King Fahd University of Petroleum & Minerals

DEPARTMENT OF MATHEMATICAL SCIENCES

Technical Report Series

TR 165

October 1994

Smoothing a Nondifferentiable Convex Function: The Technique of the Rolling Ball

Seeger, Alberto

DHAHRAN 31261 ● SAUDI ARABIA ● www.kfupm.edu.sa/math/ ● E-mail: mathdept@kfupm.edu.sa

SMOOTHING A NONDIFFERENTIABLE CONVEX FUNCTION: THE TECHNIQUE OF THE ROLLING BALL

Alberto SEEGER

King Fahd University of Petroleum and Minerals
Department of Mathematical Sciences
Dhahran 31261, Saudi Arabia

Abstract

A general formalism is introduced in order to discuss the regularization of a nondifferentiable convex function. Special attention is paid to the regularization technique which consists in rolling a ball under the epigraph of the function. We show that under suitable assumptions, the regularized function is a solution of the so–called mixed type variational PDE.

Key words: lower boundary function, infimal–convolution, regularizing kernel, subdifferential.

1991 Mathematics subject classification: 52 A 07, 52 A 41.

1 Introduction.

This paper is concerned with the smoothing or regularization of a nondifferentiable convex function $f : X \to R \cup \{+\infty\}$. Given a set A in the product space $X \times R$, we introduce the concept of A–regularization of f. This concept refers to a new function, denoted f_A, which enjoys some nice differentiability properties, whenever the set A is chosen in a suitable way.

The A–regularization formalism includes the Moreau–Yosida regularization and the Lipschitz regularization as particular cases. It also includes a smoothing technique which consists in rolling a ball under the epigraph of f. This latter technique is mentioned in a recent paper by Noll [N1], but has not yet been explored in its whole extension. The technique of the rolling ball corresponds to the particular choice $A = rB$, where B is a given ball in the product space $X \times R$, and $r \in R_+$ is a parameter which serves to measure the quality of the regularization process. In this work it is shown that the quantity

$$F(x,r) := f_{rB}(x)$$

enjoys some noteworthy differentiability properties, both as a function of the argument $x \in X$ and of the parameter $r \in R$.

The organization of the paper is as follows: In section 2 we record some preliminary results on the notion of lower boundary function. The general A–regularization formalism is introduced and discussed in section 3. In section 4 we explore in detail the technique of the rolling ball.

2 Preliminary Results on Lower Boundary Functions.

Unless otherwise specified, (X, X^*) will be a couple of locally convex topological linear spaces in duality by means of a bilinear form $\langle \cdot, \cdot \rangle : X^* \times X \to R$ (see [B2, p. 48]). Given a nonempty closed set A in the product space $X \times R$, we identify its *lower boundary* with the extended real-valued function

$$u \in X \mapsto k_A(u) := \inf_{\alpha \in R} \{\alpha + \Psi_A(u, \alpha)\}, \tag{2.1}$$

where Ψ_A stands for the indicator function of A, i.e.

$$\Psi_A(u, \alpha) := \begin{cases} 0 & \text{if } (u, \alpha) \in A, \\ +\infty & \text{otherwise} . \end{cases}$$

Strictly speaking, the lower boundary of A corresponds to the graph of k_A.

In what follows we assume that the set A is proper in the sense that it is nonempty and

$$k_A(u) > -\infty \quad \text{for all } u \in X.$$

Observe that the function k_A may take the value $+\infty$. This occurs at those points $u \in X$ for which the set

$$A_u := \{\alpha \in R : \quad (u, \alpha) \in A\}$$

is empty. Actually one has

$$\begin{aligned} \operatorname{dom} k_A : &= \{u \in X : \quad k_A(u) < +\infty\} \\ &= \{u \in X : \quad A_u \text{ is nonempty}\}. \end{aligned}$$

The Legendre–Fenchel conjugate of k_A is the function $k_A^* : X^* \to R \cup \{+\infty\}$ given by

$$k_A^*(y) := \sup_{u \in X}\{\langle y, u\rangle - k_A(u)\} \quad \text{for all } y \in X^*.$$

As a matter of computation one obtains the formula

$$k_A^*(y) = \Psi_A^*(y, -1) \quad \text{for all } y \in X^*, \tag{2.2}$$

where

$$(y, \beta) \in X^* \times R \mapsto \Psi_A^*(y, \beta) := \sup_{(u,\alpha)\in A} \{\langle y, u\rangle + \beta\alpha\}$$

stands for the support function of the set A.

If the set A is convex, then so is the function k_A (cf. [R1, Theorem 5.3]). In such a case one uses the symbol

$$\partial k_A(u) := \{y \in X^* : k_A^*(y) + k_A(u) - \langle y, u\rangle = 0\} \tag{2.3}$$

to denote the *subdifferential* of k_A at the point $u \in X$. By using Rockafellar's calculus rule on the subdifferential of a marginal function (cf, [R2, Theorem 24]), one obtains

$$\partial k_A(u) = \{y \in X^* : (y, -1) \in N_A\,[u, k_A(u)]\}, \tag{2.4}$$

for all $u \in \text{dom}\, k_A$. In the above formula the symbol

$$\begin{aligned} N_A[u, \alpha] &:= \partial\Psi_A(u, \alpha) \\ &= \{(y, \beta) \in X^* \times R : \langle (y, \beta), (z, \gamma) - (u, \alpha)\rangle \leq 0 \quad \text{for all } (z, \gamma) \in A\} \end{aligned}$$

denotes the *normal cone* to A at the point $(u, \alpha) \in A$.

3 Smoothing a Nondifferentiable Convex Function.

In this section, we assume that $f : X \to R \cup \{+\infty\}$ is a proper convex function which is not necessarily differentiable. Such class of functions will be denoted by

Conv(X). From a geometric point of view, the lack of differentiability of the function f is reflected by "corners" in its epigraph

$$\text{epi } f := \{(x, \lambda) \in X \times R : \quad f(x) \leq \lambda\}.$$

These corners can be removed by adding to the above set a suitable perturbation $A \subset X \times R$. The basic idea behind the next definition is that the Minkowski sum epi $f + A$ inherits somehow the good properties of the perturbation A.

Definition 3.1 Let $f \in$ Conv(X) and A be a proper closed set in $X \times R$. The A–regularization of f is the function $f_A : X \to \overline{R}$ whose graph is the lower boundary of epi $f + A$, i.e.

$$f_A(x) := \inf_{\lambda \in R} \{\lambda + \Psi_{\text{epi } f+A}(x, \lambda)\} \quad \text{for all } x \in X. \tag{3.1}$$

There are several equivalent ways of characterizing the function f_A. The following one is rather technical, but it has some advantages from a calculus point of view.

Proposition 3.1 *The A–regularization of f is given by*

$$f_A(x) = \inf_{(u,\alpha) \in A} \{f(x - u) + \alpha\} \quad \textit{for all } x \in X. \tag{3.2}$$

Proof. It suffices to write

$$\begin{aligned} f_A(x) &= \inf\{\lambda \in R : (x, \lambda) \in \text{epi } f + A\} \\ &= \inf\{\lambda \in R : (x, \lambda) - (u, \alpha) \in \text{ epi } f \quad \text{for some } (u, \alpha) \in A\} \\ &= \inf\{\lambda \in R : \ f(x - u) + \alpha \leq \lambda \quad \text{for some } (u, \alpha) \in A\}. \end{aligned}$$

From the above line one gets easily the formula (3.2). □

The next result has the merit of displaying the infimal–convolutive nature of the A–regularization technique. Recall that the infimal–convolution $f_1 \Box f_2$ of two proper functions $f_1, f_2 : X \to R \cup \{+\infty\}$ is defined by

$$\begin{aligned} [f_1 \Box f_2](x) &:= \inf\{f_1(x_1) + f_2(x_2) : x_1 + x_2 = x\} \\ &= \inf_{u \in X}\{f_1(x-u) + f_2(u)\}. \end{aligned}$$

Proposition 3.2 *The A–regularization of f is the infimal–convolution of f and the lower boundary function k_A, i.e.,*

$$f_A(x) = \inf_{u \in X}\{f(x-u) + k_A(u)\} \quad \textit{for all } x \in X. \tag{3.3}$$

Proof. According to Proposition 3.1 one can write

$$\begin{aligned} f_A(x) &= \inf_{(u,\alpha) \in X \times R}\{f(x-u) + \alpha + \Psi_A(u,\alpha)\} \\ &= \inf_{u \in X}\{f(x-u) + \inf_{\alpha \in R}\{\alpha + \Psi_A(u,\alpha)\}\}. \end{aligned}$$

□

Proposition 3.2 shows that only the lower boundary of A is important in the definition of f_A. In other words, the function f_A does not change if one modifies the portion of A which lies above the graph of k_A. The function k_A plays the role of a "regularizing kernel". The latter concept is not new to those readers who are familiar with the following two examples.

Example 3.1 Consider a normed space $(X, \|\cdot\|)$. If one takes $A = \operatorname{epi}\left(\frac{1}{2r}\|\cdot\|^2\right)$, then

$$x \in X \mapsto f_A(x) = \inf_{u \in X}\left\{f(u) + \frac{1}{2r}\|x-u\|^2\right\}$$

is just the Moreau–Yosida regularization of f of index r (cf. [A1]).

Example 3.2 Let $(X, \|\cdot\|)$ be a normed space. If one chooses $A = \text{epi}(r\|\cdot\|)$, then

$$x \in X \mapsto f_A(x) = \inf_{u \in X} \{f(u) + r\|x - u\|\}$$

corresponds to the Lipschitz regularization of f of index r (cf. [B1]).

Another important choice for the set A is given in the example below. A detailed analysis of this case will be carried out in section 4.

Example 3.3 Let $(X, \|\cdot\|)$ be a normed space, and let

$$B = \{(u, \alpha) \in X \times R : \ \|u\|^2 + \alpha^2 \leq 1\}$$

denote the closed ball in the product space $X \times R$. If $A = rB$ is the closed ball of radius $r \in R_+$, then the lower boundary of A is given by

$$k_A(u) = \begin{cases} -[r^2 - \|u\|^2]^{1/2} & \text{if } \|u\| \leq r, \\ +\infty & \text{otherwise} \end{cases}.$$

According to Proposition 3.2, one has

$$f_A(x) = \inf_{\|u\| \leq r} \{f(x - u) - [r^2 - \|u\|^2]^{1/2}\}.$$

Next we record some general properties of the function f_A. It helps to keep in mind that f_A is a perturbation of the function f. From this point of view, one may see the optimization problem

$$(P)_A \qquad\qquad \text{Minimize } \ \{f_A(x) : \ x \in X\}$$

as a perturbed version of

$$(P) \qquad\qquad \text{Minimize } \{f(x): \ x \in X\}.$$

The corresponding optimal–values $\inf_X f_A$ and $\inf_X f$ are related as follows:

Proposition 3.3 *Let the symbol*

$$m_A := \inf\{\alpha : (u,\alpha) \in A\} = \inf_X k_A$$

denote the smallest α–coordinate of the set $A \subset X \times R$. Then,

$$\inf_X f_A = \inf_X f + m_A. \tag{3.4}$$

Proof. According to Proposition 3.2, one has

$$\begin{aligned}
\inf_X f_A &= \inf_{x\in X} \inf_{u\in X} \{f(x-u) + k_A(u)\} \\
&= \inf_{u\in X} \{k_A(u) + \inf_{x\in X} f(x-u)\} \\
&= \inf_{u\in X} k_A(u) + \inf_X f \\
&= m_A + \inf_X f.
\end{aligned}$$

□

Definition 3.2 If the constant m_A is finite, then the normalized A–regularization of f is the function $\tilde{f}_A$ given by

$$\tilde{f}_A(x) := f_A(x) - m_A \quad \text{for all } x \in X.$$

The above normalization procedure is simply a technical device for obtaining a perturbation of f which preserves the infimal value of this function. One clearly has

$$\inf_X \tilde{f}_A = \inf_X f. \tag{3.5}$$

Next we derive a formula for computing the Legendre–Fenchel conjugate

$$y \in X^* \mapsto f_A^*(y) := \sup_{x \in X}\{\langle y, x\rangle - f_A(x)\}$$

of the function f_A.

Proposition 3.4 *For all $y \in X^*$, one has*

$$f_A^*(y) = f^*(y) + \Psi_A^*(y, -1). \tag{3.6}$$

Proof. By applying Proposition 3.2 one gets

$$f_A^* = [f \Box k_A]^* = f^* + k_A^* \quad \text{(cf. [L1])}.$$

Now it suffices to apply the formula (2.2). □.

Remark: Proposition 3.4 may be seen as generalization of Proposition 3.3. Indeed, the formula (3.4) is obtained by setting $y = 0$ in (3.6).

From now on we will consider mainly the case in which the set $A \subset X \times R$ is convex. The motivation behind this choice lies in the following trivial result:

Proposition 3.5 *If the set $A \subset X \times R$ is convex, then so is the function f_A.*

Proof. f_A is the infimal–convolution of the convex functions f and k_A. □

The first–order behavior of the convex function f_A around the point $x \in X$ is reflected by the set

$$\partial f_A(x) := \{y \in X^* : f_A^*(y) + f_A(x) - \langle y, x\rangle = 0\}. \tag{3.7}$$

The above set corresponds to the subdifferential of f_A at x (cf. [R1]).

The purpose of the next theorem is to derive a formula for computing this set. First, we need to recall two basic definitions related to the concept of subdifferentiability. Given a number $\epsilon \in R_+$, one defines the *ϵ-subdifferential* of $f : X \to R \cup \{+\infty\}$ at the point $x \in X$ as the set

$$\partial_\epsilon f(x) := \{y \in X^* : f^*(y) + f(x) - \langle y, x\rangle \leq \epsilon\}.$$

The set of *ϵ-normal directions* to $A \subset X \times R$ at the point $(u, \alpha) \in A$ is by definition

$$\begin{aligned} N_A^\epsilon[u,\alpha] &:= \partial_\epsilon \Psi_A(u,\alpha) \\ &= \{(y,\beta) \in X^* \times R : \langle (y,\beta), (z,\gamma) - (u,\alpha)\rangle \leq \epsilon \quad \text{for all} \quad (z,\gamma) \in A\}. \end{aligned}$$

For additional information on the above two notions, the reader may consult, for instance, Hiriart–Urruty [H1].

By using the characterization of f_A given in Proposition 3.1, we are able to obtain the following result.

Theorem 3.1 *Let $f \in \mathrm{Conv}(X)$ and $A \subset X \times R$ be a proper closed convex set. Suppose there exists at least one point in X at which f is continuous. If f_A is finite at $x \in X$, then*

$$\partial f_A(x) = \bigcap_{\epsilon>0} \bigcup_{(u,\alpha)\in A} \{y \in \partial_\epsilon f(x-u) : (y,-1) \in N_A^\epsilon[u,\alpha]\}. \tag{3.8}$$

Proof. According to Proposition 3.1, one has

$$f_A(x) = \inf_{(u,\alpha)\in X\times R} \{H_1(x,u,\alpha) + H_2(x,u,\alpha) + H_3(x,u,\alpha)\},$$

where

$$H_1(x,u,\alpha) = f(x-u),$$
$$H_2(x,u,\alpha) = \alpha,$$
$$H_3(x,u,\alpha) = \Psi_A(u,\alpha).$$

The fact that f is continuous at some point in X implies that

$$(H_1+H_2+H_3)^* = H_1^* \Box H_2^* \Box H_3^*.$$

The latter condition is a constraint qualification hypothesis which allows us to apply the formula (cf. Seeger [S1, Lemma 3])

$$\partial f_A(x) = \bigcap_{\epsilon>0} \bigcup_{(u,\alpha)\in X\times R} \{y \in X^* : (y,0,0) \in \partial_\epsilon H_1(x,u,\alpha) + \partial_\epsilon H_2(x,u,\alpha) + \partial_\epsilon H_3(x,u,\alpha)\}.$$

Now, as a matter of computation, one gets

$$\partial_\epsilon H_1(x,u,\alpha) = \{(v,-v,0) \in X^* \times X^* \times R : \ v \in \partial_\epsilon f(x-u)\},$$
$$\partial_\epsilon H_2(x,u,\alpha) = \{(0,0,1)\} \subset X^* \times X^* \times R,$$
$$\partial_\epsilon H_3(x,u,\alpha) = \{(0,v,\beta) \in X^* \times X^* \times R : \ (v,\beta) \in N_A^\epsilon[u,\alpha]\}.$$

Putting all the pieces together one obtains the formula (3.8). □

Formula (3.8) is quite general in the sense that it holds even if the infimum in (3.2) is not attained. A simpler formula for $\partial f_A(x)$ can be obtained if one assumes that the minimization problem (3.2) has a solution. It is easy to see that the latter condition is equivalent to the attainment of the infimum in the expression (3.3). Denote by

$$S(x) := \{u \in X : f(x-u) + k_A(u) = f_A(x)\}$$

the set of points at which the infimal-convolution (3.3) is attained.

Theorem 3.2 *Let A and f be as in Theorem 3.1. Suppose that the solution set $S(x)$ is nonempty. Then for any $u \in S(x)$, one can write*

$$\partial f_A(x) = \{y \in \partial f(x-u) : (y,-1) \in N_A[u, k_A(u)]\}. \tag{3.9}$$

Proof. A classical theorem on the subdifferential of an infimal–convolution (cf. [L1]) yields

$$\partial f_A(x) = \partial f(x-u) \cap \partial k_A(u). \tag{3.10}$$

It suffices now to apply the formula (2.4). □.

Recall that a closed convex set A is smooth at $(u, k_A(u))$ if the normal cone $N_A[u, k_A(u)]$ is a ray. This amounts to saying that $\partial k_A(u)$ contains a single element, namely the Gâteaux–differential $k'_A(u)$. One says that A has a smooth lower boundary if A is smooth at any point of the form $(u, k_A(u))$.

Corollary 3.1 . *Let A and f be as in Theorem 3.1. Suppose that A has a smooth lower boundary. Then f_A is Gâteaux–differentiable at any point $x \in X$ such that $S(x)$ is nonempty. Moreover,*

$$f'_A(x) = k'_A(u) \in \partial f(x-u), \tag{3.11}$$

where u is any element in $S(x)$.

Proof. It suffices to apply the formula (3.10). □

4 The Technique of the Rolling Ball.

In this section, X is a reflexive Banach space equipped with the norm $\|\cdot\|$, X^* is its topological dual space, and $\langle\cdot,\cdot\rangle$ refers to the duality product between X^* and X.

Let

$$B := \left\{(u,\alpha) \in X \times R : \|u\|^2 + \alpha^2 \le 1\right\}$$

denote the closed unit ball in the product space $X \times R$.

As in Section 3, we consider again a proper convex function $f : X \to R \cup \{+\infty\}$. We intend to regularize this function by using a set $A \subset X \times R$ which has a special form, namely

$$A = rB = \left\{(u,\alpha) \in X \times R : \|u\|^2 + \alpha^2 \le r^2\right\}.$$

Here $r \in R_+$ is a parameter which has been introduced to measure the quality of the regularization. The purpose of this section is to study in detail the function f_{rB}, that is to say, the rB–regularization of f. The ball B is regarded as fixed, and thus the parameter r plays the major role in our analysis. For notational convenience we write

$$F(x,r) := f_{rB}(x) \quad \text{for all} \quad r \in R_+ \quad \text{and} \quad x \in X.$$

By definition, we have

$$F(x,r) = \inf_{\lambda \in R} \left\{\lambda + \Psi_{\text{epi}\, f + rB}(x,\lambda)\right\}, \tag{4.1}$$

that is to say, $F(\cdot,r)$ is the function whose graph is the lower boundary of epi $f + rB$. Such a function $F(\cdot,r)$ has been already introduced in the area of differential geometry, although a systematic study of it has not been yet undertaken. Noll [N1] pointed out that the set epi $f + rB$ is obtained by rolling a ball of diameter r under the epigraph of f.

Example 4.1 Let $f : R \to R$ be given by $f(x) = |x|$. Then,

$$F(x,r) = \begin{cases} |x| - r\sqrt{2} & \text{if } |x| \ge r/\sqrt{2}, \\ -\sqrt{r^2 - |x|^2} & \text{if } |x| \le r/\sqrt{2}. \end{cases}$$

According to Proposition 3.1, the function $F(\cdot, r)$ admits the representation

$$F(x,r) = \inf\{f(x-u) + \alpha : \|u\|^2 + \alpha^2 \leq r^2\}. \tag{4.2}$$

As we have seen already in Example 3.3, another way of expressing the function $F(\cdot, r)$ is by means of an infimal-convolution, namely

$$F(x,r) = \inf_{\|u\| \leq r} \{f(x-u) - [r^2 - \|u\|^2]^{1/2}\}. \tag{4.3}$$

In the proposition below we list some trivial facts about the behavior of F.

Proposition 4.1 *The function F enjoys the following properties:*

(a) For all $x \in X$, $r \in R_+ \mapsto F(x,r)$ is nonincreasing;

(b) For all $x \in X$,

$$f(x) = F(x,0) = \lim_{r \to 0+} F(x,r) = \sup_{r>0} F(x,r);$$

(c) For all $x \in X$ and $r \in R_+$,

$$-\infty < F(x,r) \leq f(x) - r.$$

Proof. The proof of parts (a) and (b) is straightforward. The upper bound for $F(x,r)$, given in part (c), is obtained by setting $u = 0$ in the right hand side of (4.3). That F has never the value of $-\infty$ can be shown as follows. Let $y \in X^*$ be any point at which f^* is finite. From the Young–Fenchel inequality

$$f^*(y) + f(x) - \langle y, x \rangle \geq 0,$$

one obtains

$$f(x-u) \geq \langle y, x \rangle - f^*(y) - \langle y, u \rangle.$$

So, when $\|u\| \leq r$, one can write

$$f(x-u) - [r^2 - \|u\|^2]^{1/2} \geq \langle y, x \rangle - f^*(y) - \langle y, u \rangle - [r^2 - \|u\|^2]^{1/2}.$$

Hence

$$F(x,r) \geq \langle y,x\rangle - f^*(y) - \sup_{\|u\|\leq r} \{\langle y,u\rangle + [r^2 - \|u\|^2]^{1/2}\}.$$

It suffices now to observe that the above supremum is finite. □

As an immediate consequence of Proposition 3.3, one has the following result.

Proposition 4.2 *For all $r \in R_+$, one has*

$$\inf_{x\in X} F(x,r) = \inf_{x\in X} f(x) - r. \tag{4.4}$$

Proof. It suffices to take $A = rB$ in the formula (3.4). Observe that

$$\begin{aligned} m_{rB} &= \inf\{\alpha : (u,\alpha) \in rB\} \\ &= \inf\{\alpha : \|u\|^2 + \alpha^2 \leq r^2\} \\ &= \inf\{-[r^2 - \|u\|^2]^{1/2} : \|u\| \leq r\} \\ &= -r. \end{aligned}$$

□

As a consequence of Proposition 3.5 we see that, for each $r \in R_+$, the function $F(\cdot,r)$ is convex. In fact, one has a stronger result:

Proposition 4.3 *The function F is convex over the set $X \times R_+$.*

Proof. According to the expression (4.2), one has

$$F(x,r) = \inf_{(u,\alpha)\in X\times R}\{f(x-u) + \alpha + \Psi_{rB}(u,\alpha)\},$$

or equivalently

$$F(x,r) = \inf_{(u,\alpha)\in X\times R}\{f(x-u) + \alpha + \Psi_K(r,u,\alpha)\}, \tag{4.5}$$

where

$$K := \{(r, u, \alpha) \in R \times X \times R : \ r \geq 0, \quad (u, \alpha) \in rB\}. \tag{4.6}$$

It is not difficult to check that K is a nonempty convex set in the space $R \times X \times R$. This fact implies the convexity of the function $H : X \times R_+ \times X \times R \to R \cup \{+\infty\}$ defined by

$$H(x, r, u, \alpha) := f(x - u) + \alpha + \Psi_K(r, u, \alpha).$$

To show the convexity of F it suffices now to take into account the representation (4.5). □

Proposition 4.3 says that the function F is not only convex with respect to its first argument, but also with respect to both arguments simultaneously. The function F can be extended in a convex manner to the whole space $X \times R$ by setting simply

$$F(x, r) = \infty \quad \text{if} \quad r < 0 \quad \text{and} \quad x \in X.$$

Remark: Proposition 4.1(c) implies that $F : X \times R \to R \cup \{+\infty\}$ is a proper function. A fairly simple calculation shows that the effective domain of F is given by

$$\text{dom}\, F = \{(x, r) \in X \times R_+ : x \in \text{dom}\, f + rB_X\},$$

where $B_X := \{u \in X : \|u\| \leq 1\}$ is the closed unit ball in X.

In the next proposition we derive a formula for the Legendre–Fenchel conjugate of F. The notation $\|\cdot\|_*$ stands for the dual norm in X^* associated to the norm $\|\cdot\|$ in X, i.e.,

$$\|y\|_* := \sup_{\|x\| \leq 1} \langle y, x \rangle \quad \text{for all} \quad y \in X^*.$$

Proposition 4.4 *The Legendre–Fenchel conjugate $F^* : X^* \times R \to R \cup \{+\infty\}$ of F is given by*

$$F^*(y, s) = \begin{cases} f^*(y) & \text{if} \ \ [1 + \|y\|_*^2]^{1/2} \leq -s, \\ +\infty & \text{otherwise} . \end{cases} \tag{4.7}$$

In particular,

$$\text{dom}F^* = \{(y,s) \in X^* \times R : \; y \in \text{dom}f^* \quad \textit{and} \quad [1 + \|y\|_*^2]^{1/2} \le -s\}. \tag{4.8}$$

Proof. For all $(y,s) \in X^* \times R$ one has

$$\begin{aligned} F^*(y,s) &= \sup_{\substack{x\in X \\ r\in R}} \left\{ \langle y,x\rangle + sr - \inf_{\substack{u\in X \\ \alpha\in R}} \{f(x-u) + \alpha + \Psi_K(r,u,\alpha)\} \right\} \\ &= \sup_{x,r} \sup_{u,\alpha} \{\langle y,x\rangle + sr - f(x-u) - \alpha - \Psi_K(r,u,\alpha)\} \\ &= \sup_{x,u,\alpha} \{\langle y,x\rangle - f(x-u) - \alpha + \sup_r \{sr - \Psi_K(r,u,\alpha)\}\} \\ &= f^*(y) + \sup_{u,\alpha} \{\langle y,u\rangle - \alpha + \sup_r\{sr - \Psi_K(r,u,\alpha)\}\}. \end{aligned}$$

But

$$\begin{aligned} \sup_r \{sr - \Psi_K(r,u,\alpha)\} &= \sup_{r\ge 0} \{sr - \Psi_{rB}(u,\alpha)\} \\ &= \begin{cases} s[\alpha^2 + \|u\|^2]^{1/2} & \text{if} \quad s \le 0 \\ +\infty & \text{if} \quad s > 0. \end{cases} \end{aligned}$$

Hence, $F^*(y,s) = +\infty$ whenever $s \ge 0$. In the case $s < 0$, one has

$$F^*(y,s) = f^*(y) + \sup_{u,\alpha} \{\langle y,u\rangle - \alpha + s[\alpha^2 + \|u\|^2]^{1/2}\}.$$

The above supremum is the Legendre–Fenchel conjugate at $(y,-1)$ of the function

$$(\alpha,u) \in R \times X \to -s[\alpha^2 + \|u\|^2]^{1/2}.$$

As a matter of calculus one shows that this supremum is equal to 0 if $[1+\|y\|_*^2]^{1/2} \le -s$, and equal to $+\infty$ otherwise. □

In Proposition 4.4, the conjugation is taken with respect to both variables of the function F. Below we look rather at the conjugation of the partial function $F(\cdot, r)$.

Proposition 4.5 *For all $r \in R_+$, the Legendre–Fenchel conjugate of $F(\cdot, r)$ is given by*

$$[F(\cdot,r)]^*(y) = f^*(y) + r[1 + \|y\|_*^2]^{1/2} \quad \textit{for all} \quad y \in X^*. \tag{4.9}$$

Proof. It follows from Proposition 3.4. □

Remark. Formula (4.9) appears already in a paper by Noll [N1, p. 527]. He assumes that $X = R^n$ is a finite-dimensional space and $\|\cdot\|$ is the usual Euclidean norm.

For the sake of completeness, we also consider the conjugation of the partial function $F(x, \cdot)$.

Proposition 4.6 *For all $x \in X$, the Legendre–Fenchel conjugate of $F(x, \cdot)$ is given by*

$$[F(x,\cdot)]^*(s) = \begin{cases} -\inf_{u\in X}\{f(x-u) + [s^2-1]^{1/2}\|u\|\} & \textit{if } s \le -1, \\ +\infty & \textit{if } s > -1 \end{cases}$$

Proof. By definition, one has

$$\begin{aligned} [F(x,\cdot)]^*(s) &:= \sup_{r\in R}\{sr - F(x,r)\} \\ &= \sup_{r\ge 0}\{sr - \inf_{\|u\|\le r}\{f(x-u) - [r^2 - \|u\|^2]^{1/2}\}\}. \end{aligned}$$

By rearranging the above expression, one obtains

$$\begin{aligned} [F(x,\cdot)]^*(s) &= \sup_{r\ge 0}\sup_{\|u\|\le r}\{sr - f(x-u) + [r^2 - \|u\|^2]^{1/2}\} \\ &= \sup_{u\in X}\{K(u,s) - f(x-u)\}, \end{aligned}$$

where

$$K(u,s) := \sup_{r\in R}\{sr + [r^2 - \|u\|^2]^{1/2} : \; r \ge \|u\|\}.$$

But, as a matter of calculus,

$$K(u,s) = \begin{cases} -[s^2-1]^{1/2}\,\|u\| & \text{if } s \le -1, \\ +\infty & \text{if } s > -1. \end{cases}$$

This completes the proof of the proposition. □

The following results are related to the first–order sensitivity analysis of the convex function F. It is our intention to characterize the subdifferential

$$\partial F(x,r) = \{(y,s) \in X^* \times R : \; F^*(y,s) + F(x,r) - \langle y, x\rangle - sr = 0\}$$

of F at a given point $(x,r) \in \mathrm{dom}F$. In what follows we use the symbol

$$S(x,r) := \{u \in X : \|u\| \le r, \; f(x-u) - [r^2 - \|u\|^2]^{1/2} = F(x,r)\} \tag{4.10}$$

to denote the set of points at which the infimal–convolution (4.3) is attained. For convenience we also use the notation $k_r := k_{rB}$. As we have seen already in Example 3.3, the regularizing kernel k_r is given by

$$k_r(u) = \begin{cases} -[r^2 - \|u\|^2]^{1/2} & \text{if } \|u\| \le r\,, \\ +\infty & \text{otherwise }. \end{cases} \tag{4.11}$$

Theorem 4.1 *Let $r > 0$ and f be as in Theorem 3.1. Let $u \in X$ be in the solution set $S(x,r)$. Then, $(y,s) \in X^* \times R$ is an element of $\partial F(x,r)$ if and only if the following three conditions hold:*

$$[1 + \|y\|_*^2]^{1/2} \le -s, \tag{4.12}$$

$$\epsilon := \langle y, u\rangle + sr - k_r(u) \ge 0, \tag{4.13}$$

$$y \in \partial_\epsilon f(x-u). \tag{4.14}$$

Proof. From the Proposition 4.4 it follows that

$$(y,s) \in \partial F(x,r) \Leftrightarrow \begin{cases} \text{(a)} \quad [1+\|y\|_*^2]^{1/2} \le -s, \quad \text{and} \\ \text{(b)} \quad f^*(y) + F(x,r) - \langle y,x\rangle - sr = 0. \end{cases}$$

So, we need to show that (b) is equivalent to the conditions (4.13) and (4.14) together. Since $u \in S(x,r)$, the condition (b) can be written in the form

$$f^*(y) + f(x-u) + k_r(u) - \langle y,x\rangle - sr = 0,$$

or equivalently

$$f^*(y) + f(x-u) - \langle y, x-u\rangle = \langle y,u\rangle + sr - k_r(u). \tag{4.15}$$

To complete the proof it suffices to take into account the Young–Fenchel inequality and the definition of the ϵ–subdifferential mapping $\partial_\epsilon f$ as given in Section 3. □

It is known that the subdifferential of a bivariate convex function can be estimated in terms of the subdifferentials of the partial functions. Recall that the partial subdifferential of F with respect to the variable x is the set

$$\begin{aligned} \partial_x F(x,r) &:= \partial[F(\cdot,r)](x) \\ &= \{y \in X^* : [F(\cdot,r)]^*(y) + F(x,r) - \langle y,x\rangle = 0\} \\ &= \{y \in X^* : F(z,r) \ge F(x,r) + \langle y, z-x\rangle \quad \text{for all} \quad z \in X\}. \end{aligned}$$

In a similar way one defines the partial subdifferential of F with respect to the variable r, namely

$$\begin{aligned} \partial_r F(x,r) &:= \partial[F(x,\cdot)](r) \\ &= \{s \in R : [F(x,\cdot)]^*(s) + F(x,r) - sr = 0\} \\ &= \{s \in R : F(x,t) \ge F(x,r) + s(t-r) \quad \text{for all} \quad t \in R\}. \end{aligned}$$

According to [R3, p. 72] one has always the inclusion

$$\partial F(x,r) \subset \partial_x F(x,r) \times \partial_r F(x,r). \tag{4.16}$$

The study of the partial subdifferential $\partial_x F(x,r)$ is an interesting matter by itself. This set reflects the first–order behavior of the regularized function $F(\cdot,r)$.

Proposition 4.7 *Let $r > 0$ and f be as in Theorem 3.1. Suppose that the solution set $S(x,r)$ is nonempty. Then, for any $u \in S(x,r)$, one can write*

$$\partial_x F(x,r) = \partial f(x-u) \cap \partial k_r(u). \tag{4.17}$$

Proof. It follows from Theorem 3.2. □

With the help of the above proposition it is possible to obtain a characterization of $\partial F(x,r)$ which is much simpler than the one given in Theorem 4.1.

Theorem 4.2 *Let $r > 0$ and f be as in Theorem 3.1. Let $u \in S(x,r)$. Then, $(y,s) \in \partial F(x,r)$ if and only if the following three conditions hold:*

$$[1 + \|y\|_*^2]^{1/2} \leq -s, \tag{4.18}$$

$$\langle y, u\rangle + sr - k_r(u) = 0, \tag{4.19}$$

$$y \in \partial f(x-u). \tag{4.20}$$

Proof. By inspecting again the proof of Theorem 4.1, it should be clear that the conditions (4.18) – (4.20) imply that $(y,s) \in \partial F(x,r)$. To prove the reverse implication, take any $(y,s) \in \partial F(x,r)$. According to the inclusion (4.16), the component y belongs to $\partial_x F(x,r)$. Due to Proposition 4.7, it follows that $y \in \partial f(x-u)$. This condition says that the term on the left hand side of (4.15) is equal to zero. So does the term on the right hand side. □

Let us come back again to the discussion of the partial subdifferential $\partial_x F(x,r)$. Formula (4.17) shows that ∂k_r plays an important role in connection with the sensitivity analysis of $F(\cdot,r)$. A Gâteaux–differentiable kernel k_r can be obtained with a suitable choice of the norm $\|\cdot\|$. This choice also yields the Gâteaux–differentiability of the regularized function $F(\cdot,r)$. This fact will be stated in a more precise way in the next proposition. For convenience in our exposition, we record first the following lemma.

Lemma 4.1 *Let $r>0$ and f be as in Theorem 3.1. Consider the following three statements:*

(a) $u \in S(x,r)$;

(b) $\partial k_r(u)$ is nonempty;

(c) $\|u\| < r$.

Then, one has the relationship

$$(a) \Rightarrow (b) \Rightarrow (c).$$

Proof. (a) $\Rightarrow$ (b). If the infimal convolution $f \Box k_r$, given by (4.3), is attained at $u \in X$, then necessarily

$$\partial f(x-u) \cap \partial k_r(u) \neq \phi \quad \text{(cf. [L1])}.$$

Thus, $\partial k_r(u)$ is nonempty whenever $u \in S(x,r)$.

(b) $\Rightarrow$ (c). Let y be an arbitrary element in $\partial k_r(u)$. By definition one has

$$k_r^*(y) + k_r(u) - \langle y, u \rangle = 0,$$

or equivalently,

$$\begin{cases} \|u\| \le r, \quad \text{and} \\ r[1+\|y\|_*^2]^{1/2} - [r^2 - \|u\|^2]^{1/2} - \langle y, u \rangle = 0. \end{cases}$$

We need to prove that $\|u\| < r$. On the contrary, suppose that $\|u\| = r$. Then

$$r[1+\|y\|_*^2]^{1/2} - \langle y, u \rangle = 0,$$

or equivalently

$$[1+\|y\|_*^2]^{1/2} = \left\langle y, \frac{1}{r}u \right\rangle.$$

This clearly contradicts the fact that

$$\left\langle y, \frac{1}{r}u \right\rangle \le \|y\|_*.$$

□

Proposition 4.8 *Besides the assumptions of Proposition 4.7, suppose that the function*

$$u \in X \mapsto J(u) := \frac{1}{2}\|u\|^2$$

is Gâteaux–differentiable. Then,

(a) k_r is Gâteaux–differentiable on $\{u \in X : \|u\| < r\}$, and

$$k_r'(u) = [r^2 - \|u\|^2]^{-1/2} \ J'(u). \tag{4.21}$$

(b) $F(\cdot, r)$ is Gâteaux–differentiable at the point $x \in X$. Moreover

$$\frac{\partial F}{\partial x}(x, r) = k_r'(u) \in \partial f(x-u), \tag{4.22}$$

where $\frac{\partial F}{\partial x}(x,r)$ denotes the Gâteaux–differential of $F(\cdot, r)$ at $x \in X$, and u is any element in the set $S(x, r)$.

Proof. Part (a) follows from the expression (4.11) of k_r. To prove the part (b) it suffices to use the formula (4.17). □

Remark. The subdifferential mapping $\partial J : X \to X^*$ is usually referred to as the duality mapping between the spaces X and X^*. Gâteaux–differentiability of J amounts to single–valuedness of ∂J. It is known that J is Gâteaux–differentiable if X^* is strictly convex (see, for instance, Zeidler [Z1, Proposition 47.18]).

In the next proposition we address the question of the differentiability of $F(x,r)$ as a function of the parameter r.

Proposition 4.9 *Under the assumptions of Proposition 4.8, one can assert that the derivative of $F(x,\cdot)$ at r exists. Moreover, it is given by*

$$\frac{\partial F}{\partial r}(x,r) = r^{-1}\left[k_r(u) - \left\langle \frac{\partial F}{\partial x}(x,r), u\right\rangle\right] = -r[r^2 - \|u\|^2]^{-1/2}. \tag{4.23}$$

Proof. By combining Proposition 4.8 and the inclusion (4.16), we see that

$$\partial F(x,r) \subset \left\{\frac{\partial F}{\partial x}(x,r)\right\} \times \partial_r F(x,r).$$

Take any element (y,s) in $\partial F(x,r)$. Then necessarily

$$y = \frac{\partial F}{\partial x}(x,r).$$

Moreover, according to the equality (4.19), one can write

$$s = r^{-1}\left[k_r(u) - \left\langle \frac{\partial F}{\partial x}(x,r), u\right\rangle\right].$$

This shows that $\partial F(x,r)$ is a singleton, and that the derivative of the partial function $F(x,\cdot)$ is given by the expression (4.23). □

In the proof of Proposition 4.9 we actually show that F is Gâteaux–differentiable with respect to both variables x and r simultaneously. Moreover, the partial Gâteaux–

differentials $\frac{\partial F}{\partial x}(x,r)$ and $\frac{\partial F}{\partial r}(x,r)$ are linked to each other by means of a relation like (4.23). These comments deserve to be recorded properly.

Theorem 4.3 *Consider the same assumptions as in Proposition 4.8, and let*

$$\begin{cases} u(x,r) \in S(x,r), \\ \alpha(x,r) = k_r(u(x,r)) = -\left[r^2 - \|u(x,r)\|^2\right]^{1/2}. \end{cases} \tag{4.24}$$

Then the Gâteaux–differential

$$F'(x,r) = \left(\frac{\partial F}{\partial x}(x,r),\ \frac{\partial F}{\partial r}(x,r)\right)$$

exists at the point (x,r), and satisfies the mixed type variational PDE

$$r\frac{\partial F}{\partial r}(x,r) + \left\langle \frac{\partial F}{\partial x}(x,r),\ u(x,r)\right\rangle = \alpha(x,r). \tag{4.25}$$

The above theorem has to be put in line with a similar type of result obtained for the Moreau–Yosida regularization technique. Under suitable assumptions, the Moreau–Yosida regularization

$$G(x,r) := \inf_{u\in X}\left\{f(u) + \frac{1}{2r}\|x-u\|^2\right\} \tag{4.26}$$

is a solution of the Hamilton–Jacobi equation

$$\begin{cases} \dfrac{\partial G}{\partial r}(x,r) + \dfrac{1}{2}\left\|\dfrac{\partial G}{\partial x}(x,r)\right\|^2 = 0 \\ \lim\limits_{r\to 0^+} G(x,r) = f(x). \end{cases} \tag{4.27}$$

Further details on this question can be found in a recent paper by Attouch [A1, Section 6.2]; see also [A2], [A3], [P1].

We refer to the PDE (4.25) as a variational one of the mixed type because it involves not only the partial derivatives of the optimal–value function $F(x,r)$, but also the optimal solutions to the variational problem (4.3). This is in contrast to the Hamilton–Jacobi equation (4.27), where no reference is made to the optimal solutions of the problem (4.23).

REFERENCES

[A1] H. ATTOUCH, Viscosity solutions of minimization problems, Preprint, Laboratoire d'Analyse Convexe, Université de Montpellier II, 1993.

[A2] H. ATTOUCH, J.L. NDOUTOUME and M. THERA, Epigraphical convergence of functions and convergence of their derivatives in Banach spaces, Seminaire d'Analyse Convexe de Montpellier, 20, exposé 9, 1990.

[A3] H. ATTOUCH and R. WETS, Epigraphical analysis, in *Analyse non linéaire,* H. Attouch, J.P. Aubin, F. Clarke, I. Ekeland, Eds., 73–100, Gauthier–Villars, Paris, 1989.

[B1] G. BEER, Lipschitz regularization and the convergence of convex functions, Numer. Funct. Anal. and Optimiz., Vol. 15(1994), 31–46.

[B2] N. BOURBAKI, *Espaces vectoriels topologiques*, Chap. III, IV, and V, Hermann, Paris, 1955.

[H1] J.B. HIRIART–URRUTY, *ϵ–subdifferential calculus*, in Research Notes in Mathematics, Vol. 57, Pitman, New York, 1982.

[L1] P.J. LAURENT, Approximation et optimisation, Hermann, Paris, 1972.

[N1] D. NOLL, Generalized second fundamental forms by way of epi–derivatives, Canadian J. Math., Vol. 46, No. 3(1994).

[P1] P. PLAZANET, Contributions à l'analyse des fonctions convexes, Ph.D. Thesis, Université Paul Sabatier de Toulouse, 1990.

[R1] R.T. ROCKAFELLAR, *Convex Analysis*, Princeton Univ. Press, Princeton, 1970.

[R2] R.T. ROCKAFELLAR, *Conjugate duality and optimization*, In Regional Conference Series in Applied Mathematics, Vol. 16, SIAM Publications, 1973.

[R3] R.T. ROCKAFELLAR, The theory of subgradients and its applications to problems of optimization, Convex and nonconvex functions, Heldermann Verlag, Berlin, 1981.

[S1] A. SEEGER, *Approximate Euler–Lagrange inclusion, approximate transversality condition, and sensitivity analysis of convex parametric problems of calculus of variations,* Set–Valued Analysis, Vol. 2(1994).

[Z1] E. ZEIDLER, Nonlinear functional analysis and its applications, Vol. III, Springer–Verlag, New York, 1985.

(seg933)