lambda'(\cdot;\mu)$ are positive. Thus the condition $u \geq 0$ follows from $u_i = \Lambda'(y_i;\mu)$ and so it can be omitted. Further, from the definition of Legendre transformation for the function Λ we have $u_i y_i - \Lambda(y_i;\mu) = \Lambda_L(u_i;\mu)$ under our condition $u_i = \Lambda'(y_i;\mu)$. Therefore the last problem can be rewritten in the form

$$\max\left\{ L(x,u) + \sum_{i=1}^{m} \Lambda_L(u_i;\mu) \mid (x,u) \in \mathcal{D}_c^o \right\}. \tag{30}$$

So, from the Wolfe duality of (CP_μ^+) and (30) a duality of (CP_μ) and (30) follows.

Similarly as in the linear programming case, the previous result allows us to say that the problem (CD_μ), where $\Gamma = \Lambda_L$, is dual to the (CP_μ) and, that the corresponding function Q is dual function to T. Note that the Wolfe dual pair (CP), (CD) does not in general exhibit symmetric dual properties. Actually, we have only one-direction implication in Proposition 5(b) and this is also true for the transformed pair in Theorem 2. But there are also further results in the Wolfe dual theory (e.g. the reverse strong theorem). Due to the Wolfe duality of (CP_μ^+), (CD_μ), $\Gamma = \Lambda_L$, and the equivalence of (CP_μ) with (CP_μ^+) all these results can be adopted to the pair (CP_μ), (CD_μ).

Examples of dual transformation function in convex programming are

$$T(x;\mu) = f(x) - \mu \sum_{i=1}^{m} \ln(g_i(x)), \tag{31}$$

$$Q(x,u;\mu) = L(x,u) + \mu \sum_{i=1}^{m} \ln u_i + m\mu - m\mu \ln \mu \tag{32}$$

and

$$T(x;\mu) = f(x) - \mu \frac{1}{p} \sum_{i=1}^{m} (g_i(x))^p, \qquad p < 1, p \neq 0, \tag{33}$$

$$Q(x,u;\mu) = L(x,u) + \mu^{(1-q)} \frac{1}{q} \sum_{i=1}^{m} (u_i)^q, \qquad \frac{1}{p} + \frac{1}{q} = 1. \tag{34}$$

5 Concluding Remarks and Consequences

In this section we treat connections between the duality of the transformed problems and the basic existence and convergence statements given in Propositions 1 and 2. Note that these propositions were formulated for the convex problem (CP), where $K^o \neq \emptyset$, $K^* \neq \emptyset$, K is bounded. In the linear programming case we can apply them to both (LP_μ) and (LD_μ) since both (LP), (LD) are linear and thus also convex. However, in the convex programming case we have to be careful, because (CD_μ) need not be convex. Nevertheless, Proposition 2 is valid for (CD_μ) too, as we can see from the next proposition.

Proposition 3. *Let the set of optimal solution of (CD) be nonempty and bounded and let (CD_μ) be the corresponding transformed problem. Let (CD_μ) satisfy assumptions of Proposition* 2, *i.e.,* Γ *satisfies the assumptions formulated for* Λ *and* (x^μ, u^μ) *is an optimal solution of (CD_μ). Then the statement of Proposition* 2 *holds for problem (CD_μ) too, where* (11) *and* (12) *are replaced by*

$$\lim_{k \to 0} Q(x^k, u^k; \mu_k) = L^* \tag{35}$$

and

$$\lim_{k \to 0} L(x^k, u^k) = L^* \tag{36}$$

resp., where L^* is the optimal value of L. Moreover, if $\{\mu_k\}$ is monotonic and Γ satisfies (9)*, then the sequence $\{L(x^k, u^k)\}_{k=1}^{\infty}$ is monotonic too.*

The proof follows the line of the proof of Proposition 2 and so it can be omitted.

In the previous sections the duality between (CP_μ) and (CD_μ), $\Gamma = \Lambda_L$, was deduced for any fixed value $\mu > 0$. The main assumptions were that Λ has properties (A), (B), (C). Note that the same assumptions appear in Proposition 1 and thus an optimal solution of (CP_μ) exists for any $\mu > 0$. Then by Theorem 1, (CD_μ) with $\Gamma = \Lambda_L$ has also an optimal solution and thus it is no surprise that by Lemma 2 the "dual" function Λ_L has also properties (A), (B), (C). This result once again justifies the formulation of sufficient conditions for existence in the form (A), (B), (C). In this connection the "dual" relationship between (A) and (B) given in the proof of Lemma 2 is very interesting.

Besides the assumptions (A), (B), (C) there are also further ones in IPM which allow to state the convergence results formulated in Proposition 2. However, by Lemmas 3,4, if Λ has properties (8) and (9), (9a) or (10) from Proposition 2, then even though Λ_L has property (8), it may have none of the properties (9), (9a), (10). This enables us to formulate new sufficient conditions for convergence:

Theorem 3. *Let (CP_μ) be a transformed problem for (CP). Let the corresponding Λ have properties (A), (B), (C) and let Λ_L satisfy the assumptions of Proposition* 2 *(formulated for Λ). Then the statement* (11) *of Proposition* 2 *holds for any sequence $\{\mu_k\}$, $\mu_k > 0$, $\mu_k \to 0$.*

Proof. By Proposition 1, there exists an optimal solution of (CP_{μ_k}) for any $k > 0$. Then by Theorem 1, (x^k, u^k), where $u_i^k = (\Lambda')^{-1}(x_i^k)$, $(i = 1, \dots, m)$, is an optimal solution of (CD_{μ_k}) (with $\Gamma = \Lambda_L$). Since Λ_L and (CD_{μ_k}) satisfy the assumptions of Proposition 3 we have (35). But due to Theorem 1 we have $Q(x^k, u^k; \mu_k) = T(x^k; \mu_k)$ and so (11) is proved. □

An important property which facilitates the development of algorithms is a monotonicity of convergence. The basic convergence theorem in IPM formulated by Proposition 2 states the monotonic convergence of $f(x^k)$ to $f(x^*)$ in the case when Λ is linear in parameter. One should for convenience also use a monotonic convergence of $L(x^k, u^k)$, where u^k is given by (26) and L is Lagrangian function for (CP). But an example from [**6**] shows that this is not generally true.

In [**2**] and [**6**] it was shown that if T is of the form (31) or (33), then the corresponding $L(x^k, u^k)$ is monotonic in μ. To prove this statement, the duality between (31) or (32) and (33) or (34) resp. was used in [**2**]. Actually, if T is of the form (33), then the corresponding dual Q is of the form (34) and it should be viewed as linear function of ν (where $\nu = \mu^{1-q}$). Then we can apply Proposition 3 to the problem (CD_ν), where Q is given by (32) or (34) and so, because of linearity of Q in μ or ν we have monotonic convergence of $L(x^\mu, u^\mu)$ from Proposition 3.

Remark 3 from Section 3 states that functions Λ from Remark 2 are the only linear, satisfying (A), (B), (C) ones for which Λ_L is quasilinear again. From this we can conclude that the two types of transformations functions mentioned above (i.e. (31), (32)) are the only ones for which the monotonicity of corresponding L could be proved by argument of linearity.

In conclusion we note that the duality of transformation functions deduced in this paper for the interior point approach can be generalized to the wider class of transformation functions not necessarily interior. Namely, the only one that we really needed from assumptions (A), (B), (C) was (C) and $\Lambda' > 0$. On the other hand this duality could be treated as a special case of the abstract duality theory of Rockafellar's generalized programs from [**9**].

References

1. den Hertog D., Roos C. and Terlaky T., *Inverse Barrier Methods for Linear Programming*, Technical Report 91-27, Faculty of Mathematics and Informatics, TU Delft, NL-2628 BL Delft.

2. ______, *On the Monotonicity of the Dual Objective along Barrier Paths*, COAL Bulletin **20** (1992), 2–8.

3. Fiacco A. V. and McCormic G. P., *Nonlinear Programming, Sequential Unconstrained Minimization Techniques*, Wiley and Sons, New York, 1968.

4. Hamala M., *Quasibarrier Methods for Convex Programming*, in Survey of Mathematical Programming I, A. Prekopa, North Holland, 1979, pp. 465–477.

5. ______, *A General Approach to Interior Point Transformation Methods for Mathematical Programming*, Acta Math. Univ. Comenianae **LIV-LV** (1988), 243–266.

6. Hamala M. and Halická M., *Monotonicity of Lagrangian Function in the Parametric Interior Point Methods of Convex Programming*, Acta Math. Univ. Comenianae **LXI(1)** (1992), 41–55.

7. Jansen B., Roos C., Terlaky T. and Vial J-Ph., *Interior-Point Methodology for Linear Programming: Duality, Sensitivity Analysis and Computational Aspects*, Technical Report 93-28, Faculty of Mathematics and Informatics, TU Delft, NL-2628 BL Delft.

8. Pexider J. V., *Notiz uber Funkcionaltheoreme*, Monatsh. Math. Phys. **14** (1903), 293–301.

9. Rockafellar R. T., *Convex Analysis*, Princeton University Press, 1970.

M. Halická, Institute of Applied Mathematics, Faculty of Mathematics and Physics, Comenius University, 84215 Bratislava, Slovakia, *e-mail:* halicka@fmph.uniba.sk

M. Hamala, Department of Numerical and Optimization Methods, Faculty of Mathematics and Physics, Comenius University, 84215 Bratislava, Slovakia