Journal of Convex Analysis
Volume 19 (2012), No. 2, 569–574

A Short Derivation of the Conjugate of a Supremum Function

Stephen M. Robinson*
University of Wisconsin-Madison, Department of Industrial and Systems Engineering, 1513 University Avenue, Madison, WI 53706-1539, USA
smrobins@wisc.edu

Dedicated to Claude Lemaréchal, with appreciation for his fundamental contributions to practical convex optimization.

Received: August 24, 2010
Revised manuscript received: October 09, 2010

We give a very short derivation of the convex conjugate of the function whose value is the pointwise supremum of a finite collection of proper convex functions.

Keywords: Pointwise supremum, convex conjugate

2000 Mathematics Subject Classification: 52A41, 90C25

1. Introduction

This note deals with the problem of representing the conjugate of the function taking a point $x \in \mathbb{R}^n$ to the supremum of the values at x of a finite collection of proper convex functions $f_1, \ldots, f_k$ from $\mathbb{R}^n$ to the extended real line $\bar{\mathbb{R}}$. Treatments in standard texts, such as [7, Theorem 16.5] or [6, Theorems 2.4.4, 2.4.7], use the convex-hull operator on the conjugates of the functions in question. However, this method requires taking the closure of the resulting convex-hull representation.

Attempts to remove this closure operation have led to various devices, such as assuming that the effective domains of the f_i have a common closure [7, p. 149] or, more restrictively, that all of these effective domains equal $\mathbb{R}^n$ [6, p. 68]. These devices limit the applicability of the result, but they are necessary in the framework adopted in these treatments.

In the more general setting of infinite-dimensional spaces, several authors have given formulas for the conjugate of the pointwise supremum of two convex functions. Fitzpatrick and Simons [3] gave such a formula in their Equation (2.1), and they

*This material is based on work supported in part by the Air Force Research Laboratory under awards FA9550-07-1-0389 and FA9550-10-1-0101, and in part by the National Science Foundation under Grant DMS-0807893. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author and do not necessarily reflect the views of the sponsoring agencies.

ISSN 0944-6532 / $ 2.50 © Heldermann Verlag

also credited a second such formula, in their Equation (2.3), to a paper of Traoré and Volle [9, Section 7]. The first of these formulas contains a limiting operation, and the second is of min-sup type. Boţ and Wanka [1] studied this problem, citing the paper of Fitzpatrick and Simons as well as the book of Zălinescu [10], in which two applicable results appear in Corollaries 2.8.11 (for k functions) and 2.8.12 (for two functions). In addition, Hantaute and López [4] and then Hantaute, López, and Zălinescu [5] have studied the subdifferential of the supremum function, as have Combari, Laghdir, and Thibault [2, Theorem 5.12].

We show here that in $\mathbb{R}^n$ one can avoid the closure operation by computing the conjugate using a functional form that is similar to, but not the same as, the convex-hull operation. The only required regularity condition is the standard qualification that the relative interiors of the effective domains of the functions f_k have a common point. Of the results cited above, the formula we obtain seems closest to the result of Traoré and Volle and to Corollary 2.8.12 of Zălinescu, but it accommodates any finite number of functions and its proof (Theorem 3.2) is very short. This brevity results from a connection with epigraphs also employed by Combari *et al.* [2, Proposition 5.11].

The following section explains a redefinition of certain functional operations that we will use in computing the conjugate, and it also establishes some notational conventions. Section 3 develops the representation of the conjugate and applies the result to an illustrative example from [3].

2. Redefining left and right multiplication

In what follows we use for the most part the notation of Rockafellar [7], but with a few important exceptions. The indicator function of a subset C of $\mathbb{R}^n$, evaluated at x, is $I_C(x)$: this has the value 0 if $x \in C$ and $+\infty$ otherwise. Its conjugate, the support function of C, is I^*_C. Following [8, Section 1.H], we also use # instead of $\Box$ to denote the operation of infimal convolution (epi-addition) that is dual to ordinary addition of functions.

A more substantial change is the redefinition of the functional operations λf and $f\lambda$ that Rockafellar defines in [7, pp. 34–35]. The following definitions depart from those of Rockafellar only when $\lambda = 0$. As in [7], $f0^+$ denotes the recession function of f.

Definition 2.1. Let $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ be a proper convex function. For $\lambda \in [0, +\infty)$ and $x \in \mathbb{R}^n$,

$$(\lambda f)(x) := \begin{cases} \lambda f(x) & \text{if } \lambda > 0, \\ I_{\mathrm{cl\,dom}\, f}(x) & \text{if } \lambda = 0, \end{cases} \tag{1}$$

and

$$(f\lambda)(x) := \begin{cases} \lambda f(\lambda^{-1}x) & \text{if } \lambda > 0, \\ f0^+(x) & \text{if } \lambda = 0. \end{cases} \tag{2}$$

When f is closed, the redefined functions of Definition 2.1 obey the same conjugacy relations in the variable x as do those originally defined in [7]: that is, $(\lambda f)^* = (f^*\lambda)$

and $(f\lambda)^* = (\lambda f^*)$.

If we pass from treating the functional operations λf and $f\lambda$ as functions of x, with f and λ held fixed, to fixing only f and considering these operations as functions of (x, λ), then the operation λf is generally not convex (e.g., take $f(x) = x$). However, with $f\lambda$ the situation is different as we will shortly see. The transition from considering these operations as functions of x to considering them as functions of (x, λ) also reveals an advantage of the functions in Definition 2.1 over the original definitions: namely, the former are lower semicontinuous as functions of (x, λ) while the latter generally are not. The proof of lower semicontinuity for (λf) is straightforward and we omit it. Traoré and Volle [9, p. 145] used the definition in (2) for $f\lambda$, but their definition for λf differs from (1) and is in general not lower semicontinuous even when f is closed.

The proof that if f is a closed proper convex function then $(f\lambda)(x)$ is a closed proper *convex* function of (x, λ) is an immediate consequence of the following proposition, by taking f and λ to be the f^* and ξ^* in the proposition. However, that proposition gives additional information that we will use in the next section: namely, that the function of (x^*, ξ^*) that equals $(f^*\xi^*)(x^*)$ for $\xi^* \in [0, +\infty)$ and $+\infty$ for $\xi^* < 0$ is, up to a sign change, the support function of the epigraph of f. This is a slight variation on [7, Cor. 13.5.1], but we give a direct proof because it is simpler to do so than to adapt the original proof.

Proposition 2.2. *If $f : \mathbb{R}^n \to \bar{\mathbb{R}}$ is a proper convex function, then*

$$I^*_{\operatorname{epi} f}(x^*, -\xi^*) = \begin{cases} (f^*\xi^*)(x^*) & \text{if } \xi^* \in [0, +\infty), \\ +\infty & \text{if } \xi^* < 0. \end{cases}$$

Proof. We have

$$\begin{aligned} I^*_{\operatorname{epi} f}(x^*, -\xi^*) &= \sup_{x,\xi} \{\langle x^*, x\rangle + \langle -\xi^*, \xi\rangle - I_{\operatorname{epi} f}(x, \xi)\} \\ &= \sup_{x,\xi} \{\langle x^*, x\rangle + \langle -\xi^*, \xi\rangle \mid f(x) \le \xi\}. \end{aligned}$$

If $\xi^* < 0$ then the supremum is $+\infty$ because we can take ξ to be as large as we please, whereas if $\xi^* = 0$ then the supremum is $I^*_{\operatorname{dom} f}(x^*)$, which is $(f^*0)(x^*)$. Finally, if $\xi^* > 0$ then for each $x \in \operatorname{dom} f$ we can make $\langle x^*, x\rangle + \langle -\xi^*, \xi\rangle$ as large as possible by choosing $\xi = f(x)$. We then have

$$\begin{aligned} I^*_{\operatorname{epi} f}(x^*, -\xi^*) &= \sup_{x} \{\langle x^*, x\rangle - \xi^* f(x)\} \\ &= \xi^* \sup_{x} \{\big\langle (\xi^*)^{-1} x^*, x\big\rangle - f(x)\} \\ &= \xi^* f^*[(\xi^*)^{-1} x^*] \\ &= (f^*\xi^*)(x^*). \end{aligned}$$

□

3. Computing the conjugate

Given a finite collection $f_1, \dots, f_K$ of proper convex functions from $\mathbb{R}^n$ to $\bar{\mathbb{R}}$, we will compute the conjugate of the function $\sup_{k=1}^K f_k$ under the condition that $\cap_{k=1}^K \operatorname{ri} \operatorname{dom} f_k \neq \emptyset$. As we will use the epigraphs of the f_k in the proof, we need to relate this condition to a condition on the relative interiors of the epigraphs.

Lemma 3.1. *Let $f_1, \dots, f_K$ be convex functions on $\mathbb{R}^n$ and define $\Pi : \mathbb{R}^{n+1} \to \mathbb{R}^n$ by $\Pi(x_1, \dots, x_{n+1}) = (x_1, \dots, x_n)$. Then*

$$\cap_{k=1}^K \operatorname{ri} \operatorname{dom} f_k = \Pi(\cap_{k=1}^K \operatorname{ri} \operatorname{epi} f_k). \tag{3}$$

Proof. For each k we have $\operatorname{dom} f_k = \Pi(\operatorname{epi} f_k)$. Properties of the relative interior then imply that

$$\cap_{k=1}^K \operatorname{ri} \operatorname{dom} f_k = \cap_{k=1}^K \operatorname{ri} \Pi(\operatorname{epi} f_k) = \cap_{k=1}^K \Pi(\operatorname{ri} \operatorname{epi} f_k). \tag{4}$$

It is immediate that $\Pi(\cap_{k=1}^K \operatorname{ri} \operatorname{epi} f_k) \subset \cap_{k=1}^K \Pi(\operatorname{ri} \operatorname{epi} f_k)$. On the other hand, if $x \in \cap_{k=1}^K \Pi(\operatorname{ri} \operatorname{epi} f_k)$ then for each k there is some real ξ_k with $(x, \xi_k) \in \operatorname{ri} \operatorname{epi} f_k$. As K is finite we can set $\xi = \max_{k=1}^K \xi_k$; then for each k, $(x, \xi) \in \operatorname{ri} \operatorname{epi} f_k$. Therefore $x \in \Pi(\cap_{k=1}^K \operatorname{ri} \operatorname{epi} f_k)$, so that $\cap_{k=1}^K \Pi(\operatorname{ri} \operatorname{epi} f_k) = \Pi(\cap_{k=1}^K \operatorname{ri} \operatorname{epi} f_k)$. Combining this with (4) yields (3). □

Here is the main result. It represents the value of the conjugate of the supremum as a linearly-constrained minimum of a sum of K closed proper convex functions that are explicitly computable if the f_k^* are known.

Theorem 3.2. *Let $f_1, \dots, f_K$ be proper convex functions from $\mathbb{R}^n$ to $\bar{\mathbb{R}}$. If*

$$\cap_{k=1}^K \operatorname{ri} \operatorname{dom} f_k \neq \emptyset, \tag{5}$$

then for each $x^* \in \mathbb{R}^n$,

$$\left(\sup_{k=1}^K f_k \right)^* (x^*) = \inf \left\{ \sum_{k=1}^K (f_k^* \lambda_k^*)(x_k^*) \mid \sum_{k=1}^K x_k^* = x^*, \ \sum_{k=1}^K \lambda_k^* = 1, \ \lambda_k^* \geq 0 \right\}. \tag{6}$$

Further, for each x^* there are $x_1^*, \dots, x_K^*$ and $\lambda_1^*, \dots, \lambda_K^*$ that attain the infimum on the right in (6) (possibly at $+\infty$).

Proof. First, rewrite the left side of (6) in terms of epigraphs:

$$\begin{aligned}
\left(\sup_{k=1}^K f_k \right)^* (x^*) &= \sup_{x,\xi} \left\{ \left\langle \begin{pmatrix} x^* \\ -1 \end{pmatrix}, \begin{pmatrix} x \\ \xi \end{pmatrix} \right\rangle \mid \begin{pmatrix} x \\ \xi \end{pmatrix} \in \cap_{k=1}^K \operatorname{epi} f_k \right\} \\
&= \sup_{x,\xi} \left\{ \left\langle \begin{pmatrix} x^* \\ -1 \end{pmatrix}, \begin{pmatrix} x \\ \xi \end{pmatrix} \right\rangle - \sum_{k=1}^K I_{\operatorname{epi} f_k} \begin{pmatrix} x \\ \xi \end{pmatrix} \right\} \\
&= \left(\sum_{k=1}^K I_{\operatorname{epi} f_k} \right)^* \begin{pmatrix} x^* \\ -1 \end{pmatrix}.
\end{aligned}$$

The effective domains of the functions $I_{\operatorname{epi} f_k}$ are the sets $\operatorname{epi} f_k$. By combining the hypothesis in (5) with Lemma 3.1 we find that $\cap_{k=1}^K \operatorname{ri} \operatorname{epi} f_k \neq \emptyset$. Recalling the duality between addition and infimal convolution [7, Theorem 16.4], we conclude that

$$\begin{aligned}\left(\sum_{k=1}^K I_{\operatorname{epi} f_k}\right)^* \begin{pmatrix} x^* \\ -1 \end{pmatrix} &= \left(\#_{k=1}^K I^*_{\operatorname{epi} f_k}\right) \begin{pmatrix} x^* \\ -1 \end{pmatrix} \\ &= \inf\left\{ \sum_{k=1}^K I^*_{\operatorname{epi} f_k} \begin{pmatrix} x_k^* \\ -\lambda_k^* \end{pmatrix} \mid \sum_{k=1}^K \begin{pmatrix} x_k^* \\ -\lambda_k^* \end{pmatrix} = \begin{pmatrix} x^* \\ -1 \end{pmatrix} \right\} \\ &= \inf\left\{ \sum_{k=1}^K (f_k^* \lambda_k^*)(x_k^*) \mid \sum_{k=1}^K x_k^* = x^*,\ \sum_{k=1}^K \lambda_k^* = 1,\ \lambda_k^* \geq 0 \right\},\end{aligned}$$

where we used Proposition 2.2 to derive the last equality. The duality theorem also ensures the existence of x_k^* and λ_k^* attaining the infimum. □

We demonstrate an application of Theorem 3.2 by using it to calculate the conjugate of a supremum function given by Fitzpatrick and Simons [3, Remark 3], who showed that the convex-hull operator (without the closure operation) fails to give the correct value for the conjugate at $x^* = 0$.

Example 3.3. Take $n = 2$ and $k = 2$, with

$$f_1(x) = I^*_{\mathbb{R}_- \times \{1\}}(x) = \begin{cases} x_2 & \text{if } x_1 \geq 0 \\ +\infty & \text{otherwise,} \end{cases}$$

and

$$f_2(x) = I^*_{\{(1,0)\}}(x) = x_1.$$

Then

$$f(x) := \sup\{f_1(x), f_2(x)\} = \begin{cases} \max\{x_1, x_2\} & \text{if } x_1 \geq 0, \\ +\infty & \text{otherwise.} \end{cases}$$

We have $\operatorname{dom} f_1 = \mathbb{R}_+ \times \mathbb{R}$ and $\operatorname{dom} f_2 = \mathbb{R}^2$, so the regularity condition in (5) holds.

As f is closed, convex and positively homogeneous it is the support function of some closed convex set C, and then $f^* = I_C$. Thus, to determine f^* we need only find C.

A calculation shows that for nonnegative λ_1^* and λ_2^* one has

$$f_1^* \lambda_1^* = I_{\mathbb{R}_- \times \{\lambda_1^*\}}, \qquad f_2^* \lambda_2^* = I_{\{(\lambda_2^*, 0)\}}.$$

For nonnegative λ_1^* and λ_2^* that sum to 1, $f_1^* \lambda_1^*(x_1^*) + f_2^* \lambda_2^*(x_2^*)$ attains its minimum of zero if and only if $x_1^* = (\nu, \lambda_1^*)$ and $x_2^* = (\lambda_2^*, 0)$, where $\nu \leq 0$. A point x^* will thus lie in C if and only if it has the form

$$x^* = x_1^* + x_2^* = (\nu + \alpha, 1 - \alpha) \ \text{ for } \alpha \in [0, 1] \text{ and } \nu \leq 0.$$

Hence C is an unbounded trapezoid equal to the sum of the halfline $\mathbb{R}_- \times \{0\}$ and the line segment whose endpoints are $(1, 0)$ and $(0, 1)$.

Acknowledgements. I thank an anonymous referee for helpful comments including additional references.

References

[1] R. I. Boţ, G. Wanka: The conjugate of the pointwise maximum of two convex functions revisited, J. Glob. Optim. 41 (2008) 625–632.

[2] C. Combari, M. Laghdir, L. Thibault: Sous-différentiels de fonctions convexes composées, Ann. Sci. Math. Qué. 18(2) (1994) 119–148.

[3] S. P. Fitzpatrick, S. Simons: On the pointwise maximum of convex functions, Proc. Amer. Math. Soc. 128 (2000) 3553–3561.

[4] A. Hantoute, M. A. López: A complete characterization of the subdifferential set of the supremum of an arbitrary family of convex functions, J. Convex Analysis 15 (2008) 831–858.

[5] A. Hantoute, M. A. López, C. Zălinescu: Subdifferential calculus rules in convex analysis: a unifying approach via pointwise supremum functions, SIAM J. Optim. 19 (2008) 863–882.

[6] J.-B. Hiriart-Urruty, C. Lemaréchal: Convex Analysis and Minimization Algorithms II, Grundlehren der mathematischen Wissenschaften 306, Springer, Berlin (1993).

[7] R. T. Rockafellar: Convex Analysis, Princeton University Press, Princeton (1970).

[8] R. T. Rockafellar, R. J.-B. Wets: Variational Analysis, Grundlehren der mathematischen Wissenschaften 317, Springer, Berlin (1998).

[9] S. Traoré, M. Volle: On the level sum of two convex functions on Banach spaces, J. Convex Analysis 3 (1996) 141–151.

[10] C. Zălinescu: Convex Analysis in General Vector Spaces, World Scientific, Singapore (2002).