Annals of Mathematics, **169** (2009), 661–674

The concept of duality in convex analysis, and the characterization of the Legendre transform

By Shiri Artstein-Avidan and Vitali Milman*

Abstract

In the main theorem of this paper we show that any involution on the class of lower semi-continuous convex functions which is order-reversing, must be, up to linear terms, the well known Legendre transform.

1. Introduction

The notion of duality is one of the central concepts both in geometry and in analysis. At the same time, it is usually defined in a very concrete way, using very concrete structures. It turns out, however, that in many of the central examples of duality in geometry and analysis, the standard definitions arise, explicitly, from two very simple and natural properties: *involution*, and *order reversion*. This basic fact, which, it seems, was not discovered until now despite the very common use of duality throughout mathematical research, can be seen in the following result regarding the Legendre transform, which, together with its consequences, is the main subject of this paper.

Denote the class of lower-semi-continuous convex functions $\phi : \mathbb{R}^n \to \mathbb{R} \cup \{\pm\infty\}$ by $Cvx(\mathbb{R}^n)$ (so that the only function attaining the value $-\infty$ is the constant $-\infty$ function). Denote by $\langle \cdot, \cdot \rangle$ the standard scalar product on $\mathbb{R}^n$. Recall the definition of the classical Legendre transform $\mathcal{L} : Cvx(\mathbb{R}^n) \to Cvx(\mathbb{R}^n)$ given by

$$(\mathcal{L}\phi)(x) = \sup_y \left(\langle x, y \rangle - \phi(y) \right). \tag{1}$$

For background on the Legendre transform and its many applications, see, e.g., [Arn], [Ho] and [Ro]. The following theorem shows why it is, in some sense, the only natural transform to associate with duality of convex functions.

*The authors were both partially supported by BSF grant no. 2006079. This first named author was partially supported by the ISF grant no. 865/07, and the second named author was partially supported by an ISF grant.

THEOREM 1. *Assume a transform* $\mathcal{T} : Cvx(\mathbb{R}^n) \to Cvx(\mathbb{R}^n)$ (*defined on the whole domain* $Cvx(\mathbb{R}^n)$) *satisfies*

1. $\mathcal{T}\mathcal{T}\phi = \phi$,
2. $\phi \le \psi$ *implies* $\mathcal{T}\phi \ge \mathcal{T}\psi$.

Then, $\mathcal{T}$ *is essentially the classical Legendre transform*; *namely there exists a constant* $C_0 \in \mathbb{R}$, *a vector* $v_0 \in \mathbb{R}^n$, *and an invertible symmetric linear transformation* $B \in \mathrm{GL}_n$ *such that*

$$(\mathcal{T}\phi)(x) = (\mathcal{L}\phi)(Bx + v_0) + \langle x, v_0\rangle + C_0.$$

Thus, the two seemingly weak conditions, which we are used to associating with any geometric 'duality' transform, are already enough to imply, up to an additive linear function (compensated by a translation) and a symmetric linear transformation, the concrete form of the Legendre transform.

It turns out that even the involution condition (1) is stronger than needed, and variants of Theorem 1 where this condition is replaced by a weaker one will be presented in Section 5 below.

The main part of this note will be dedicated to proving this theorem and some of its applications. However, we have obtained several more results of this type, and in fact, in every form of geometric and functional dualities we were interested in, we have been able to prove a similar result. These cases include duality for positive functions concave on their support (and, more generally, s-concave functions), and subclasses of this class, log-concave duality and more. The case of indicator functions of compact convex bodies with 0 in the interior was settled recently by Böröczky and Schneider [BSc]. We present the general picture, as we view it, and some of these additional results, without proofs, in the last section of this note. For a more elaborate discussion we refer the reader to the announcement paper [AM1].

2. First observation: Interchanging inf and sup

The class of convex functions is closed under the operation of taking the supremum of a family of functions, but not under infimum. To this end we introduce the notation for 'regularized infimum', $\hat{\inf}$, which will denote the largest convex function which is less than or equal to each of the functions in the family. More precisely,

$$\hat{\inf}_\alpha\{f_\alpha\} = \sup\{f : f \in Cvx(\mathbb{R}^n),\ \text{and}\ f \le f_\alpha\ \forall\alpha\},$$

and so clearly $\hat{\inf}_\alpha\{f_\alpha\} \in Cvx(\mathbb{R}^n)$.

For later use, the following lemma is given under weaker conditions than those of Theorem 1.

LEMMA 2. *Assume we are given a transformation $\mathcal{T} : Cvx(\mathbb{R}^n) \to Cvx(\mathbb{R}^n)$, invertible, satisfying*

a. $\phi \le \psi$ *implies* $\mathcal{T}\phi \ge \mathcal{T}\psi$,

b. $\mathcal{T}\phi \le \mathcal{T}\psi$ *implies* $\phi \ge \psi$.

Then $\mathcal{T}(\hat{\inf}(f_\alpha)) = \sup(\mathcal{T} f_\alpha)$, *and vice versa.*

Proof. Indeed, by surjectivity we have that $\sup_\alpha(\mathcal{T} f_\alpha) = \mathcal{T}h$ for some $h \in Cvx(\mathbb{R}^n)$, so $\mathcal{T}h \ge \mathcal{T} f_\alpha$ for all α, and so by (b) we must have $h \le f_\alpha$ for all α, and since $h \in Cvx(\mathbb{R}^n)$ we have $h \le h' = \hat{\inf}_\alpha(f_\alpha)$. Next, as $h' \le f_\alpha$ for all α, we must have by (a.) that $\mathcal{T}h' \ge \mathcal{T} f_\alpha$ for all α, and so $\mathcal{T}h' \ge \sup_\alpha \mathcal{T} f_\alpha = \mathcal{T}h$, and be condition (b) we get that $h' \le h$. We thus have equality $h = h'$.

The other side is proved in the exact same way, $\hat{\inf}_\alpha(\mathcal{T} f_\alpha) = \mathcal{T}h$ for some $h \in Cvx(\mathbb{R}^n)$, so, $\mathcal{T}h \le \mathcal{T} f_\alpha$ for all α, and so by (b) we must have $h \ge f_\alpha$ for all α, and since $h \in Cvx(\mathbb{R}^n)$ we have $h \ge h' = \sup_\alpha(f_\alpha)$. However, by (a.) $\mathcal{T}h' \le \mathcal{T} f_\alpha$ for every α, so that $\mathcal{T}h' \le \hat{\inf}_\alpha(\mathcal{T} f_\alpha) = \mathcal{T}h$ and by (b) again, $h' \ge h$, so there is equality $h = h'$. $\square$

3. Second observation: Linear and delta-functions suffice

For simplicity of notation, for $\theta \in \mathbb{R}^n$ we will denote the function $-\log \delta_\theta(x)$ by $D_\theta(x)$, that is,

$$D_\theta(\theta) = 0 \text{ and } D_\theta(y) = +\infty \quad \text{for } y \ne \theta.$$

we will refer to these functions throughout the paper as 'delta-type functions'. These functions clearly belong to $Cvx(\mathbb{R}^n)$, and so do their parallels (or shifts) $D_\theta + c$. Moreover, any function can be expressed as the infimum of such shifted functions, namely $f(x) = \inf_y(D_y(x) + f(y))$. Taking this into account, the following theorems are quite straightforward.

LEMMA 3. *Assume a* 1-1 *and onto transform* $\mathcal{T} : Cvx(\mathbb{R}^n) \to Cvx(\mathbb{R}^n)$ (*defined on the whole domain* $Cvx(\mathbb{R}^n)$) *satisfies*

1. $\phi \le \psi$ *implies* $\mathcal{T}\phi \ge \mathcal{T}\psi$,
2. $\mathcal{T}\phi \le \mathcal{T}\psi$ *implies* $\phi \ge \psi$,
3. *There exist* $C_0, \in \mathbb{R}$, $C_1 > 0$, $B \in \mathrm{GL}_n$ *and* $v_0, v_1 \in \mathbb{R}^n$ *such that for any* θ *and* c

$$\left(\mathcal{T}(D_\theta + c)\right)(x) = \langle B\theta + v_1, x\rangle + \langle v_0, \theta\rangle - C_1 c + C_0.$$

Then $\mathcal{T}$ *is a variant of the Legendre transform defined by*

$$(\mathcal{T}\phi)(x) = C_0 + \langle v_1, x\rangle + C_1(\mathcal{L}\phi)(B'x + v_0').$$

for $B' = B^*/C_1 \in \mathrm{GL}_n$ *and* $v_0' = v_0/C_1$.

Proof. Let

$$\phi(x) = \left(\hat{\inf}_y(\phi(y) + D_y)\right)(x),$$

so that by Lemma 2 we have that

$$(\mathcal{T}\phi)(x) = \sup_y \mathcal{T}(\phi(y) + D_y)(x),$$

which by (3) equals

$$\begin{aligned}(\mathcal{T}\phi)(x) &= \sup_y(\langle By + v_1, x\rangle + \langle v_0, y\rangle - C_1\phi(y) + C_0) \\ &= C_0 + \langle v_1, x\rangle + \sup_y(\langle y, v_0 + B^*x\rangle - C_1\phi(y)) \\ &= C_0 + \langle v_1, x\rangle + C_1 \sup_y(\langle y, B'x + v_0'\rangle - \phi(y)) \\ &= C_0 + \langle v_1, x\rangle + C_1(\mathcal{L}\phi)(B'x + v_0').\end{aligned}$$

□

Secondly, we will need a version which applies to the involution condition, which is stronger than the second condition above.

LEMMA 4. *Assume a transform $\mathcal{T} : Cvx(\mathbb{R}^n) \to Cvx(\mathbb{R}^n)$ (defined on the whole domain $Cvx(\mathbb{R}^n)$) satisfies*

1. $\mathcal{T}\mathcal{T}\phi = \phi$,
2. $\phi \le \psi$ *implies* $\mathcal{T}\phi \ge \mathcal{T}\psi$,
3. *There exist $C_0 \in \mathbb{R}$, $C_1 > 0$, $v_0, v_1 \in \mathbb{R}^n$ and $B \in \mathrm{GL}_n$ such that for any θ and c*

$$\mathcal{T}(D_\theta + c) = \langle B\theta + v_1, \cdot\rangle + \langle v_0, \theta\rangle - C_1 c + C_0.$$

Then B must be symmetric, $B = B^*$, $C_1 = 1$, $v_0 = v_1$, and $\mathcal{T}$ is a variant of the Legendre transform, given by

$$(\mathcal{T}\phi)(x) = C_0 + \langle v_0, x\rangle + (\mathcal{L}\phi)(Bx + v_0).$$

Proof. Lemma 3 implies that the transform has the form

$$(\mathcal{T}\phi)(x) = C_0 + \langle v_1, x\rangle + C_1(\mathcal{L}\phi)(B'x + v_0')$$

for $B' = B^*/C_1 \in \mathrm{GL}_n$ and $v_0' = v_0/C_1$.

We will apply the transform twice to a general function ϕ. Apply first the Legendre transform to $\mathcal{T}\phi$:

$$\begin{aligned}(\mathcal{L}\mathcal{T}\phi)(x) &= \sup_y \left(\langle x, y\rangle - (\mathcal{T}\phi)(y)\right)\\ &= -C_0 + \sup_y \left(\langle x - v_1, y\rangle - C_1(\mathcal{L}\phi)(B'y + v_0')\right)\\ &= -C_0 + \sup_z \left(\langle x - v_1, (B')^{-1}(z - v_0')\rangle - C_1(\mathcal{L}\phi)(z)\right)\\ &= -C_0 - \langle x - v_1, (B')^{-1}v_0'\rangle + C_1(\mathcal{L}\mathcal{L}\phi)((\frac{(B')^{-1}}{C_1})^*(x - v_1))\\ &= -C_0 - \langle x - v_1, (B^*)^{-1}v_0\rangle + C_1\phi(B^{-1}(x - v_1)).\end{aligned}$$

We then use the formula defining $\mathcal{T}$ and get that

$$\begin{aligned}(\mathcal{T}\mathcal{T}\phi)(x) &= C_0 + \langle v_1, x\rangle + C_1(\mathcal{L}(\mathcal{T}\phi))(B'x + v_0')\\ &= C_0 + \langle v_1, x\rangle - C_1(C_0 + \langle B'x + v_0' - v_1, (B^*)^{-1}v_0\rangle)\\ &\quad + C_1^2\phi(B^{-1}(B'x + v_0' - v_1)).\end{aligned}$$

This last expression must equal $\phi(x)$ for every x and ϕ. Letting $\phi = 0$ we see that for every x

$$C_0 + \langle v_1, x\rangle - C_1C_0 - C_1\langle B'x + v_0' - v_1, (B^*)^{-1}v_0\rangle = 0$$

which in particular means that $B(B^*)^{-1}v_0 = v_1$. The formula becomes

$$\phi(x) = C_1^2\phi(B^{-1}(B'x + v_0' - v_1)).$$

Obviously this means that $C_1 = 1$, say, by looking at a constant function. This leaves us with $B^{-1}(B'x + v_0' - v_1) = x$, which implies that $(v_0 =)v_0' = v_1$, and that $B = B^*$, as required. $\square$

4. Linear and delta-functions, and a proof of Theorem 1

We now proceed to the proof of the main theorem. To this end, we show the following theorem, which together with Lemma 4 gives Theorem 1.

THEOREM 5. *Assume an invertible transform $\mathcal{T} : Cvx(\mathbb{R}^n) \to Cvx(\mathbb{R}^n)$ satisfies*

1. *$\phi \le \psi$ implies $\mathcal{T}\phi \ge \mathcal{T}\psi$,*
2. *$\mathcal{T}\phi \le \mathcal{T}\psi$ implies $\phi \ge \psi$.*

Then, there exist constants $C_0 \in \mathbb{R}, C_1 > 0$, two vectors $v_0, v_1 \in \mathbb{R}^n$, and $B \in \mathrm{GL}_n$ such that for any θ and c,

$$\mathcal{T}(D_\theta + c) = \langle B\theta + v_1, \cdot\rangle + \langle v_0, \theta\rangle - C_1c + C_0.$$

The proof is composed of several steps: first we show that delta functions must be mapped to linear functions and vice versa. Then we show that this correspondence is in essence linear, and then we interpret this linearity.

Step One. We will show that for any $\theta \in \mathbb{R}^n$ and $c \in \mathbb{R}$ there exist $a_{\theta,c} \in \mathbb{R}^n, A_{\theta,c} \in \mathbb{R}$ such that

$$D_\theta + c = \mathcal{T}(\langle a_{\theta,c}, \cdot\rangle + A_{\theta,c}).$$

Assume $\mathcal{T}\phi = D_\theta + c$. (We know there exists such a ϕ, from surjectivity). We claim that each two linear functionals h_1, h_2 satisfying $h_i \le \phi$, $i = 1, 2$, differ by a constant from one another. This means that ϕ itself is linear, since any convex function which is not linear has at least two different supporting hyperplanes. Indeed, assume ϕ had two linear functionals $h_i = \langle \cdot, a_i\rangle + A_i$, for $i = 1, 2$, which are below it. Since $\mathcal{T}$ is 1-1, the two functions $\mathcal{T}h_1$ and $\mathcal{T}h_2$ are different. On the other hand, they both have to be greater than or equal to $D_\theta + c$ at each point x, and so can be $\neq \infty$ only at θ. From this we conclude that they are both equal to $D_\theta + c_i$ for some constants c_1 and c_2, respectively. But then, from Lemma 2, $\mathcal{T}(\max(h_1, h_2)) = D_\theta + \min(c_1, c_2)$, which contradicts the injectivity unless $\max(h_1, h_2) = h_i$ for one of the $i = 1, 2$, which means that they are parallel.

Step Two. In the same way we may show that for every $a \in \mathbb{R}^n$ and $A \in \mathbb{R}$ there exist $\theta_{a,A}$ and $c_{a,A}$ such that

$$D_{\theta_{a,A}} + c_{a,A} = \mathcal{T}(\langle a, \cdot\rangle + A).$$

(That is, the class of linear functions is mapped, onto, the class of delta-type functions.) Indeed, denote $\phi_{a,A} = \langle a, \cdot\rangle + A$. All we have to show is that the support of $\mathcal{T}\phi_{a,A}$ is one point. This follows from the fact that every two functions which are greater than $\mathcal{T}\phi_{a,A}$ are images of two functions which are less than $\phi_{a,A}$, and are hence, in particular, comparable (one is greater than the other). However, if the support of $\mathcal{T}\phi_{a,A}$ includes two different points, x_1 and x_2, then above it are, for some constants α_1 and α_2, both $D_{x_1} + \alpha_1$ and $D_{x_2} + \alpha_2$, which are not comparable.

Step Three. In the same way, for any $\theta \in \mathbb{R}^n$ and $c \in \mathbb{R}$ there exist $a'_{\theta,c} \in \mathbb{R}^n, A'_{\theta,c} \in \mathbb{R}$ such that

$$\mathcal{T}(D_\theta + c) = \langle a'_{\theta,c}, \cdot\rangle + A'_{\theta,c},$$

and moreover this mapping is onto, namely for every $a \in \mathbb{R}^n$ and $A \in \mathbb{R}$ there exist $\theta_{a,A} \in \mathbb{R}^n$ and $c_{a,A} \in \mathbb{R}$ such that

$$\mathcal{T}(D_{\theta_{a,A}} + c_{a,A}) = \langle a, \cdot\rangle + A.$$

Indeed, this can be achieved directly, but one may instead notice that the conditions regarding $\mathcal{T}$ and $\mathcal{T}^{-1}$ are identical, and so by the first two steps applied to $\mathcal{T}^{-1}$ instead of $\mathcal{T}$ we get the required conclusion.

Step Four. From the above it follows that for any θ and c there exists another c' such that $\mathcal{T}(D_\theta + c) = \mathcal{T}(D_\theta) + c'$, and that for any $a \in \mathbb{R}^n$ and $A \in \mathbb{R}$ there exists $A' \in \mathbb{R}$ with $\mathcal{T}(\langle a, \cdot\rangle + A) = \mathcal{T}(\langle a, \cdot\rangle) + A'$ (simply because comparable pairs are mapped to comparable pairs). Note that from anti-monotonicity, the signs of c and c' must be opposite, and also of A and A', a fact which we will use later.

Thus the above also implies that for any θ

$$\mathcal{T}\{D_\theta + c : c \in \mathbb{R}\} = \{\mathcal{T} D_\theta + c : c \in \mathbb{R}\},$$

and that for any a

$$\mathcal{T}\{\langle a, \cdot\rangle + A : A \in \mathbb{R}\} = \{\mathcal{T}\langle a, \cdot\rangle + A : A \in \mathbb{R}\}.$$

More precisely, define the two mappings $G_1, G_2 : \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$ in the following way: $G_1(\theta, c) = (a', A')$ where $\mathcal{T}(D_\theta + c) = \langle a', \cdot\rangle + A'$, and $G_2(a, A) = (\theta, c)$ where $\mathcal{T}(\langle a, \cdot\rangle + A) = D_\theta + c$. Then G_1 maps the interval $\{(\theta, tc) : t \in [t_1, t_2]\}$ to an interval $\{(a_c, A) : A \in [A_1, A_2]\}$, and G_2 maps an interval $\{(a, A) : A \in [A_1, A_2]\}$ to an interval $\{(\theta_a, tc) : t \in [t_1, t_2]\}$.

Step Five. We will now show that the mappings G_i map any interval to an interval. For fixed $\theta_0, c_0, \theta_1, c_1$, denoting $\mathcal{T}(D_{\theta_0} + c_0) = \phi_0$ and $\mathcal{T}(D_{\theta_1} + c_1) = \phi_1$, we show that

$$\begin{aligned}\mathcal{T}\{D_\theta + c_\theta : (\theta, c_\theta) &= \lambda(\theta_0, c_0) + (1-\lambda)(\theta_1, c_1), 0 \le \lambda \le 1\}\\ &= \{\mu\phi_0 + (1-\mu)\phi_1 : 0 \le \mu \le 1\}.\end{aligned}$$

This implies that the mapping G_1 defined in Step Four above, maps intervals in $\mathbb{R}^{n+1}$ to intervals. We then show a corresponding fact for G_2^{-1}.

Indeed, look at $\hat{\min}(D_{\theta_0} + c_0, D_{\theta_1} + c_1)$. This is a function which is linear in the interval $[\theta_0, \theta_1]$ (joining (θ_0, c_0) and (θ_1, c_1)) and ∞ elsewhere. Its transform is, by Lemma 2, equal to $\max(\phi_1, \phi_2)$. For every $0 \le \lambda \le 1$, the function $D_\theta + c_\theta$ for $\theta = \lambda\theta_0 + (1-\lambda)\theta_1$ and $c_\theta = \lambda c_0 + (1-\lambda)c_1$ is clearly above this $\hat{\min}$, and thus its image under $\mathcal{T}$ is below $\max(\phi_1, \phi_2)$. However, its image is a linear function $\langle a, \cdot\rangle + A$, and so $a \in [a_0, a_1]$ where $a_i = \nabla\phi_i$ (recall that ϕ_i are linear). That is, for some $0 \le \mu \le 1$ we have $a = \mu a_0 + (1-\mu)a_1$. Secondly, the graph of $\mathcal{T}(D_\theta + c_\theta)$ must touch the graph of $\max(\phi_1, \phi_2)$, otherwise for some positive c'' we have $\mathcal{T}(D_\theta + c_\theta) + c'' \le \max(\phi_1, \phi_2)$ and then (applying $\mathcal{T}^{-1}$), $D_\theta + c_\theta - c''' \ge \hat{\min}(D_{\theta_1}, D_{\theta_2})$ which does not hold for any positive c'''. The fact that it touches the graph implies that also $A = \mu A_0 + (1-\mu)A_1$, thus, $\mathcal{T}(D_\theta + c_\theta) \in [\phi_1, \phi_2]$, as required.

The same considerations can be applied to $\mathcal{T}^{-1}$, and so G_2^{-1} maps intervals to intervals (which will imply the same for G_2, as we will see in the next step).

Step Six. The conclusion of Step Five implies that the mapping $G_1 : \mathbb{R}^{n+1} \to \mathbb{R}^{n+1}$ maps all straight lines in $\mathbb{R}^{n+1}$ to straight lines, and, what is formally stronger, maps intervals to intervals. Notice that $n+1 \geq 2$, and thus we can apply the following fact, called the fundamental fact of affine geometry:

FACT. *For $m \geq 2$, an injective mapping $F : \mathbb{R}^m \to \mathbb{R}^m$ which maps all straight lines to straight lines, must be an affine linear map* (*i.e.*, *there is some $B \in \mathrm{GL}_m$ such that $F(x) = F(0) + Bx$*).

Remark 6. The above fact is well known, and plays a basic role in projective geometry, see for example [Art]. For the convenience of the reader we provide the proof for the formally weaker statement which we use (namely, that such a map, mapping intervals to intervals, is affine linear). To prove it one first notices that the case $F(0) = 0$ and $m = 2$ is enough, since the fact that straight lines are mapped to straight lines implies that affine two-dimensional subspaces are mapped to affine two-dimensional subspaces, and a translation is always possible. In $m = 2$ one first notices that for linearly independent x any y, one has that $F(x)$ and $F(y)$ are linearly independent (or it would contradict injectivity). Second, one uses the fact that parallel lines must be mapped to parallel lines (since they must not intersect), to see that $F(x+y)$ must lie on both lines $\{F(x+cy) : c \in \mathbb{R}\} = F(x) + cF(y)$ and $\{F(cx+y) : c \in \mathbb{R}\} = F(y) + cF(x)$. These intersect at $F(x) + F(y)$ only; hence $F(x+y) = F(x) + F(y)$. We are almost done, since, taking x linearly independent of y and such that $x+y$ is linearly independent of y as well, we have

$$F(x) = F(x+y-y) = F(x+y) + F(-y) = F(x) + F(y) + F(-y),$$

which implies $F(-y) = -F(y)$, and so

$$F(x+x) = F(x+y) + F(x-y) = 2F(x).$$

Similarly for any $p \in \mathbb{N}$ we have $F(px) = pF(x)$. Assume that $F(ax) = bF(x)$ and that $\frac{p}{q} \leq a \leq \frac{p+1}{q}$, then $F(aqx) = qF(ax) = qbF(x)$ but, from the conditions on intervals,

$$F(aqx) \in [F(px), F((p+1)x)] = [pF(x), (p+1)F(x)],$$

so that $b \in [p/q, (p+1)/q]$. Thus, $a = b$, and the mapping is truly linear $F(ax) = aF(x)$.

We would also like to remark that this fact, although relatively simple, has very interesting consequences not only for geometry but also in various other fields of mathematics; for example it has surprisingly appeared in decision theory applications communicated to us by David Schmeidler [BGSS].

Thus, G_1 is given by an affine linear map, $G_1(\theta, c) = B_1(\theta, c) + V_1$ for $B_1 \in \mathrm{GL}_{n+1}, V_1 \in \mathbb{R}^{n+1}$. Similarly, G_2^{-1} is affine linear, and therefore so is G_2, so we write that $G_2(a, A) = B_2(a, A) + V_2$ for $B_2 \in \mathrm{GL}_{n+1}, V_2 \in \mathbb{R}^{n+1}$.

Moreover, Step Four implies that the a-coordinate of $G_1(\theta, c)$ does not depend on c, and that the θ-coordinate of G_2 does not depend on A. Therefore B_1 and B_2, in matrix form, have zeros in all the entries of their last column except for the $(n+1)$-th entry.

Step Seven. We thus see that for some $B \in \mathrm{GL}_n$, $v_0, v_1 \in \mathbb{R}^n$, $C_1, C_0 \in \mathbb{R}$, we have $G_1(\theta, c) = (B\theta + v_1, \langle v_0, \theta\rangle - C_1 c + C_0)$. (Indeed, simply let B be the first $n \times n$ block of B_1, $V_1 = (v_1, C_0)$, and (v_0, C_1) be the $(n+1)^{th}$ row of B_1.) From the reversion of order it is clear that $C_1 > 0$.

That is,

$$\mathcal{T}(D_\theta + c) = \langle \cdot, B\theta + v_1\rangle + \langle v_0, \theta\rangle - C_1 c + C_0.$$

This completes the proof of Theorem 5. □

Proof of Theorem 1. The above equation, namely the conclusion of Theorem 5, is exactly the form given in the statement of Lemma 4, and so we arrive at the conclusions of that lemma, namely that $C_1 = 1$, B is symmetric, and that for any function ϕ

$$(\mathcal{T}\phi)(x) = C_0 + \langle v_0, x\rangle + (\mathcal{L}\phi)(Bx + v_0)).$$

This completes the proof of Theorem 1. □

5. Two consequences and two remarks

When the condition of involution is replaced by the weaker fact that both $\mathcal{T}$ and its inverse are order reversing, one produces a similar result, with some additional terms. More precisely

THEOREM 7. *Assume an invertible transform $\mathcal{T} : Cvx(\mathbb{R}^n) \to Cvx(\mathbb{R}^n)$ satisfies*

1. *$\phi \le \psi$ implies $\mathcal{T}\phi \ge \mathcal{T}\psi$,*
2. *$\mathcal{T}\phi \le \mathcal{T}\psi$ implies $\phi \ge \psi$.*

Then, there exist constants $C_0 \in \mathbb{R}, C_1 > 0$, two vectors $v_0, v_1 \in \mathbb{R}^n$, and an invertible linear transformation $B \in \mathrm{GL}_n$ such that

$$(\mathcal{T}\phi)(x) = C_0 + \langle v_1, x\rangle + C_1(\mathcal{L}\phi)(Bx + v_0).$$

Proof. Combine Theorem 5 with Lemma 3, and change notation slightly. □

The above theorem is useful for determining all the order *preserving* 1-1 and onto maps from $Cvx(\mathbb{R}^n)$ to itself. Each such 1-1 and onto order preserving map is composed of a translation in the argument, multiplication by a positive constant, and addition of an affine linear function. More precisely, we may conclude the following

COROLLARY 8. *Assume an invertible transform $\mathcal{F} : Cvx(\mathbb{R}^n) \to Cvx(\mathbb{R}^n)$ satisfies*

1. $\phi \le \psi$ *implies* $\mathcal{F}\phi \le \mathcal{F}\psi$,
2. $\mathcal{F}\phi \le \mathcal{F}\psi$ *implies* $\phi \le \psi$.

Then there exist constants $C_0 \in \mathbb{R}, C_1 > 0$, *two vectors* $v_0, v_1 \in \mathbb{R}^n$, *and an invertible linear transformation* $B \in \mathrm{GL}_n$ *such that*

$$(\mathcal{F}\phi)(x) = C_0 + \langle v_1, x\rangle + C_1\phi(Bx + v_0).$$

Proof. Assume $\mathcal{F}$ is given, and consider the transformation $\mathcal{T} = \mathcal{F} \circ \mathcal{L} : Cvx(\mathbb{R}^n) \to Cvx(\mathbb{R}^n)$. It is invertible, order reversing, and also its inverse (which is $\mathcal{L} \circ \mathcal{F}^{-1}$) is order reversing, so we are in the conditions of Theorem 7. Therefore $\mathcal{T}$ is of the form

$$(\mathcal{T}\phi)(x) = C_0 + \langle v_1, x\rangle + C_1(\mathcal{L}\phi)(Bx + v_0).$$

To reconstruct $\mathcal{F}$, simply observe that $\mathcal{F} = \mathcal{T} \circ \mathcal{L}$, and also that

$$(\mathcal{T}\mathcal{L}\phi)(x) = C_0 + \langle v_1, x\rangle + C_1\phi(Bx + v_0). \qquad \square$$

It should be remarked here that one could begin with a direct proof of Corollary 8, about order-preserving transformations, and receive as a consequence Theorem 7 and the main theorem. This is another way to approach the subject, and the proofs in both approaches are quite similar.

Remark 9. The two conditions of Corollary 8 can be replaced by the following two conditions:

1. $\mathcal{F}(\hat{\min}(\phi, \psi)) = \hat{\min}(\mathcal{F}\phi, \mathcal{F}\psi)$,
2. $\mathcal{F}(\max(\phi, \psi)) = \max(\mathcal{F}\phi, \mathcal{F}\psi)$.

These two conditions are equivalent to the two conditions above, and should be compared with results of Gruber [Gr1], [Gr2] regarding the lattice of convex bodies and of normed spaces. However, we do have the additional assumption that $\mathcal{F}$ is invertible.

Similarly, in Theorem 7, we may replace the two conditions by the equivalent

1. $\mathcal{T}(\hat{\min}(\phi, \psi)) = \max(\mathcal{T}\phi, \mathcal{T}\psi)$,
2. $\mathcal{T}(\max(\phi, \psi)) = \hat{\min}(\mathcal{T}\phi, \mathcal{T}\psi)$.

Remark 10. It turns out that an even stronger theorem than Theorem 7 and Corollary 8 holds, when the condition of order reversion, or preservation, is replaced by the condition of 'comparability-preservation'. More precisely, if one merely assumes, concerning the invertible transform, that both it, and its inverse, map comparable pairs of functions in $Cvx(Rn)$, already it must be either order-reversing for all functions or order-preserving for all functions, and so must be either a Legendre-type transform as in Theorem 7, or an identity-type transform as in Corollary 8. For more details see the announcement paper [AM1].

6. Discussion and additional results

As explained in the introduction, Theorem 1, regarding the Legendre transform, is one (perhaps the main) example of a more general concept of duality, which can be applied not only to the class $Cvx(\mathbb{R}^n)$, but also to various other classes. The concept of duality is captured in the following definition, which, in all the examples we have considered, together with a specified class of functions, implies a concrete form of the duality transform, as can be seen in Theorem 1, and in the other theorems we quote below.

Definition 11 (The concept of duality). A transform $\mathcal{T}$ generates a *duality transform* on a set of functions $\mathcal{S}$ defined on $\mathbb{R}^n$ if the following two properties are satisfied:

1. For any $f \in \mathcal{S}$ we have $\mathcal{T}\mathcal{T}f = f$,
2. For any two functions in $\mathcal{S}$ satisfying $f \leq g$. we have $\mathcal{T}f \geq \mathcal{T}g$

First we would like to remark that, by the same reasoning as in Lemma 2, these two conditions imply that the regularized infimum is transformed to the regularized supremum and vice versa (provided that these regularizations are well defined in the class).

One important example, mentioned in the introduction, is that of polarity of compact convex bodies with zero in their interior (or of their indicator functions, as is more appropriate in out setting). For this class, Böröczky and Schneider showed in [BSc] a classification of transforms which satisfy the following condition: the transform of a convex hull of two convex sets, is the intersection of the transforms of the sets, and vice versa. As a conclusion from this, together with the equivalent of Lemma 2, one gets that an order-reversing involution defined on this class must be, up to a linear transformation, the usual polarity for convexity, given by

$$K^\circ = \{x : \sup_{y \in K} \langle x, y \rangle \leq 1\}.$$

Two related papers which study transformations which preserve the lattice of centrally symmetric (compact) convex bodies with respect to intersection of bodies and to the convex hull of their union, and also for the lattice of norms, are Gruber [Gr1] and [Gr2]. The latter implies, as was noticed in [BSc], the corresponding duality result for centrally symmetric convex bodies with 0 in their interior. For the case of closed convex sets with 0 inside but possibly on the boundary, see [AM2].

A direct application of Theorem 1 is a corresponding theorem for duality of log-concave functions. The class of log-concave functions, defined below, is studied in probability and in convex geometry, see [Bor], [Ba], and for more recent developments see e.g. [Kl], [M]. Denote by $\mathrm{LC}(\mathbb{R}^n)$ (for "log-concave") the class of upper semi-continuous non-negative functions with convex *non-empty* support, such that on their support, their logarithm is concave. The concept of duality for this class was studied in a joint paper of the authors and B. Klartag [AKM] where the following definition was given for the dual of a log-concave function:

$$f^\circ(x) = \inf_{y\in\mathbb{R}^n} \frac{e^{-\langle x,y\rangle}}{f(y)}. \tag{2}$$

Theorem 1 implies that this definition, up to linear terms, is the only definition corresponding to abstract duality, i.e., if one demands the duality operation to satisfy the abstract duality properties of Definition 11. Indeed:

Corollary 12. *Assume a transform $\mathcal{T} : \mathrm{LC}(\mathbb{R}^n) \to \mathrm{LC}(\mathbb{R}^n)$ (defined on the whole domain $\mathrm{LC}(\mathbb{R}^n)$) satisfies*

1. $\mathcal{T}\mathcal{T}f = f$,
2. $f \le g$ *implies* $\mathcal{T}f \ge \mathcal{T}g$.

Then, there exist a constant $0 < C_0 \in \mathbb{R}$, a vector $v_0 \in \mathbb{R}^n$, and an invertible symmetric linear transformation $B \in \mathrm{GL}_n$ such that $\mathcal{T}$ is defined as follows:

$$(\mathcal{T}f)(x) = C_0 e^{-\langle v_0,x\rangle} \inf_y \frac{e^{-\langle Bx+v_0,y\rangle}}{f(y)}.$$

(Notice that $(\mathcal{T}f)(x) = C_0 e^{-\langle v_0,x\rangle} f^\circ(Bx+v_0)$ for the functional duality f° defined above in (2).)

Other classes are of interest in geometry, and we quote just one additional result, namely a corresponding theorem for s-concave functions.

Define for $s > 0$ the set $\mathrm{Conc}_s(\mathbb{R}^n)$ to be the set of all upper semi-continuous non-negative functions on $\mathbb{R}^n$ which are s-concave, namely have convex support which includes 0, and $f^{1/s}$ is concave on the support. In [AKM] the following definition was given for duality in this class (and in several related classes):

$$\mathcal{L}_s f = \inf_{\{y: f(y)>0\}} \frac{\left(1 - \frac{\langle x,y\rangle}{s}\right)_+^s}{f(y)}. \tag{3}$$

It turns out (but requires an independent, and quite involved, proof) that again this transform is uniquely defined, up to some linear terms, if we require it to satisfy abstract duality as in Definition 11. The following theorem's proof, and several other variants of it, will appear in [AM2]:

THEOREM 13. *Assume* $n \geq 2$ *and we are given a transform* $\mathcal{T} : \mathrm{Conc}_s(\mathbb{R}^n) \to \mathrm{Conc}_s(\mathbb{R}^n)$ (*defined on the whole domain*) *satisfying*

1. $\mathcal{T}\mathcal{T}f = f$,
2. $f \leq g$ *implies* $\mathcal{T}f \geq \mathcal{T}g$.

Then, there exist a constant $C_0 \in \mathbb{R}$ *and an invertible symmetric linear transformation* $B \in \mathrm{GL}_n$ *such that*

$$(\mathcal{T}f)(x) = C_0 \inf_{(y: f(y)>0)} \frac{(1 - \langle x,y\rangle)_+^s}{f(By)}.$$

Finally, we describe one more related theorem. Before stating it, let us note that the Legendre transform has another remarkable property, *a priori* not connected with the fact that it reverses order. With the definition of the inf-convolution of two functions $f, g \in Cvx(\mathbb{R}^n)$, given by

$$(f \Box g)(z) = \inf_{x+y=z} \left(f(x) + g(y)\right),$$

the Legendre transform exchanges the sum of two functions with their inf-convoltuion, namely

$$\mathcal{L}f + \mathcal{L}g = \mathcal{L}(f \Box g)$$

for all $f, g \in Cvx(\mathbb{R}^n)$. (We need some convention to decide what to do with $-\infty + \infty$. We omit the justification for the following reasonable agreement: $(-\infty)\Box f \equiv -\infty$ for all f, and $(-\infty) + f = -\infty$ for all $f \not\equiv +\infty$, and, as functions, $-\infty + \infty \equiv +\infty$.)

Thus, the theorems above imply that any involutive transform on $Cvx(\mathbb{R}^n)$ which reverses order, essentially (after linear corrections) satisfies a relation of this form. The opposite is also true, and we can prove the following theorem (for details see the announcement paper [AM1])

THEOREM 14. *Assume that there is a transform* $\mathcal{T} : Cvx(\mathbb{R}^n) \to Cvx(\mathbb{R}^n)$ (*defined on the whole domain*) *satisfying*

1. $\mathcal{T}\mathcal{T}f = f$,
2. $\mathcal{T}f + \mathcal{T}g = \mathcal{T}(f \Box g)$.

Then, there exists a symmetric linear transformation $B \in \mathrm{GL}_n$ such that for all f

$$(\mathcal{T}f)(x) = \sup_{y\in\mathbb{R}^n} \left(\langle Bx, y\rangle - f(y)\right).$$

Tel Aviv University, Tel Aviv, Israel
E-mail addresses: shiri@post.tau.ac.il
milman@post.tau.ac.il

References

[Arn] V. I. Arnold, *Mathematical Methods of Classical Mechanics*, translated from Russian by K. Vogtmann and A. Weinstein, *Grad. Texts in Math.* **60**, Springer-Verlag, New York, 1978.

[Art] E. Artin, *Geometric Algebra*, *Interscience Tracts in Pure and Applied Mathematics*, Number 3, Interscience Publishers, Inc. New York, 1957.

[AKM] S. Artstein-Avidan, B. Klartag, and V. Milman, The Santaló point of a function, and a functional form of Santaló inequality, *Mathematika* **54** (2004), 33–48.

[AM1] S. Artstein-Avidan and V. Milman, A characterization of the concept of duality, *Electronic Research Anouncements in Mathematical Sciences*, *AIMS* **14** (2007), 48–65.

[AM2] S. Artstein-Avidan and V. Milman, The concept of duality for measure projections of convex bodies, *J. Funct. Ananl.* **254** (2008), 2648–2666.

[Ba] K. Ball, Logarithmically concave functions and sections of convex sets in $\mathbb{R}^n$, *Studia Math.* **88** (1988), 69–84.

[BGSS] A. Billot, I. Gilboa, D. Samet, and D. Schmeidler, Probabilities as similarily-weighted frequencies, *Econometrica* **73** (2005), 1125–1136.

[Bor] C. Borell, Convex measures on locally convex spaces, *Arikiv Math.* bf 12 (1974), 239–252.

[BSc] K. Böröczky and R. Schneider, A characterization of the duality mapping for convex bodies, *Geom. Funct. Anal* **18** (2008), 657–667.

[Gr1] P. Gruber, The endomorphisms of the lattice of convex bodies, *Abh. Math. Sem. Univ. Hamburg* **61** (1991), 121–130.

[Gr2] ———, The endomorphisms of the lattice of norms in finite dimensions, *Abh. Math. Sem. Univ. Hamburg* **62** (1992), 179–189.

[Ho] L. Hörmander, *Notions of Convexity*, *Progr. in Math.* **127**, Birkhäuser, Boston (1994).

[Kl] B. Klartag, Marginals of geometric inequalities, in *Geometric Aspects of Functional Analysis*, *Lecture Notes in Math.* **1910** (2007), 133–166, Springer-Verlag, New York.

[M] V. Milman, Geometrization of probability, in *Geometry and Dynamics of Groups and Spaces* (M. Kapranov, S. Kolyadi, Yu. Manin, and P. Moree, eds.), *Progr. in Math.* **265**, 569–588, Birkhauser-Verlag, Basel (2007).

[Ro] R. T. Rockafellar, *Convex Analysis*, *Princeton Landmarks in Mathematics and Physics*, Princeton Univ. Press, Princeton, NJ (1970).

(Received May 1, 2007)