Conjugating the inverse of a concave function.

J.-B. HIRIART-URRUTY
Université Paul Sabatier
118, route de Narbonne
31062 Toulouse cedex 4, France.
jbhu@cict.fr

Dedicated to C. LEMARECHAL on the occasion of his 60th birthday (April 1, 2004).

Abstract: This note is devoted to the clarification of the relationship between the LEGENDRE-FENCHEL conjugate of $\frac{1}{f}$ and that of $-f$ when $f$ is a positive concave function.

Keywords: Convex functions, LEGENDRE-FENCHEL conjugate, subdifferential.

2000 Mathematics Subject Classification: 26B25, 52A41, 90C25.

1 Introduction

The LEGENDRE-FENCHEL conjugate (or transform) of a function $f : X \to \mathbb{R} \cup \{+\infty\}$ is a function defined on the topological dual space $X^*$ of $X$ as

$$p \in X^* \mapsto f^*(p) := \sup_{x \in X} [(p, x) - f(x)]. \quad (1)$$

In convex analysis the conjugacy operation $f \leadsto f^*$ plays a central role, therefore a large body of calculus rules have been developed for it; they can
be found in any book on the subject. There however are some particular calculus rules which have been considered only recently, see for example ([5, 3]). The present note is devoted to clarifying the calculus rule giving the conjugate of $\frac{1}{f}$ in terms of that $-f$ when $f$ turns out to be a positive concave function. At first glance this situation can be viewed as a particular instance of a general calculus rule concerning a convex function post-composed with an increasing convex function ([2], Section 2.5 in chapter X): $x \mapsto (-f)(x)$ post-composed with $0 > y \mapsto \frac{-1}{y}$. We nevertheless provide a self-contained proof, insisting on the distinctive features of the resulting formula.

2 The conjugate of $\frac{1}{f}$.

The context of our work is the following one:

- $X$ is a (real) Banach space; by $X^*$ we denote the topological dual space of $X$, and $(p, x) \in X^* \times X \mapsto \langle p, x \rangle$ stands for the duality pairing.
- $f : X \to \mathbb{R} \cup \{-\infty\}$ is a concave upper-semicontinuous (or closed) function, strictly positive on $C := \{x \in X \mid f(x) > -\infty\}$ (assumed nonempty).

If we formulate this assumption in a way more familiar to practitioners of convex analysis or optimization, this gives: $-f : X \to \mathbb{R} \cup \{+\infty\}$ is a convex lower-semicontinuous (or closed) function, strictly negative on $\text{dom} (-f) = C$.

If $C$ turns out to be the whole of $X$, then $f$ is constant on $X$, so that this situation is not of much interest. In applications, $C$ happens to be a bounded (convex) set of $X$, on which $f$ is strictly positive.

The assumptions listed above are in force throughout the paper.

The inverse function of $f$, denoted as $\frac{1}{f}$, is defined on $X$ as follows:

$$
(\frac{1}{f})(x) := \begin{cases} 
\frac{1}{f(x)} & \text{if } x \in C \\
+\infty & \text{otherwise}
\end{cases}
$$

A classical and easily proved result is that $\frac{1}{f}$ is now convex on $X$, with domain $C$. 

2
The theorem below gives the expression of \( (\frac{1}{f})^\ast \) in terms of that of \((-f)^\ast\).

**Theorem 1.** For all \( p \in X^\ast \),

\[
(\frac{1}{f})^\ast(p) = \min \left\{ \left[ \alpha(-f)^\ast(\frac{p}{\alpha}) - 2\sqrt{\alpha} \right]_{\alpha > 0}, \sigma_C(p) \right\},
\]

where \( \sigma_C \) denotes the support function of \( C \). If \( p \) belongs to the cone generated by \( \text{dom}(-f)^\ast \) (i.e. if \( p \in \mathbb{R}^\ast_+ \text{dom}(-f)^\ast \)), then

\[
(\frac{1}{f})^\ast(p) = \inf_{\alpha > 0} \left[ \alpha(-f)^\ast(\frac{p}{\alpha}) - 2\sqrt{\alpha} \right].
\]

**Proof.** We mimic here the proof devised in [2] (p. 69 of volume II). By definition,

\[
-(\frac{1}{f})^\ast(p) = \inf_{x \in X} \left[ \frac{1}{f(x)} - \langle p, x \rangle \right] = \inf_{x \in C} \left[ -\frac{1}{f(x)} - \langle p, x \rangle \right]
\]

\[
= \inf_{x \in X, r > 0} \left[ -\frac{1}{r} - \langle p, x \rangle \mid (-f)(x) \leq r \right]
\]

(because \( r \mapsto -\frac{1}{r} \) is increasing on \((-\infty, 0))\).

Let us define

\[ f_1 : (x, r) \in X \times \mathbb{R} \mapsto f_1(x, r) := \begin{cases} -\langle p, x \rangle - \frac{1}{r} & \text{if } x \in X \text{ and } r < 0, \\ +\infty & \text{otherwise}; \end{cases} \]

\[ f_2 := i_{\text{epi}(-f)} \] (indicator function of the epigraph of \(-f\)).

Thus, (5) can be written as

\[
-(\frac{1}{f})^\ast(p) = \inf_{(x, r) \in X \times \mathbb{R}} [f_1(x, r) + f_2(x, r)].
\]

We then have to compute the conjugate of a sum of functions, however in a favorable context since \( \text{int}(\text{dom} f_1) = X \times (-\infty, 0) \) and \( \text{dom} f_2 = \text{epi}(-f) \) overlap. According to the classical Fenchel duality theorem

\[
(\frac{1}{f})^\ast(p) = \min_{(s, \alpha) \in X^\ast \times \mathbb{R}} [f_1^\ast(-s, \alpha) + f_2^\ast(s, -\alpha)].
\]

3
The computation of the above two conjugate functions is easy and gives:

\[ f_1^*(-s, \alpha) = -2\sqrt{\alpha} \text{ if } s = p \text{ and } \alpha \geq 0, \quad +\infty \text{ otherwise}; \]

\[ f_2^*(s, -\alpha) = \sigma_{\text{epi}(-f)}(s, -\alpha) = \begin{cases} 
\alpha(-f)^*(\frac{s}{\alpha}) & \text{if } \alpha > 0, \\
\sigma_C(s) & \text{if } \alpha = 0, \\
+\infty & \text{if } \alpha < 0.
\end{cases} \]

Plugging these results into (6) yields (3).

Observe that the function

\[ \alpha \in \mathbb{R} \mapsto \Theta(\alpha) := \begin{cases} 
\alpha(-f)^*(\frac{p}{\alpha}) - 2\sqrt{\alpha} & \text{if } \alpha > 0, \\
\sigma_C(p) & \text{if } \alpha = 0, \\
+\infty & \text{if } \alpha < 0
\end{cases} \]

is convex and lower-semicontinuous; its value at 0, that is \( \sigma_C(p) \), is the limit of \( \Theta(\alpha) \) when \( \alpha \in \text{dom } \Theta \to 0^+ \).

If \( p \) belongs to \( \mathbb{R}_+^* \text{dom } (-f)^* \), the domain of \( \Theta \) cannot reduce to \( \{0\} \), whence

\[ \min_{\alpha \geq 0} \Theta(\alpha) = \inf_{\alpha > 0} \Theta(\alpha), \]

that is to say (4). \( \square \)

**Remarks 2.**

- It may happen that \( p_0 \in \text{dom } (\frac{1}{f})^* \) but \( p_0 \notin \mathbb{R}_+^* \text{dom } (-f)^* \). In that case, (4) is invalid and the minimal value in (3) is achieved "at the limit \( \alpha_0 = 0 \)"; thus \( (\frac{1}{f})^*(p_0) = \sigma_C(p_0) \). See Example 5 below for an illustration of such a situation.

- According to formula (3), \( \text{dom } (\frac{1}{f})^* \) contains \( \mathbb{R}_+^* \text{dom } (-f)^* \). We will see later on that the closed (convex) cones generated by \( \text{dom } (-f)^* \) and \( \text{dom } (\frac{1}{f})^* \) are equal. Accordingly, \( p_0 \in \text{dom } (\frac{1}{f})^* \) belongs to \( \mathbb{R}_+^* \text{dom } (-f)^* \) "generically", i.e. at the exception of some "boundary-situations" such as that described above.

In view of the formulas (3) and (4) on \( (\frac{1}{f})^*(p) \), on may ask the following questions:
Could we delineate those $p$ for which the infimum in (4) is achieved for some $\alpha > 0$?
If so, is there any way of determining such $\alpha$ in terms of the given $p$?
Before tackling these questions, it is worthwhile to consider the next simple examples in order to grasp what can be expected and what not.

**Example 3.** Let $C = [-1, +1]$ and

$$f : x \in \mathbb{R} \mapsto f(x) := \begin{cases} 2 - |x| & \text{if } x \in C, \\ -\infty & \text{otherwise}. \end{cases}$$

Then, $(-f)^*$ and $(\frac{1}{f})^*$ are even functions with:

$$(-f)^*(p) = \begin{cases} 2 & \text{if } 0 \leq p \leq 1, \\ p + 1 & \text{if } p \geq 1; \end{cases}$$

$$(\frac{1}{f})^*(p) = \begin{cases} -1/2 & \text{if } 0 \leq p \leq 1/4, \\ 2(p - \sqrt{p}) & \text{if } 1/4 \leq p \leq 1, \\ p - 1 & \text{if } p \geq 1. \end{cases}$$

Let, for instance, $\frac{1}{4} \leq p_0 \leq 1$. The minimum value in the right-hand side of the formula

$$(\frac{1}{f})^*(p_0) = \min \left\{ \alpha(-f)^*(\frac{p_0}{\alpha}) - 2\sqrt{\alpha}, \ p_0 \right\}$$

5
is achieved for $\alpha_0 = p_0$.

**Example 4.** Let $C = [0, +\infty)$ and

$$f : x \in \mathbb{R} \mapsto f(x) := \begin{cases} 2 - e^{-x} & \text{if } x \geq 0, \\ -\infty & \text{otherwise.} \end{cases}$$

Then, an easy calculation leads to

$$(-f)^*(p) = \begin{cases} 1 & \text{if } p \leq 1, \\ -p \log(-p) + p + 2 & \text{if } -1 \leq p \leq 0, \\ +\infty & \text{if } p \geq 0; \end{cases}$$

while the explicit expression of $\left(\frac{1}{f}\right)^*(p)$ is fairly complicated (see however Figure 2 for a sketch of its graph). Let $p_0 = 0$. Then the minimum value in the right-hand side of the formula

$$\left(\frac{1}{f}\right)^*(0) = \min \left\{ [2\alpha - 2\sqrt{\alpha}]_\alpha > 0, 0 \right\}$$

is achieved for $\alpha_0 = \frac{1}{4}$. 
This is a general rule. Suppose $M := \sup_{x \in C} f(x) < +\infty$ (as it is the case in Example 3 and the present one). Then the minimum value in the right-hand side of the formula expressing $\left(\frac{1}{f}\right)^*(0)$ is achieved at $\alpha_0 = \frac{1}{M^2}$; the corresponding value is $-\frac{1}{M}$, as expected.

**Example 5.** Let $C = [0, +\infty)$ and

$$f : x \in \mathbb{R} \mapsto f(x) := \begin{cases} x + 1 & \text{if } x \geq 0, \\ -\infty & \text{otherwise}. \end{cases}$$

Then:

$$(-f)^*(p) = \begin{cases} 1 & \text{if } p \leq -1, \\ +\infty & \text{if } p > -1; \end{cases}$$

$$\left(\frac{1}{f}\right)^*(p) = \begin{cases} -1 & \text{if } p \leq -1, \\ -p - 2\sqrt{-p} & \text{if } -1 \leq p \leq 0, \\ +\infty & \text{if } p > 0. \end{cases}$$

Let $p_0 = 0$. Here $p_0 \in \text{dom} \left(\frac{1}{f}\right)^*$ but $p_0 \notin \mathbb{R}^+_* \text{dom} (-f)^*$. This is an example where

$$\alpha(-f)^*(\frac{p_0}{\alpha}) - 2\sqrt{\alpha} = +\infty \text{ for all } \alpha > 0,$$

whence the minimum value in the formula (3) yielding $\left(\frac{1}{f}\right)^*(p_0)$ is achieved "at the limit $\alpha_0 = 0" \text{ and has the value } \sigma_C(p_0) = 0.
To answer the questions posed as an introduction to the examples listed above, we need to explore furthermore the relationship between \( \text{dom} \left( \frac{1}{f} \right)^* \) and \( \text{dom} \left( -f \right)^* \) by establishing a link between the subdifferential of \( \frac{1}{f} \) that of \(-f\). The connecting formula, as expected, is the one given in the statement below.

**Proposition 6.** For all \( x \in C \),

\[
\partial \left( \frac{1}{f} \right)(x) = \frac{\partial (-f)(x)}{[f(x)]^2}.
\]

**Proof.** We proceed to compare the directional derivates \( \left( \frac{1}{f} \right)'(x, d) \) and \( (-f)'(x, d) \) for all \( d \in X \).

**First case** \( x + td \notin C \) for all \( t > 0 \).

In that case, both \( \left( \frac{1}{f} \right)'(x, d) \) and \( (f)'(x, d) \) equal \(+\infty\).

**Second case** \( x + td \in C \) for some \( t > 0 \).

Thus, the line-segment \([x, x + td]\) is entirely contained in \( C \); therefore it comes from the assumption made on \( f \) (an upper-semicontinuous concave function) that the function (of the real variable) \( t \mapsto f(x + td) \) is continuous on \([0, t]\). We infer from that,

\[
\frac{1}{f(x + td)} - \frac{1}{f(x)} = \frac{(-f)(x + td) - (-f)(x)}{t} \frac{t}{f(x + td)f(x)} \rightarrow \lim_{t \to 0^+} \frac{(-f)(x + td) - (-f)(x)}{t},
\]

whence

\[
\left( \frac{1}{f} \right)'(x, d) = \frac{(-f)'(x, d)}{[f(x)]^2}.
\]

In summary, the equality (8) holds for all \( d \in X \). Since we have for any convex function \( \varphi \) on \( X \)

\[
\partial \varphi(x) = \left\{ p \in X^* \mid \langle p, d \rangle \leq \varphi'(x, d) \text{ for all } d \in X \right\},
\]

the announced relationship (7) readily follows from (8). \(\square\)
Not all the \( p \) in \( \text{dom} \left( \frac{1}{f} \right)^* \) are in \( \text{Im} \partial \left( \frac{1}{f} \right) \); however we have

\[
\text{Im} \partial \left( \frac{1}{f} \right) \subset \text{dom} \left( \frac{1}{f} \right)^* \subset \text{Im} \partial \left( \frac{1}{f} \right)
\]

(9)

(the second inclusion follows from the approximation theorem of BRØNDSTED-ROCKAFELLAR (1965)). For those \( p \) which are in \( \text{Im} \partial \left( \frac{1}{f} \right) \), we are able to provide \( \alpha > 0 \) at which the infimum is achieved in the formula (4).

**Theorem 7.** Let \( p_0 \in \text{Im} \partial \left( \frac{1}{f} \right) \), and consider \( x_0 \) such that \( p_0 \in \partial \left( \frac{1}{f} \right)(x_0) \). Then,

\[
\left( \frac{1}{f} \right)^*(p_0) = \frac{(-f)^* \{ \langle f(x_0) \rangle^2 p_0 \}}{|f(x_0)|^2} - \frac{2}{f(x_0)}. \tag{10}
\]

In other words: we are in a situation where \( p_0 \) belongs to the convex cone \( \mathbb{R}_+^* \text{dom} \left( f \right)^* \), and the minimizer in the right-hand side of (4) is \( \alpha_0 = \frac{1}{|f(x_0)|^2} \).

**Proof.** We have \( p_0 \in \partial \left( \frac{1}{f} \right)(x_0) \) and, according to (7), \( |f(x_0)|^2 p_0 \in \partial (-f)(x_0) \). Using the characterization of the subdifferential of \( \varphi \) in terms of its conjugate \( (s_0 \in \partial \varphi(x_0) \text{ if and only if } \varphi^*(s_0) + \varphi(x_0) - \langle s_0, x_0 \rangle = 0) \), we have:

\[
(-f)^* \{ \langle f(x_0) \rangle^2 p_0 \} - f(x_0) + \langle f(x_0), p_0, x_0 \rangle = 0, \tag{11}
\]

\[
\left( \frac{1}{f} \right)^*(p_0) + \frac{1}{f(x_0)} - \langle p_0, x_0 \rangle = 0. \tag{12}
\]

Dividing (11) by \( |f(x_0)|^2 \) and comparing the resulting equality to (12), we derive (10).

When \( p_0 \in \text{Im} \partial \left( \frac{1}{f} \right) \), we clearly are in a situation where \( p_0 \in \mathbb{R}_+^* \text{dom} \left( -f \right)^* \) since (according to (7)) \( p_0 \in \mathbb{R}_+^* \text{Im} \partial (-f) \subset \mathbb{R}_+^* \text{dom} \left( -f \right)^* \). Then the (strictly) convex function

\[
\alpha > 0 \mapsto \alpha (-f)^* \left( \frac{p_0}{\alpha} \right) - 2 \sqrt{\alpha}
\]
is minimized at \( \alpha_0 = \frac{1}{[f(x_0)]^2} \).

\[
\square
\]

**Remarks 8.** - Even if \( p_0 \notin \operatorname{Im} \partial \left( \frac{1}{f} \right) \), it may happen that the infimum in the right-hand side of (4) is achieved at some \( \alpha_0 > 0 \), but such \( \alpha_0 \) is not necessarily \( \frac{1}{[f(x_0)]^2} \) for some \( x_0 \in C \). Indeed, consider again Example 4 and \( p_0 = 0 \). We note that

\[
p_0 \notin \operatorname{Im} \partial \left( \frac{1}{f} \right), \quad p_0 \in \mathbb{R}^*_+ \operatorname{dom} (-f)^*,
\]

\[
(\frac{1}{f})^*(p_0) = \inf_{\alpha > 0} [\alpha(-f)^* \left( \frac{p_0}{\alpha} \right) - 2\sqrt{\alpha}]
\]

\[
= [\alpha(-f)^* \left( \frac{p_0}{\alpha_0} \right) - 2\sqrt{\alpha_0}] \quad \text{for} \quad \alpha_0 = \frac{1}{4},
\]

but there is no \( x_0 \in C \) such that \( \frac{1}{4} = \frac{1}{[f(x_0)]^2} \) (such an \( x_0 \) is "rejected at the infinity on \( C'' \)).

- We have:

\[
\operatorname{dom} \left( \frac{1}{f} \right)^* = \operatorname{Im} \partial \left( \frac{1}{f} \right), \quad \operatorname{dom} (-f)^* = \operatorname{Im} \partial (-f)
\]

(see the comments about (9));

\[
\mathbb{R}^*_+ \operatorname{Im} \partial (-f) = \mathbb{R}^*_+ \operatorname{Im} \partial \left( \frac{1}{f} \right)
\]

(this results from (7)).

Combining (13) and (14) gives rise to the following relationship between \( \operatorname{dom} \left( \frac{1}{f} \right)^* \) and \( \operatorname{dom} (-f)^* \):

\[
\mathbb{R}^*_+ \operatorname{dom} \left( \frac{1}{f} \right)^* = \mathbb{R}^*_+ \operatorname{dom} (-f)^*
\]

- There are several possible situations where \( \operatorname{Im} \partial \left( \frac{1}{f} \right) = \operatorname{dom} \left( \frac{1}{f} \right)^* \); one of them is when \( X \) is reflexive and \( C \) is bounded. Indeed, in that case, \( \left( \frac{1}{f} \right)^* \)
is continuous throughout $X^*$ and $Im \partial \left( \frac{1}{f} \right) = dom \left( \frac{1}{f} \right)^* = X^*$ ([4], Corollary 7G); thus formula (10) holds true at any $p_0 \in X^*$.

The expression (10) for \( \left( \frac{1}{f} \right)^*(p) \), more comfortable and easier to handle than (3) (provided one can solve the equation $p \in \partial \left( \frac{1}{f} \right)(x)$), would then allow us to pursue further the study of possible relations between the mathematical objects (from the viewpoint of convex analysis) associated with the convex functions $-f$ and $\frac{1}{f}$.

References


