DOCTORAL THESIS 1994:139 D

DEPARTMENT OF MATHEMATICS

ISSN 0348 – 8373

Luleå University of Technology
S - 971 87 Luleå, Sweden

ISRN HLU - TH - T - - 139 - D - - SE

A Study of the Operation of Infimal Convolution

by

THOMAS STRÖMBERG

TEKNISKA HÖGSKOLAN I LULEÅ
LULEÅ UNIVERSITY OF TECHNOLOGY

A Study of the Operation of Infimal Convolution

THOMAS STRÖMBERG

Institutionen för matematik, Högskolan i Luleå
Doktorsavhandling 1994:139 D, ISSN 0348-8373.

Akademisk avhandling som med vederbörligt tillstånd av tekniska fakultetsnämnden vid Högskolan i Luleå för avläggande av teknologie doktorsexamen framlägges till offentlig granskning. Avhandlingen kommer att försvaras fredagen den 11:e mars 1994 klockan 13.00 i hörsal A117 vid Högskolan i Luleå.

FAKULTETSOPPONENT: Professor CHRISTER KISELMAN, Uppsala universitet.

BETYGSNÄMND: Docent LECH MALIGRANDA, Högskolan i Luleå; professor ANDERS MELIN, Lunds universitet; professor HANS WALLIN, Umeå universitet.

ORDFÖRANDE VID DISPUTATIONEN: Biträdande professor LARS ERIK PERSSON, Högskolan i Luleå. (Lars Erik Persson har varit handledare för respondenten.)

A Study of the Operation of Infimal Convolution

by

THOMAS STRÖMBERG

Department of Mathematics, Luleå University of Technology, S-971 87 Luleå, Sweden

A Study of the Operation of Infimal Convolution

DOCTORAL DISSERTATION

by

THOMAS STRÖMBERG

Doctoral Thesis, Department of Mathematics, Luleå University, 1994.

To be publicly examined in lecture hall A117, Luleå University, on March 11, at 1 p.m., for the degree of Doctor of Technology.

Abstract

This thesis consists of five papers (A–E), which examine the operation of infimal convolution and discuss its close connections to unilateral analysis, convex analysis, inequalities, approximation, and optimization. In particular, we attempt to provide a detailed investigation for both the convex and the non-convex case, including several examples.

Paper (A) is both a survey of and a self-contained introduction to the operation of infimal convolution. In particular, we discuss the infimal value and minimizers of an infimal convolute, infimal convolution on subadditive functions, sufficient conditions for semicontinuity or continuity of an infimal convolute, "exactness," regularizing effects, continuity of the operation of infimal convolution, and approximation methods based on infimal convolution.

A Young-type inequality, closely connected to the operation of infimal convolution, is studied in paper (B). The main results obtained are an equivalence theorem and a representation formula.

In paper (C) we consider coercive, convex, proper, and lower semicontinuous functions on a reflexive Banach space. For the infimal convolution of such functions we establish, in particular, different formulae. Moreover, we demonstrate the possibility of using the formulae obtained for solving special types of Hamilton–Jacobi equations. Furthermore, the operation of infimal convolution is interpreted from a physical viewpoint.

Paper (D) presents properties of infimal convolution of functions that are uniformly continuous on bounded sets. In particular, we present regularization procedures by means of infimal convolution. The role of growth conditions on the functions under consideration is essential.

Finally, in paper (E) we study semicontinuity, continuity, and differentiability of the infimal convolute of two convex functions. Moreover, under certain geometric conditions, the classical Moreau–Yosida approximation process is, roughly speaking, extended to the non-convex case.

Key words: Infimal convolution, epigraphical addition, convex function, regularization, duality, hyperspace topologies, Moreau–Yosida approximation, inequalities, Hamilton–Jacobi equation.

1991 Mathematics Subject Classification: 26D20, 41A65, 46N10, 49J27, 49J45, 49J50, 49L25, 52A40, 52A41, 54B20, 65K10.

Papers summarized in this dissertation:

(A) T. STRÖMBERG, An introduction to the operation of infimal convolution, Research report (57 pages), Department of Mathematics, Luleå University, 1994, *submitted.*

(B) T. STRÖMBERG, An operation connected to a Young-type inequality, *Math. Nachr.* **159** (1992), 227–243.

(C) T. STRÖMBERG, Representation formulae for infimal convolution with applications, *in* "Analysis, Algebra, and Computers in Mathematical Research—Proceedings of the Twenty-First Nordic Congress of Mathematicians" (M. Gyllenberg, L. E. Persson, Eds.), Marcel Dekker, to appear in April 1994.

(D) T. STRÖMBERG, On the epigraphical sum of functions uniformly continuous on bounded sets, *Sém. Anal. Convexe Montpellier* **22** (1992), 19.1–19.16.

(E) T. STRÖMBERG, On the operation of infimal convolution and regularization of lower semicontinuous functions, Research report (17 pages), Department of Mathematics, Luleå University, 1993, *submitted.*

Preface

This thesis consists of the following five papers:

(A) T. STRÖMBERG, An introduction to the operation of infimal convolution, Research report (57 pages), Department of Mathematics, Luleå University, 1994, *submitted.*

(B) T. STRÖMBERG, An operation connected to a Young-type inequality, *Math. Nachr.* **159** (1992), 227–243.

(C) T. STRÖMBERG, Representation formulae for infimal convolution with applications, *in* "Analysis, Algebra, and Computers in Mathematical Research—Proceedings of the Twenty-First Nordic Congress of Mathematicians" (M. Gyllenberg, L. E. Persson, Eds.), Marcel Dekker, to appear in April 1994.

(D) T. STRÖMBERG, On the epigraphical sum of functions uniformly continuous on bounded sets, *Sém. Anal. Convexe Montpellier* **22** (1992), 19.1–19.16.

(E) T. STRÖMBERG, On the operation of infimal convolution and regularization of lower semicontinuous functions, Research report (17 pages), Department of Mathematics, Luleå University, 1993, *submitted.*

Paper (A) is the most recent paper, while papers (B–E) are arranged in chronological order.

Each of these papers examines, from different perspectives, the operation of infimal convolution which we proceed to define. Suppose that X is a commutative monoid for which, except in (B), we use additive notation. Given two extended real-valued functions f and g defined on X, then their *infimal convolute* $f\Box g\colon X \to [-\infty,+\infty]$ sends each $x \in X$ to

$$(f\Box g)(x) = \inf\{f(y)+g(z) \mid y,z \in X,\ y+z=x, \text{ and } f(y)<+\infty,\ g(z)<+\infty\}.$$

(Adopting the convention that $\inf\emptyset = +\infty$.)

PAPER (A). This paper is both a survey of and a self-contained introduction to the operation of infimal convolution. In the first chapter, in particular, we present a list of examples which serves as the motivation for the study of the operation.

Chapter 2 contains a detailed investigation of geometric properties, infimal values and minimizers of an infimal convolute, infimal convolution on subadditive functions, basic regularizing effects, sufficient conditions for semicontinuity or continuity of an infimal convolute, and "exactness."

We discuss the important convex case in Chapter 3. We focus our interest on semicontinuity, continuity, and differentiability. In particular, we point out that infimal convolution has a smoothing feature, at least when X is reflexive, but that this applies in general only to first order differentiability.

In Chapter 4, we consider continuity of the mapping $(f, g) \mapsto f \Box g$, or $f \mapsto f \Box g$, with respect to modern notions of "variational convergence" (for instance, epi-convergence, slice convergence, and Attouch–Wets convergence) of functions as identified with their epigraphs.

Finally, the extremely important Moreau–Yosida approximation process and related methods are studied in Chapter 5. We review the classical case of approximation of a convex function on a Hilbert space as well as present results for more general situations.

In particular, I have independently established several results (for instance, Theorem 2.7, Theorem 3.7 (in part), Theorem 4.4, Theorem 4.8, Theorem 5.10, Theorem 5.11, Theorem 5.12). However, in my opinion, my main contribution is to provide a review article which unifies, complements or extends existing results.

PAPER (B). In this article we consider infimal convolution for the case where X is the *multiplicative* group $]0, +\infty[$. To be precise, we examine the restriction, denoted $\oplus$, of infimal convolution to the set Φ of all continuous and nondecreasing functions $\varphi\colon]0, +\infty[\to]0, +\infty[$ such that $\varphi(t) \to 0$ as $t \to 0+$ and $\varphi(t) \to +\infty$ as $t \to +\infty$. Explicitly,

$$(\varphi \oplus \psi)(s) = \inf\{\varphi(t) + \psi(u) |\ t, u \in]0, +\infty[, \text{ and } s = tu\}$$

for any $\varphi, \psi \in \Phi$ and all $s \in]0, +\infty[$. Clearly, the *Young-type inequality*

$$(\varphi \oplus \psi)(tu) \leq \varphi(t) + \psi(u)$$

holds for all $t, u \in]0, +\infty[$. The study of such inequalities was the purpose of this paper.

The main results are an equivalence theorem (Theorem 1) and a representation formula (Theorem 2).

PAPER (C). Here we consider coercive convex proper lower semicontinuous functions on a reflexive Banach space. For the infimal convolution of such functions we establish, in particular, different formulae (Theorem 1, Theorem 2). Moreover, we demonstrate the possibility of using the formulae obtained for solving special types of Hamilton–Jacobi equations (Proposition 2). Furthermore, the operation of infimal convolution is interpreted from a physical viewpoint.

PAPER (D). We study infimal convolution of functions, defined on an arbitrary normed vector space, that are uniformly continuous on bounded sets (Propositions 1 & 2, Theorem 1). In particular, we present regularization procedures by means of infimal convolution (Theorem 2, Proposition 3). The role of growth conditions on the functions under consideration is essential.

PAPER (E) Basic semicontinuity, continuity and differentiability properties of the infimal convolute of two convex functions are presented (Propositions 1 & 2, Theorems 1 &

2). Moreover, assuming X to be a locally uniformly rotund Banach space the dual of which is also locally uniformly rotund, the classical Moreau–Yosida approximation process is, roughly speaking, extended to the non-convex case (Theorem 3).

Acknowledgements. I am indebted to Professor Lars Erik Persson, my supervisor, for his constant support and for inspiring discussions. Moreover, I would like to express my gratitude to Professor Lech Maligranda who (together with Lars Erik) suggested that I should study inequalities of the Young-type, which was the point of departure in my investigations. Finally, I would like to thank Professors Henryk Hudzik (Poznań), Alois Kufner (Prague), and Hedy Attouch (Montpellier) for their hospitality in connection with my visits and for all their suggestions which have helped me in my research.

A

An Introduction to the Operation of Infimal Convolution

THOMAS STRÖMBERG

Department of Mathematics, Luleå University of Technology, S-971 87 Luleå, Sweden

Abstract

This paper is both a self-contained introduction to and a survey of the operation of infimal convolution. We consider the convex as well as the non-convex case. In particular, we provide a detailed investigation of the regularizing effects of infimal convolution, and study continuity properties of the operation with respect to modern notions of "variational convergence." Moreover, we examine regularization methods based on infimal convolution. Several examples are included and some well-known results are complemented, unified or extended in various ways.

Contents

Chapter 1

Introduction and preliminaries

In this paper we will consider the operation of infimal convolution. The purpose is to provide a survey of the subject as well as to complement or extend existing results. For the convenience of the reader we will give a self-contained presentation. Therefore, in particular, we have collected in this chapter, in detail, background notations and definitions, and developed several motivating examples in order to make it possible for the reader to fully appreciate the material.

1.1 Introduction

The first systematic study of the operation of infimal convolution was performed by J.-J. Moreau, see [59, 60, 61, 62, 63, 64]. However, the operation appears in W. Fenchel's monograph [31], and in some sense in the paper [34] by F. Hausdorff. Moreover, the operation is examined in the monographs [2, 41, 49, 72] by H. Attouch, A. D. Ioffe & V. M. Tihomirov, P.-J. Laurent, and R. T. Rockafellar, respectively.

In the following definition, which is the starting-point in our study, the symbol "$\dot{+}$" stands for "upper addition" of extended real numbers in the sense of J.-J. Moreau [63]. Moreover, by an "additive monoid" we mean a commutative monoid for which we use additive notation.

Definition. Suppose that X is an additive monoid. The binary operation of *infimal convolution* is defined on the set $\bar{\mathbb{R}}^X$ of all extended real-valued functions on X as follows: For any two functions $f, g: X \to \bar{\mathbb{R}}$, the *infimal convolute* $f \Box g$ assigns to each $x \in X$ the extended real number

$$\begin{aligned}(f \Box g)(x) &= \inf\{f(y) \dot{+} g(z) \mid y, z \in X, \text{ and } y + z = x\} \qquad (1.1)\\ &= \inf\{f(y) + g(z) \mid y, z \in X,\ y + z = x, \text{ and } f(y) < +\infty,\ g(z) < +\infty\}.\end{aligned}$$

The infimal convolution of f with g is said to be *exact at x* provided the infimum (1.1) is achieved, and *exact* if it is exact at each element of X.

The terminology "convolution" stems from the fact that, when X is an additive group, $(f \square g)(x)$ is given by

$$(f \square g)(x) = \inf_{y \in X} (f(y) \dot{+} g(x-y)),$$

a formula similar to that defining integral convolution, and "infimal" reminds of the infimum appearing in the definition. However, infimal convolution is also known as "epigraphical addition" [7] because, geometrically, performing infimal convolution of f with g amounts to the adding of the strict epigraphs of f and g.

It is clear from the definition of infimal convolution that, given two extended real-valued functions f and g on X, $f \square g$ is the largest function h for which

$$h(y+z) \leq f(y) \dot{+} g(z) \text{ for all } y, z \in X. \tag{1.2}$$

For any subset A of X, $\delta_A : X \to \bar{\mathbb{R}}$ denotes the *indicator function* of A: $\delta_A(x) = 0$ if $x \in A$, $\delta_A(x) = +\infty$ if $x \in X \setminus A$. If A and B are two subsets of X, then

$$\delta_A \square \delta_B = \delta_{A+B}.$$

Therefore, infimal convolution includes Minkowski addition of subsets of X.

1.2 Preliminaries

Throughout this text, $\mathbb{N}$ and $\mathbb{R}$ denote the set of natural and real numbers, respectively. Moreover, $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ stands for the set of extended real numbers, while $\mathbb{R}_+ = [0, +\infty[$ denotes the non-negative real numbers, and $\mathbb{P} =]0, +\infty[$ the positive real numbers. Since we shall constantly be dealing with extended real-valued functions, we need notions of addition of extended real numbers. We follow J.-J. Moreau [63] and define the operation of *upper addition* $\dot{+}$ (respectively, *lower addition* $\underset{\cdot}{+}$) as the commutative extension of usual addition obtained by taking $(-\infty) \dot{+} (+\infty) = +\infty$ (respectively, $(-\infty) \underset{\cdot}{+} (+\infty) = -\infty$). Of course, we write $+$ for both of these notions of addition when they coincide, that is, when the sum of $+\infty$ and $-\infty$ never occurs.

With any function $f: X \to \bar{\mathbb{R}}$ we associate its

- *epigraph*

$$\operatorname{epi} f = \{(x, \alpha) \in X \times \mathbb{R} \mid f(x) \leq \alpha\};$$

- *strict epigraph*

$$\operatorname{epi}_s f = \{(x, \alpha) \in X \times \mathbb{R} \mid f(x) < \alpha\};$$

- *sublevel set at height* α, $\alpha \in \mathbb{R}$,

$$\operatorname{slv}(f; \alpha) = \{x \in X \mid f(x) \leq \alpha\};$$

- *essential domain*

$$\operatorname{dom} f = \{x \in X \mid f(x) < +\infty\};$$

- *infimal value* $\inf f = \inf_{x \in X} f(x) \in \bar{\mathbb{R}}$;
- set

$$\arg\min f = \{x \in X \mid f(x) = \inf f\}$$

of minimizers.

Recall that f is called *proper* if it does not take on the value $-\infty$ and is not the constant $+\infty$.

For any $t \in \mathbb{P}$ we write $t \bullet f$ for the function $tf(\cdot/t)$ when X is a vector space. It satisfies the identity $\operatorname{epi}(t \bullet f) = t \operatorname{epi} f$.

Suppose that X is a topological space. Then, f is lower semicontinuous if and only if $\operatorname{epi} f$ is a closed subset of $X \times \mathbb{R}$. Dually, f is upper semicontinuous if and only if $\operatorname{epi}_s f$ is an open subset of $X \times \mathbb{R}$. The *lower semicontinuous regularization* $\operatorname{cl} f : X \to \bar{\mathbb{R}}$ of f is the function whose epigraph is equal to the closure of the epigraph of f in $X \times \mathbb{R}$:

$$\operatorname{epi}(\operatorname{cl} f) := \overline{\operatorname{epi} f}.$$

A pointwise formulation reads

$$(\operatorname{cl} f)(x) = \sup_{V \in \mathcal{N}(x)} \inf_{\xi \in V} f(\xi) \text{ at each } x \in X,$$

where $\mathcal{N}(x)$ denotes the filter of neighborhoods of x. Obviously, for any $x \in X$, $(\operatorname{cl} f)(x) \leq f(x)$, and the equality $(\operatorname{cl} f)(x) = f(x)$ holds if and only if f is lower semicontinuous at x.

Given a net $\langle A_\lambda \rangle_{\lambda \in \Lambda}$ of subsets of X, then its *lower closed limit* $\liminf_{\lambda \in \Lambda} A_\lambda$ (respectively, *upper closed limit* $\limsup_{\lambda \in \Lambda} A_\lambda$) is the set of all $x \in X$ for which each neighborhood of x meets A_λ for all λ in some residual (respectively, cofinal) subset of Λ. That these sets are both closed, is immediate from the formulas (see G. Beer [15, Proposition 5.2.2])

$$\begin{aligned}
\liminf_{\lambda \in \Lambda} A_\lambda &= \bigcap \left\{ \overline{\textstyle\bigcup_{\lambda \in \Sigma} A_\lambda} \,\middle|\, \Sigma \text{ is a cofinal subset of } \Lambda \right\}, \\
\limsup_{\lambda \in \Lambda} A_\lambda &= \bigcap \left\{ \overline{\textstyle\bigcup_{\lambda \in \Sigma} A_\lambda} \,\middle|\, \Sigma \text{ is a residual subset of } \Lambda \right\}.
\end{aligned}$$

The net $\langle A_\lambda \rangle_{\lambda \in \Lambda}$ is said to converge to A in the sense of *Kuratowski-Painlevé* if

$$\limsup_{\lambda \in \Lambda} A_\lambda \subseteq A \subseteq \liminf_{\lambda \in \Lambda} A_\lambda. \tag{1.3}$$

If $\langle A_\lambda \rangle_{\lambda \in \Lambda}$ converges to A in the above sense, then we actually have equality throughout (1.3) and A is a closed set.

We need a few concepts and notations of a common use in the field of convex analysis. An extended real-valued function f defined on a vector space is convex if and only if epi f (or, equivalently, $\mathrm{epi}_s f$) is a convex set. Suppose that X is a locally convex topological vector space over $\mathbb{R}$. The topological dual space of X is represented by X^*. The weak* topology on X^* is denoted by w^*. By $\Gamma(X)$ we understand the set of all proper convex lower semicontinuous functions from X into $\mathbb{R} \cup \{+\infty\}$, and $\Gamma^*(X^*)$ stands for the proper convex weak* lower semicontinuous functions on X^*.

Suppose that $f: X \to \bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ is an extended real-valued function on X. The *conjugate function* $f^*: X^* \to \bar{\mathbb{R}}$ of f is defined by the assignment

$$f^*(x^*) = \sup_{x \in X}(\langle x^*, x\rangle - f(x)) \text{ for each } x^* \in X^*.$$

Symmetrically, for any $\varphi: X^* \to \bar{\mathbb{R}}$, φ^* is the extended real-valued function on X that to each $x \in X$ assigns the extended real number

$$\varphi^*(x) = \sup_{x^* \in X^*} (\langle x^*, x\rangle - \varphi(x^*)).$$

It is easily verified that $f^{**} \leq f$. It is a classical result that $f = f^{**}$ if and only if either $f \in \Gamma(X)$ or f is the constant $-\infty$ or the constant $+\infty$. The map $f \mapsto f^*$, which is usually called the *Young–Fenchel transform* or the *Legendre–Fenchel transform*, is a bijection of $\Gamma(X)$ onto $\Gamma^*(X^*)$.

The *subdifferential* ∂f of f is

$$\partial f = \{(x, x^*) \in X \times X^* \mid f(x) \in \mathbb{R}, \text{ and } \forall y \in X: \langle x^*, y - x\rangle \leq f(y) - f(x)\},$$

which is viewed upon as a multifunction from X into X^*. In terms of infimal convolution,

$$\partial f = \{(x, x^*) \in X \times X^* \mid f(x) \in \mathbb{R}, \text{ and } (f \Box x^*)(x) = f(x)\}.$$

Yet another characterization of ∂f reads

$$\partial f = \{(x, x^*) \in X \times X^* \mid \langle x^*, x\rangle = f(x) + f^*(x^*)\}.$$

A convex proper function f is subdifferentiable at $x \in \operatorname{dom} f$, that is to say $\partial f(x) \neq \emptyset$, if f is continuous at x. Then $\partial f(x)$ is a convex and weak* compact subset of X^*.

If $\varepsilon \in \mathbb{R}_+$, then the *$\varepsilon$-subdifferential* $\partial_\varepsilon f$ is the following enlargement of ∂f:

$$\partial_\varepsilon f = \{(x, x^*) \in X \times X^* \mid f(x) \in \mathbb{R}, \text{ and } \forall y \in X: \langle x^*, y - x\rangle - \varepsilon \leq f(y) - f(x)\}.$$

Clearly,

$$\partial f(x) = \bigcap_{\varepsilon \in \mathbb{P}} \partial_\varepsilon f(x) \text{ for each } x \in X.$$

If $f \in \Gamma(X)$, $x \in \operatorname{dom} f$ and $\varepsilon \in \mathbb{P}$, then $\partial_\varepsilon f(x)$ is nonempty, convex and weak* closed.

A proper extended real-valued function f on X is called *cs-convex* provided that, for each sequence $\langle x_n \rangle_{n \in \mathbb{N}}$ in X and each sequence $\langle t_n \rangle_{n \in \mathbb{N}}$ of numbers in $[0,1]$ with $\sum_{n=0}^{\infty} t_n = 1$, the following implication holds:

$$\sum_{n=0}^{\infty} t_n x_n \text{ convergent } \Rightarrow f\left(\sum_{n=0}^{\infty} t_n x_n\right) \leq \liminf_{N \to \infty} \sum_{n=0}^{N} t_n f(x_n).$$

Every cs-convex function is convex, and every element of $\Gamma(X)$ is cs-convex.

For any nonempty subset A of X, we put

$$\operatorname{cone} A = \{\lambda x | \ \lambda \in \mathbb{R}_+ \text{ and } x \in A\}.$$

When $(X, \|\cdot\|)$ is a normed vector space we write $(X^*, \|\cdot\|_*)$ for the normed dual, $B(x, \rho)$ for the closed ball centered at x and of radius ρ, and U for the closed unit ball in X, that is, $U = B(0,1)$. If $x \in X$ and A is a nonempty subset of X, then $d(x, A) = \inf_{a \in A} \|x - a\|$ denotes the *distance* from x to A. Moreover, a function $f: X \to \bar{\mathbb{R}}$ is called *weakly coercive* (respectively, *coercive*) provided

$$\text{as } \|x\| \to +\infty, \ f(x) \to +\infty \text{ (respectively, } f(x)/\|x\| \to +\infty).$$

Recall that X is called *rotund* provided each open segment in U is disjoint from the surface Σ of U. We say that X is *locally uniformly rotund* if

$$\inf\{1 - \|(x+y)/2\| \,|\, y \in \Sigma, \text{ and } \|x - y\| \leq \varepsilon\} > 0 \text{ for each } x \in \Sigma \text{ and } \varepsilon \in]0,2],$$

see M. M. Day [25].

1.3 Introductory examples

This section consists of a list of examples where it is natural to consider the operation of infimal convolution.

EXAMPLE. (An economical example.) Suppose that a person P wishes to buy x quantities of a certain product which is supplied by two manufacturers A and B. (Depending on the type of product, x is either a nonnegative real number or a natural number.) Suppose moreover that the price at which manufacturer A (respectively, B) provides y (respectively, z) quantities of the product is equal to $f(y)$ (respectively, $g(z)$). Naturally, P is then interested in the following:

(i) The infimum of the total cost $f(y) + g(z)$ under the constraint $y + z = x$. Put differently, P wishes to calculate $(f \Box g)(x)$.

(ii) The set of solutions to the infimal problem (i).

Assume that $f, g: \mathbb{R}_+ \to \mathbb{R}_+$ are concave continuous unbounded functions with $f(0) = 0 = g(0)$, which is reasonable in the continuous case. Then $f \square g$ is equal to $f \wedge g$, the meet of f and g. Indeed,

$$(f \square g)(x) = \min_{y \in [0,x]} (f(y) + g(x-y)),$$

and $y \mapsto f(y) + g(x-y)$ is a concave function which implies that the above minimum is achieved at $y = 0$ or $y = x$. Thus

$$(f \square g)(x) = \min\{f(x), g(x)\} = (f \wedge g)(x).$$

However, it makes sense to consider other types of price functions f and g from $\mathbb{R}_+$ into $\mathbb{R}_+$, as well as to study the discrete case. The latter involves infimal convolution of nonnegative sequences.

EXAMPLE. (Convex optimization.) Suppose that X is a locally convex Hausdorff topological vector space. Let $f, g: X \to \mathbb{R} \cup \{+\infty\}$ be convex functions and f^*, g^* their respective conjugate functions. It is of fundamental interest in the duality theory of convex optimization to know when we can state

$$\inf_{x \in X} (f(x) + g(x)) + \min_{y^* \in X^*} (f^*(y^*) + g(-y^*)) = 0.$$

In terms of conjugate function this amounts to say that

$$(f+g)^*(0) = (f^* \square g^*)(0)$$

with the infimal convolution of f^* and g^* exact at 0, that is, with the infimum defining $(f^* \square g^*)(0)$ achieved. According to a result due to R. T. Rockafellar [71] (an extension of the classical Fenchel's duality theorem [31]), we know that $(f+g)^* = f^* \square g^*$ with exact infimal convolution provided that there exists a point in $\operatorname{dom} f \cap \operatorname{dom} g$ at which at least one of f and g is continuous. For further generalizations as well as complements to Fenchel's duality theorem, see for instance [6, 42, 73, 77].

EXAMPLE. (The sum of convex sets.) The Minkowski sum of two convex closed subsets of the plane $\mathbb{R}^2$ admits a (local) description by means of infimal convolution of convex functions on $\mathbb{R}$. The idea is, given a boundary point a of a closed convex set A with nonempty interior, to choose an affine coordinate system such that A agrees near a with the epigraph of some convex function, and to make use of the fact that addition of (strict) epigraphs is the operation of infimal convolution. Consult C. O. Kiselman [44].

EXAMPLE. (A Hamilton–Jacobi equation.) Let X be a Eucledian space. Suppose that $\mathcal{H}: X \to \mathbb{R}$ is convex with $\mathcal{H}(p)/|p| \to +\infty$ as $|p|$ tends to $+\infty$, and that $f: X \to \mathbb{R}$ is lower semicontinuous and bounded from below. The "viscosity solution" of the Hamilton–Jacobi equation

$$\begin{aligned} \dot{u}(x,t) + \mathcal{H}(Du(x,t)) &= 0, \quad (x,t) \in X \times \mathbb{P}, \\ \lim_{t\to 0} u(x,t) &= f(x), \quad x \in X, \end{aligned}$$

is given by $u(\cdot,t) = f\Box(t \bullet \mathcal{H}^*)$, $t \in \mathbb{P}$. Study P. L. Lion's monograph [51].

EXAMPLE. (Regularization.) An important tool in nonlinear analysis is the method of regularization in which a proper and lower semicontinuous function $f: X \to \mathbb{R} \cup \{+\infty\}$, defined on a normed space $(X, \|\cdot\|)$, is approximated by means of infimal convolution of f with members of a parameterized family $\{g_t |\, t \in \mathbb{P}\}$, such that when the parameter t approaches zero from above, the approximant $f\Box g_t$ approaches f pointwise. Particularly important in the literature are the following parameterized families of kernels:

(i) $\{\|\cdot\|^2/(2t) |\, t \in \mathbb{P}\}$;

(ii) $\{\|\cdot\|/t |\, t \in \mathbb{P}\}$;

(iii) $\{\delta_{tU} |\, t \in \mathbb{P}\}$.

The kernel functions in (ii) were probably first used by F. Hausdorff [34] who attributes the definition to M. Pasch. (To be more correct, their definition is in the setting of an arbitrary metric space.) Let f be a proper function on X. Put, for each $t \in \mathbb{P}$, $f_t = f\Box t^{-1}\|\cdot\|$. Suppose that $f_{t_0}(x_0) > -\infty$ for some $(t_0, x_0) \in \mathbb{P} \times X$. Then, provided $t \in]0, t_0]$, f_t is real-valued and Lipschitz continuous with constant $1/t$, and it is the greatest minorant of f with this property. Moreover, at each $x \in X$, $f_t(x) \uparrow f(x)$ as $t \downarrow 0$. See for instance J.-B. Hiriart-Urruty [35] or S. Fitzpatrick & R. R. Phelps [32].

The above procedure of regularization enables us to characterize different notions of convergence (epi-convergence, slice convergence, Attouch–Wets convergence) of a net of convex and lower semicontinuous functions by means of the corresponding net of regularized functions, consult [2, 15, 86].

EXAMPLE. (Extension of Lipschitz functions.) Let $(X, \|\cdot\|)$ be a normed vector space. Suppose that S is a nonempty subset of X and that $f: S \to \mathbb{R}$ is Lipschitz continuous on S with constant k. Let F denote the extension by $+\infty$ to X: $F(x) = f(x)$ if $x \in S$, and $F(x) = +\infty$ if $x \in X \setminus S$. Then $F\Box k\|\cdot\|$ is real-valued, Lipschitz continuous with constant k and coincides with f on S. Moreover, $F\Box k\|\cdot\|$ is the largest extension of f which is globally Lipschitz continuous with constant k. More precisely, if $g: X \to \mathbb{R} \cup \{+\infty\}$ is a proper function, $k \in \mathbb{P}$, and $g\Box k\|\cdot\|$ takes on one real value, then $g\Box k\|\cdot\|$ is the largest Lipschitz continuous (real-valued) function with constant k that g majorizes. See the paper [35] by J.-B. Hiriart-Urruty.

EXAMPLE. (Ekeland's Variational Principle.) Let $(X, \|\cdot\|)$ be a Banach space. Suppose that $f: X \longrightarrow \mathbb{R} \cup \{+\infty\}$ is lower semicontinuous with $\inf f \in \mathbb{R}$. The celebrated Ekeland's Variational Principle principle states that to any $\varepsilon \in \mathbb{P}$, any $x \in X$ with $f(x) \leq \inf f + \varepsilon$, and any $\delta \in \mathbb{P}$, there exists $\tilde{x} \in \operatorname{dom} f \cap B(x, \varepsilon/\delta)$ such that

(a) $f(\tilde{x}) + \delta\|x - \tilde{x}\| \leq f(x)$;

(b) $f(y) + \delta\|\tilde{x} - y\| > f(\tilde{x})$ whenever $y \neq \tilde{x}$.

In particular, part (b) may be expressed, in terms of infimal convolution, by saying that $(f\Box\delta\|\cdot\|)(\tilde{x}) = f(\tilde{x})$ with "strictly exact" infimal convolution (that is, the infimum defining $(f\Box\delta\|\cdot\|)(\tilde{x})$ is uniquely attained). See I. Ekeland [27, 28] or R. R. Phelps [70].

EXAMPLE. (Moreau's Decomposition Theorem.) Suppose that X is a Hilbert space. A classical decomposition theorem due to J.-J. Moreau [64] states that

$$f\Box\tfrac{1}{2}\|\cdot\|^2 + f^*\Box\tfrac{1}{2}\|\cdot\|^2 = \tfrac{1}{2}\|\cdot\|^2 \text{ for any } f \in \Gamma(X).$$

As observed in [40], by choosing f as the indicator function of a closed convex cone C, f^* is the indicator function of the polar cone C° to C, $f\Box\frac{1}{2}\|\cdot\|^2$ is half of the square of the distance function to C, so that the conclusion of the theorem reads as a kind of a Pythagoras' theorem:

$$d(\cdot, C)^2 + d(\cdot, C^\circ)^2 = \|\cdot\|^2.$$

If g, h are proper extended real-valued functions on X, we define $h\triangle g: X \longrightarrow \bar{\mathbb{R}}$ by the formula

$$(h\triangle g)(y) = \sup_{z\in X}(h(y+z) \dot{+} (-g(z))) \text{ for all } y \in X.$$

Clearly, $(f\Box g)\triangle g \leq f$ for any $f, g: X \longrightarrow \bar{\mathbb{R}}$, compare with (1.2). The operation $(h, g) \mapsto h\triangle g$ is important in the study of "infimal convolution equations" [38, 57] and is sometimes called "deconvolution." According to a formula due to J.-B. Hiriart-Urruty [37],

$$h\triangle g = (h^* \dot{+} (-g^*))^* \text{ if } g, h \in \Gamma(X).$$

Moreau's theorem has been revisited by J.-B. Hiriart-Urruty and P. Plazanet [40]: If $\varphi \in \Gamma(X)$ and $\psi \in \Gamma(X)$ satisfy $\varphi + \psi = \frac{1}{2}\|\cdot\|^2$, then

(a) there exists exactly one f in $\Gamma(X)$ such that $\varphi = f\Box\frac{1}{2}\|\cdot\|^2$ and $\psi = f^*\Box\frac{1}{2}\|\cdot\|^2$;

(b) the equalities $f = \varphi\triangle\frac{1}{2}\|\cdot\|^2 = \psi^* - \frac{1}{2}\|\cdot\|^2$ hold.

EXAMPLE. (The K-functional.) Let $(X_0, \|\cdot\|_{X_0})$ and $(X_1, \|\cdot\|_{X_1})$ be normed vector spaces, both subspaces of some common Hausdorff topological vector space X. Then the sum $X_0 + X_1$ may be endowed with the norm $\|\cdot\|_{X_0+X_1}$ sending each $x \in X_0 + X_1$ to

$$\|x\|_{X_0+X_1} = \inf\{\|x_0\|_{X_0} + \|x_1\|_{X_1} \mid x_0 \in X_0,\ x_1 \in X_1, \text{ and } x_0 + x_1 = x\}.$$

If we denote by p_k $(k \in \{0,1\})$ the extension of $\|\cdot\|_{X_k}$ by $+\infty$ to the whole of X, that is to say

$$p_k(x) = \begin{cases} \|x\|_{X_k} & \text{if } x \in X_k, \\ +\infty & \text{if } x \in X \setminus X_k, \end{cases}$$

then $p_0 \Box p_1$ is an extension of $\|\cdot\|_{X_0+X_1}$ (by $+\infty$) to X. In the theory of interpolation spaces, the K-functional $K: (X_0 + X_1) \times \mathbb{P} \to \mathbb{R}_+$ plays a central role [18]. In terms of infimal convolution,

$$K(\cdot, t) = [p_0 \Box t p_1]|_{X_0+X_1} \text{ for each } t \in \mathbb{P}.$$

Moreover, the so-called E-functional $E: (X_0 + X_1) \times \mathbb{P} \to \mathbb{R}_+$ may be written

$$E(\cdot, t) = [p_0 \Box \delta_{tU_1}]|_{X_0+X_1} \text{ for each } t \in \mathbb{P},$$

where U_1 stands for the unit ball in X_1. More explicitly,

$$E(x,t) = \inf\{\|x_0\|_{X_0} |\, x_0 \in X_0,\ x_1 \in X_1,\ \|x_1\|_{X_1} \leq t, \text{ and } x_0 + x_1 = x\}$$

for all $(x,t) \in (X_0 + X_1) \times \mathbb{P}$. The K-functional and E-functional are closely connected (see [54, Proposition 1]):

$$K(x,t) = \inf_{s \in \mathbb{P}} (st + E(x,s)) \text{ for all } (x,t) \in (X_0 + X_1) \times \mathbb{P}.$$

Several explicit descriptions of these functionals can be found in the paper [54] by L. Maligranda and L. E. Persson.

EXAMPLE. (A Young-type inequality.) Let us consider the set Φ of all continuous nondecreasing functions $\varphi: \mathbb{P} \to \mathbb{P}$ such that $\varphi(t) \to 0$ as $t \to 0+$ and $\varphi(t) \to +\infty$ as $t \to +\infty$. Given two functions φ and ψ in Φ, put

$$(\varphi \oplus \psi)(s) = \inf\{\varphi(t) + \psi(u) |\, t, u \in \mathbb{P}, \text{ and } s = tu\}$$

for all $s \in \mathbb{P}$. Evidently, the *Young-type inequality*

$$(\varphi \oplus \psi)(tu) \leq \varphi(t) + \psi(u)$$

holds for all $t, u \in \mathbb{P}$. The operation $\oplus$ is the restriction of infimal convolution, with X equal to the *multiplicative* group $\mathbb{P}$, to Φ. The operation $\oplus$ has been investigated by C. Bylka and W. Orlicz in [23], by L. Maligranda and L. E. Persson in [52, 53], and by the author in [78].

Chapter 2

Elementary properties

In this chapter we present and discuss some elementary properties and examples of infimal convolution. In particular, we discuss the infimal value and corresponding minimizers of an infimal convolute, infimal convolution of subadditive functions, semicontinuity and continuity properties, and exactness.

Notation. Throughout this chapter, X stands for an additive monoid.

2.1 Basic facts

We shall frequently make use of the following result:

Theorem 2.1 $(\bar{\mathbb{R}}^X, \Box)$ *is a commutative monoid with the indicator function* $\delta_{\{0\}}$ *acting as the neutral element.*

PROOF. A proof may be constructed in a straightforward manner; see for instance J.-J. Moreau [63, Section 6.3].

EXAMPLE. Suppose that X is an additive group with more than one element, and consider the equation

$$f \Box 0 = 0,$$

where 0 denotes the zero function on X. We obviously have $f \Box 0 = 0$ if and only if $\inf f = 0$, since $f \Box 0$ is the constant $\inf f$. Thus there exist many solutions to the equation; the smallest one is the zero function. Moreover, it is clear that the equation $f \Box 0 = h$ has no solution if the given function h is nonconstant. In particular, this example shows that $(\bar{\mathbb{R}}^X, \Box)$ is not a group.

Theorem 2.2 *Let f and g be functions from X into $\bar{\mathbb{R}}$. The following assertions hold.*

(a) $\operatorname{dom} f \Box g = \operatorname{dom} f + \operatorname{dom} g$;

(b) $\operatorname{epi}_s f\Box g = \operatorname{epi}_s f + \operatorname{epi}_s g$;

(c) $\operatorname{epi} f\Box g \supseteq \operatorname{epi} f + \operatorname{epi} g$, *and equality holds if and only if the infimal convolution is exact at each* $x \in X$ *where* $(f\Box g)(x) \in \mathbb{R}$.

(d) $\operatorname{epi} f\Box g$ *is equal to the "vertical closure" of* $\operatorname{epi} f + \operatorname{epi} g$, *that is,*

$$\operatorname{epi} f\Box g = \{(x,\alpha) \in X \times \mathbb{R} |\ \alpha \geq \inf\{a \in \mathbb{R} |\ (x,a) \in \operatorname{epi} f + \operatorname{epi} g\}\}.$$

(e) *If* X *is a topological additive group, then*

$$\overline{\operatorname{epi} f\Box g} = \overline{\operatorname{epi} f + \operatorname{epi} g},$$

where the bars denote closure in the product topology of $X \times \mathbb{R}$.

PROOF. Assertions (a) and (b) are proved in [63, Section 8c] and [63, Section 7b], respectively.

The first part of (c) is a simple consequence of the definition of $f\Box g$ [63, Section 7b]. For the second part, suppose first that

$$\operatorname{epi} f\Box g = \operatorname{epi} f + \operatorname{epi} g.$$

Take $x \in X$ with $(f\Box g)(x) \in \mathbb{R}$. Then $(x, (f\Box g)(x))$ belongs to $\operatorname{epi} f\Box g$ and, hence, admits a decomposition

$$(x, (f\Box g)(x)) = (y, \beta) + (z, \gamma), \text{ where } (y,\beta) \in \operatorname{epi} f \text{ and } (z,\gamma) \in \operatorname{epi} g.$$

Then,

$$(f\Box g)(x) = \beta + \gamma \geq f(y) + g(z),$$

and, trivially,

$$(f\Box g)(x) \leq f(y) + g(z).$$

Therefore, $f\Box g$ is exact at x:

$$(f\Box g)(x) = f(y) + g(z).$$

In order to establish the reverse implication, suppose that $f\Box g$ is exact at each $x \in X$ where $(f\Box g)(x)$ is finite. Let (x,α) be an element of $\operatorname{epi} f\Box g$. If $(x,\alpha) \in \operatorname{epi}_s f\Box g$, then, by (b),

$$(x,\alpha) \in \operatorname{epi}_s f + \operatorname{epi}_s g \subseteq \operatorname{epi} f + \operatorname{epi} g.$$

Otherwise, $(x,\alpha) = (x, (f\Box g)(x))$. Hence, $(f\Box g)(x) \in \mathbb{R}$ and there exist $y, z \in X$ with $y + z = x$ and $(f\Box g)(x)$ equal to $f(y) + g(z)$. In particular, $f(y)$ and $g(z)$ are finite numbers. Thus,

$$(x,\alpha) = (x, (f\Box g)(x)) = (y, f(y)) + (z, g(z)) \in \operatorname{epi} f + \operatorname{epi} g.$$

A proof of part (d) can be found in [63, Section 7b].

(e) The inclusion

$$\overline{\operatorname{epi} f \Box g} \supseteq \overline{\operatorname{epi} f + \operatorname{epi} g}$$

is a direct consequence of (c). Moreover, denoting by "vcl" the operation of "vertical closure," we have, according to (d),

$$\operatorname{epi} f \Box g = \operatorname{vcl}(\operatorname{epi} f + \operatorname{epi} g) \subseteq \overline{\operatorname{epi} f + \operatorname{epi} g},$$

from which the remaining inclusion

$$\overline{\operatorname{epi} f \Box g} \subseteq \overline{\operatorname{epi} f + \operatorname{epi} g}$$

follows. □

Since infimal convolution is highly applicable in connection with optimization (in fact, performing the infimal convolution of f with g is an optimization problem in itself), we are interested in the relations between the infimal value, as well as the points at which this infimum is achieved, of $f \Box g$ and those of f and g. Some basic properties of infimal values and minimizers are collected in our next statement.

Theorem 2.3 *Let $f, g\colon X \to \bar{\mathbb{R}}$. Then the following statements hold:*

(a) $\inf f \Box g = \inf f \dot{+} \inf g$.

(b) $\arg\min f \Box g \supseteq \arg\min f + \arg\min g$; *that is, if $\bar{y}$ minimizes f and $\bar{z}$ minimizes g, then the sum $\bar{y} + \bar{z}$ minimizes $f \Box g$.*

(c) *Suppose that $f \Box g$ is exact and that* $\inf f$ *and* $\inf g$ *are both real numbers. Then*

$$\arg\min f \Box g = \arg\min f + \arg\min g.$$

(d) *Let X be a topological additive group. Suppose that f and g are sequentially lower semicontinuous, and that* $\inf f$ *and* $\inf g$ *are both real numbers. Moreover, suppose that every minimizing sequence for f has a convergent subsequence. Then*

$$\arg\min f \Box g = \arg\min f + \arg\min g.$$

REMARK. The assumptions on f in (d) are often in the literature expressed by saying that the problem of minimizing f is "well-posed in the generalized sense," see for instance A. L. Dontchev & T. Zolezzi [26]. For such a function f it is readily verified that $\arg\min f$ is a sequentially compact closed nonempty set.

PROOF. (a) and (b) are self-evident consequences of the definition of $f \Box g$ [7, Theorem 2.3]. In view of (b), for (c) and (d), it suffices to prove that $\arg\min f \Box g$ is a subset of $\arg\min f + \arg\min g$.

(c) Suppose that $\bar{x}$ minimizes $f\Box g$. Since $f\Box g$ is exact there exist $\bar{y}$ and $\bar{z}$ in X with $\bar{y}+\bar{z}=\bar{x}$ and $(f\Box g)(\bar{x}) = f(\bar{y})+g(\bar{z})$. By (a), $f(\bar{y})+g(\bar{z}) = \inf f + \inf g$, and it follows that $\bar{y}$ and $\bar{z}$ minimize, respectively, f and g.

(d) Assume that $f\Box g$ has a minimum at $\bar{x} \in X$ and let $\langle y_n \rangle$ and $\langle z_n \rangle$ be sequences in X such that for each $n \in \mathbb{N}$

$$y_n + z_n = \bar{x} \text{ and } f(y_n) + g(z_n) \leq (f\Box g)(\bar{x}) + 2^{-n} = \inf f + \inf g + 2^{-n}.$$

Then we must have $f(y_n) \to \inf f$ as n tends to infinity. By hypothesis, $\langle y_n \rangle$ has a convergent subsequence $\langle y_{n_k} \rangle$, say $y_{n_k} \to \bar{y}$. Hence, $z_{n_k} \to \bar{z}$ where $\bar{y}+\bar{z}=\bar{x}$. Since f and g are sequentially lower semicontinuous,

$$\begin{aligned} f(\bar{y})+g(\bar{z}) &\leq \liminf_{k\to\infty}(f(y_{n_k}) + g(z_{n_k})) \\ &\leq \liminf_{k\to\infty}(\inf f + \inf g + 2^{-n_k}) \\ &= \inf f + \inf g. \end{aligned}$$

Hence, $\bar{y}$ minimizes f and $\bar{z}$ minimizes g, so that $\bar{x} = \bar{y}+\bar{z}$ belongs to $\arg\min f + \arg\min g$. $\Box$

EXAMPLE. Let $(X, \|\cdot\|)$ be a normed vector space. Suppose that $f\colon X \to \mathbb{R} \cup \{+\infty\}$ is lower semicontinuous with $\inf f \in \mathbb{R}$. Suppose moreover that $\psi\colon \mathbb{R}_+ \to \mathbb{R}_+$ is continuous nondecreasing with $\psi(0) = 0$ and $\psi(r) > 0$ if $r > 0$. Put $f_\psi = f\Box(\psi \circ \|\cdot\|)$. Then, by statements (a) and (d) of the previous proposition,

$$\inf f_\psi = \inf f \text{ and } \arg\min f_\psi = \arg\min f.$$

This means that from the perspective of minimizing f we might as well work with the function f_ψ. The latter function enjoys frequently more regularity than f, see for instance [80].

Theorem 2.4 *Let f and g be extended real-valued functions on X. If a is an additive real-valued function on X, then*

$$\sup(a - f\Box g) = \sup(a-f) \dot{+} \sup(a-g).$$

PROOF. The proof consists of the following calculation:

$$\begin{aligned} \sup_{x\in X}(a(x) - (f\Box g)(x)) &= \sup_{y,z\in X}(a(y+z) - (f(y) \dot{+} g(z))) \\ &= \sup_{y\in X}(a(y) - f(y)) \dot{+} \sup_{z\in X}(a(z) - g(z)). \end{aligned}$$

We have arrived at the desired identity. $\Box$

2.2 Infimal convolution of subadditive functions

We say that a function $f: X \to \bar{\mathbb{R}}$ is *subadditive* if

$$f(y+z) \leq f(y) \dot{+} f(z) \text{ for all } y, z \in X.$$

It turns out that subadditivity admits a nice characterization in terms of and behaves well with respect to infimal convolution. In the following theorem, $f \wedge g$ denotes the *meet* of the extended real-valued functions f and g on X, $(f \wedge g)(x) = \min\{f(x), g(x)\}$ for each $x \in X$.

Theorem 2.5 *Let $f, g: X \to \bar{\mathbb{R}}$. Then the following statements hold:*

(a) *f is subadditive if and only if $f \square f \geq f$. If $f(0) = 0$, then f is subadditive if and only if $f \square f = f$.*

(b) *If $f \wedge g$ is subadditive, then*

$$f \square g \geq f \wedge g.$$

If, in addition, $f(0) = 0 = g(0)$, then

$$f \square g = f \wedge g.$$

(c) *If f and g are both subadditive, then $f \square g$ is subadditive. If, in addition, $f(0) = 0 = g(0)$, then $f \square g$ is the largest subadditive minorant of $f \wedge g$; in particular,*

$$f \square g = f \wedge g \Leftrightarrow f \wedge g \text{ is subadditive.}$$

PROOF. Part (a) is self-evident [63, Section 6e].

(b) If $f \wedge g$ is subadditive, then

$$f \square g \geq (f \wedge g) \square (f \wedge g) \geq f \wedge g.$$

Suppose that $f(0) = 0 = g(0)$. Then $f \square g \leq f \wedge g$. Indeed,

$$(f \square g)(x) \leq \min\{f(x) + g(0), f(0) + g(x)\} = (f \wedge g)(x) \text{ for all } x \in X.$$

(c) Since f and g are subadditive, $f \square f \geq f$ and $g \square g \geq g$. Thus,

$$(f \square g) \square (f \square g) = (f \square f) \square (g \square g) \geq f \square g.$$

Hence, by (a), $f \square g$ is subadditive. Suppose that $f(0) = 0 = g(0)$. Then $f \square g \leq f \wedge g$. Let h be a subadditive minorant of $f \wedge g$. This means that $h \leq h \square h$ and $h \leq f \wedge g$. Then

$$h \leq h \square h \leq f \square g.$$

We conclude that $f \Box g$ is the largest subadditive minorant of $f \wedge g$. $\Box$

EXAMPLE. Let X be a group, $g: X \to \mathbb{R}$ be additive and $f: X \to \mathbb{R} \cup \{+\infty\}$ be proper. Then

$$(f \Box g)(x) = g(x) + \inf(f - g) \text{ for all } x \in X,$$

so that $f \Box g$ is either equal to the constant $-\infty$ or equal to g plus a real constant. If also f is real-valued and additive with $f \neq g$, then $f \Box g$ is everywhere equal to $-\infty$.

EXAMPLE. Let A and B be two nonempty subsets of the normed vector space $(X, \|\cdot\|)$. Then $(\delta_A \Box \|\cdot\|)(x) = d(x, A)$, the distance from x to A, for each $x \in X$. Moreover, $d(\cdot, A) \Box d(\cdot, B) = d(\cdot, A + B)$, because

$$(\delta_A \Box \|\cdot\|) \Box (\delta_B \Box \|\cdot\|) = (\delta_A \Box \delta_B) \Box (\|\cdot\| \Box \|\cdot\|) = \delta_{A+B} \Box \|\cdot\|.$$

Here we have used the fact that $\|\cdot\|$ is subadditive.

EXAMPLE. Let $f, g: \mathbb{R}_+ \to \mathbb{R}_+$ be nondecreasing, and continuous at the origin with $f(0) = 0 = g(0)$. Suppose that $x \mapsto f(x)/x$ and $x \mapsto g(x)/x$ are nonincreasing on $\mathbb{P}$. (This is the case for instance if f and g are concave.) Then f and g are continuous and subadditive. Moreover, $f \wedge g$ is nondecreasing and $x \mapsto (f \wedge g)(x)/x$ is nonincreasing on $\mathbb{P}$ and, consequently, $f \wedge g$ is subadditive. Thus, by Theorem 2.5 (b), $f \Box g = f \wedge g$.

EXAMPLE. Let X be a vector space. Suppose that both f and g are seminorms on X, and put

$$A = \{y \in X \mid f(y) < 1\} \text{ and } B = \{z \in X \mid g(z) < 1\}.$$

Then, $f \Box g$ is a seminorm. Actually, $f \Box g$ is equal to $\mu_{\mathrm{co}(A \cup B)}$, the Minkowski functional of the convex hull of $A \cup B$.

PROOF. According to Theorem 2.5 (c), $f \Box g$ is subadditive. The remaining axiom for a seminorm is easily checked.

We have $f = \mu_A$ and $g = \mu_B$. Put

$$C = \{x \in X \mid (f \Box g)(x) < 1\} \cup A \cup B.$$

Then

$$\{x \in X \mid (f \Box g)(x) < 1\} \subseteq C \subseteq \{x \in X \mid (f \Box g)(x) \leq 1\}$$

and, hence, $f \Box g = \mu_C$ (see [75, Theorems 1.34 &1.35]), so it suffices to show that $C = \mathrm{co}(A \cup B)$. By the definition of $f \Box g$ and C,

$$\begin{aligned} \mathrm{co}(A \cup B) &= \bigcup_{t \in [0,1]} (tA + (1-t)B) \\ &= \bigcup_{t \in]0,1[} (tA + (1-t)B) \cup A \cup B \end{aligned}$$

$$
\begin{aligned}
&= \bigcup_{t\in\,]0,1[} \left(t\{y \in X \mid f(y) < 1\} + (1-t)\{z \in X \mid g(z) < 1\}\right) \cup A \cup B \\
&= \bigcup_{t\in\,]0,1[} \left(\{y \in X \mid f(y) < t\} + \{z \in X \mid g(z) < 1-t\}\right) \cup A \cup B \\
&= \{x \in X \mid (f\Box g)(x) < 1\} \cup A \cup B \\
&= C,
\end{aligned}
$$

and the proof is complete. □

2.3 Upper semicontinuity, and uniform continuity

A main feature of infimal convolution is its regularizing effect. Namely, frequently under very weak assumptions on f, some types of regularity properties of g are carried over to $f\Box g$. Traditionally, this has been one of the main reasons for the study of infimal convolution.

Theorem 2.6 *Let X be a topological additive group, and $f, g\colon X \to \bar{\mathbb{R}}$. If g is upper semicontinuous, then $f\Box g$ is upper semicontinuous and equal to $(\operatorname{cl} f)\Box g$.*

PROOF. Upper semicontinuity of g means that $\operatorname{epi}_s g$ is an open subset of $X \times \mathbb{R}$. Therefore, $\operatorname{epi}_s f\Box g$ is open, being the sum of $\operatorname{epi}_s f$ and $\operatorname{epi}_s g$. In other words, $f\Box g$ is upper semicontinuous.

Since,

$$\operatorname{epi}_s f \subseteq \operatorname{epi}_s(\operatorname{cl} f) \subseteq \operatorname{epi}(\operatorname{cl} f) = \overline{\operatorname{epi} f} = \overline{\operatorname{epi}_s f},$$

we have

$$\operatorname{epi}_s f + \operatorname{epi}_s g \subseteq \operatorname{epi}_s(\operatorname{cl} f) + \operatorname{epi}_s g \subseteq \overline{\operatorname{epi}_s f} + \operatorname{epi}_s g. \tag{2.1}$$

We actually have equality throughout (2.1) because $\operatorname{epi}_s g$ is an open set. (Write $A = \operatorname{epi}_s f$ and $\mathcal{O}$ for the open set $\operatorname{epi}_s g$. Then, $A + \mathcal{O} = \bar{A} + \mathcal{O}$.) In particular,

$$\operatorname{epi}_s f + \operatorname{epi}_s g = \operatorname{epi}_s(\operatorname{cl} f) + \operatorname{epi}_s g,$$

which is the desired identity $f\Box g = (\operatorname{cl} f)\Box g$. □

In our next statement we consider uniform continuity. If g is a real-valued function defined on a topological vector space X, then we may associate with it the function $m_g\colon X \to \mathbb{R}_+ \cup \{+\infty\}$ defined by the assignments $m_g(0) = 0$ and

$$m_g(x) = \sup\{|g(v) - g(w)| \mid v, w \in X, \text{ and } v - w = x\} \text{ for each } x \in X \setminus \{0\}.$$

As is readily verified, a necessary and sufficient condition for g to be uniformly continuous is that m_g be continuous at 0. The function m_g may be thought of as a modulus of uniform continuity for g. Note that the obvious inequality

$$|g(v) - g(w)| \le m_g(v - w) \text{ for all } v, w \in X,$$

means that $g\Box m_g = g$. Moreover, for any $x, y \in X$,

$$\begin{aligned} m_g(x+y) &= \sup_{w\in X} |g((x+y)+w) - g(w)| \\ &= \sup_{w\in X} |(g(x+(y+w)) - g(y+w)) + (g(y+w) - g(w))| \\ &\le \sup_{w\in X} |g(x+(y+w)) - g(y+w)| + \sup_{w\in X} |g(y+w) - g(w)| \\ &= m_g(x) + m_g(y), \end{aligned}$$

that is, m_g is subadditive. It is readily checked that m_g is symmetric: $m_g(-x) = m_g(x)$ for all $x \in X$.

Theorem 2.7 *Let X be a topological vector space. Suppose that $g: X \longrightarrow \mathbb{R}$ is uniformly continuous and that $f: X \longrightarrow \mathbb{R} \cup \{+\infty\}$ is proper. The following hold:*

(a) *m_g is nowhere equal to $+\infty$.*

(b) *Either $f\Box g$ is the constant $-\infty$, or $f\Box g$ is real-valued and uniformly continuous with $m_{f\Box g} \le m_g$.*

PROOF. (a) Let

$$V = \{x \in X |\, m_g(x) < 1\},$$

which is a neighborhood of zero since m_g is continuous at zero with $m_g(0) = 0$. If n is a positive integer and $x \in nV$, then, by the subadditivity of m_g, $m_g(x) \le n m_g(x/n) < n$. Therefore, m_g is bounded by n on nV which together with the fact that X is equal to $\bigcup\{nV |\, n = 1, 2, 3, \ldots\}$ show that m_g cannot take on the value of $+\infty$.

(b) We have

$$(f\Box g)\Box m_g = f\Box(g\Box m_g) = f\Box g. \tag{2.2}$$

Suppose that the value of $f\Box g$ is $-\infty$ at some point. Then $(f\Box g)\Box m_g$ is everywhere $-\infty$ since m_g is nowhere $+\infty$. Hence, by (2.2), $f\Box g$ is the constant $-\infty$. If $f\Box g$ is nowhere $-\infty$, then it must be real-valued. Then (2.2) implies that

$$\forall x, y \in X \quad (f\Box g)(x) \le (f\Box g)(y) + m_g(x-y),$$

so that (m_g is symmetric)

$$\forall x, y \in X \quad |(f\Box g)(x) - (f\Box g)(y)| \le m_g(x-y).$$

The previous inequality shows that $f\Box g$ is uniformly continuous; in fact, "at least as uniformly continuous as g" in the sense that $m_{f\Box g} \le m_g$. □

Corollary 2.1 *Let X be a normed vector space. Suppose that $g: X \longrightarrow \mathbb{R}$ is Hölder continuous with exponent p, $p \in]0, 1]$, and constant k, and that $f: X \longrightarrow \mathbb{R} \cup \{+\infty\}$ is proper. Then $f\Box g$ is either equal to the constant $-\infty$ or real-valued and Hölder continuous with exponent p and constant k.*

PROOF. The continuity assumption on g means that $m_g \leq k\|\cdot\|^p$, $\|\cdot\|$ being the norm on X. Hence, by the previous proposition, $f\Box g$ is either the constant $-\infty$ or real-valued with $m_{f\Box g} \leq k\|\cdot\|^p$. □

Infimal convolution of Lipschitz continuous functions has been thoroughly studied in [35, 36].

2.4 Exactness, lower semicontinuity, and continuity

Recall that the infimal convolution of f with g is said to be exact at x if the infimum (1.1) is achieved, and exact if it is exact at each element of X. The following theorem considers the case where X is the real line. It is a slight modification of a statement by P. D. Lax [50].

Theorem 2.8 *Let f be a proper extended real-valued function and g a real-valued strictly convex function, both defined on $\mathbb{R}$. Put*

$$Y(x; f, g) = \{y \in \mathbb{R} |\ (f\Box g)(x) = f(y) + g(x-y)\} \text{ for each } x \in \mathbb{R}.$$

Then $x \mapsto Y(x; f, g)$ is a monotone mapping. In particular, $Y(x; f, g)$ has at most one element for every real x outside a subset which is at most countable.

REMARK. Evidently, to say that the infimal convolution $f\Box g$ is exact at x is the same as to say that $Y(x; f, g) \neq \emptyset$.

PROOF. Suppose that x_1 and x_2 are distinct real numbers such that the sets $Y(x_1; f, g)$ and $Y(x_2; f, g)$ are both nonempty. Take $y_i \in Y(x_i, f, g)$ for $i \in \{1, 2\}$. Then

$$\begin{aligned} f(y_1) + g(x_1 - y_1) &\leq f(y_2) + g(x_1 - y_2), \\ f(y_2) + g(x_2 - y_2) &\leq f(y_1) + g(x_2 - y_1). \end{aligned}$$

Note that the numbers $f(y_1)$ and $f(y_2)$ are finite since $(f\Box g)(x_1)$ and $(f\Box g)(x_2)$ are finite. By addition of the above inequalities, we get the inequality

$$g(x_2 - y_2) - g(x_1 - y_2) \leq g(x_2 - y_1) - g(x_1 - y_1). \tag{2.3}$$

We may assume that $x_1 < x_2$. We shall show that $y_1 \leq y_2$. Inequality (2.3) may be written in the form

$$g((x_1 - y_2) + (x_2 - x_1)) - g(x_1 - y_2) \leq g((x_1 - y_1) + (x_2 - x_1)) - g(x_1 - y_1),$$

from which it follows that $x_1 - y_2 \leq x_1 - y_1$ and, hence, $y_1 \leq y_2$. We have used the fact that the function $x \mapsto g(x + h) - g(x)$, where $h = x_2 - x_1 > 0$, is strictly increasing. □

Suppose that X is a topological space. If $f: X \to \mathbb{R} \cup \{+\infty\}$ is proper lower semicontinuous and has a compact nonempty sublevel set then f has a minimum value. This classical and extremely important result due to Weierstrass may be used to formulate conditions ensuring the infimal convolution to be exact. Moreover, under certain compactness conditions, it is possible to prove that $f \Box g$ is lower semicontinuous. Part (a) of the following theorem is due to J.-J. Moreau [62, Section 4e].

Theorem 2.9 *Let X be a Hausdorff topological additive group. Assume that $g: X \to \mathbb{R} \cup \{+\infty\}$ is proper and has compact sublevel sets, and that $f: X \to \mathbb{R} \cup \{+\infty\}$ is proper and bounded from below.*

(a) *Suppose that f is lower semicontinuous. Then the infimal convolution $f \Box g$ is exact, and the infimal convolute $f \Box g$ is lower semicontinuous. If both f and g have compact sublevel sets, then $f \Box g$ has compact sublevel sets.*

(b) *Suppose that g is continuous. Then $f \Box g$ is continuous.*

PROOF. For a proof of part (a), see [49, 62]. (b) By Theorem 2.6, $f \Box g$ is upper semicontinuous and equal to $f \Box (\operatorname{cl} g)$. Moreover, by (a), $f \Box (\operatorname{cl} g)$ is lower semicontinuous. Thus, $f \Box g$ is continuous. □

Corollary 2.2 *Let X be a Eucledian space. Let $g: X \to \mathbb{R} \cup \{+\infty\}$ be proper, lower semicontinuous and weakly coercive. Let $f: X \to \mathbb{R} \cup \{+\infty\}$ be proper and bounded from below.*

(a) *Suppose that f is lower semicontinuous. Then the infimal convolution $f \Box g$ is exact, and $f \Box g$ is lower semicontinuous. If, in addition, f is weakly coercive, then $f \Box g$ is weakly coercive.*

(b) *Suppose that g is continuous. Then $f \Box g$ is continuous.*

PROOF. The assumption $g(x) \to +\infty$ as $|x| \to +\infty$ means that g has bounded sublevel sets. Moreover, the hypothesis that g is lower semicontinuous means that g has closed sublevel sets. Hence, $\operatorname{slv}(g; \alpha)$ is compact for each $\alpha \in \mathbb{R}$. Thus, the previous theorem may be applied, and it yields the conclusions of (a) and (b). □

Corollary 2.3 *Let X be a reflexive Banach space. Let $f, g: X \to \mathbb{R} \cup \{+\infty\}$ be proper, weakly coercive, lower semicontinuous with convex sublevel sets. Then the infimal convolution $f \Box g$ is exact, and $f \Box g$ is weakly lower semicontinuous and weakly coercive.*

PROOF. Our assumptions imply that f and g have bounded closed convex, hence weakly compact, sublevel sets. Therefore, the proof follows by applying Theorem 2.9 with X supplied with its weak topology. □

EXAMPLE. The coerciveness assumption in Corollary 2.2 is essential as witnessed by the following calculation. Let f and g be defined on the real line by

$$f(x) = \sin(\pi x^2) \text{ and } g(x) = -\sin(\pi x^2), \text{ for all } x \in \mathbb{R}.$$

Then f and g are bounded infinitely smooth functions. We claim that $f \Box g$ fails to be lower semicontinuous at the origin. Obviously, $(f \Box g)(0) = 0$. Put, for each $n \in \mathbb{N}$,

$$y_n = (2n + \tfrac{3}{2})^{1/2}, \ z_n = -(2n + \tfrac{1}{2})^{1/2}, \text{ and } x_n = y_n + z_n.$$

Then, $x_n \to 0$ as $n \to \infty$, and

$$(f \Box g)(x_n) \leq f(y_n) + g(z_n) = -2 \text{ for each } n \in \mathbb{N}.$$

Our claim follows.

The continuity statement in Corollary 2.2 does not really depend on compactness of sublevel sets, but rather on growth conditions and uniform continuity on bounded sets.

Theorem 2.10 *Let X be a normed space. Suppose that $g: X \to \mathbb{R}$ is weakly coercive, and uniformly continuous on each bounded subset of X, and that $f: X \to \mathbb{R} \cup \{+\infty\}$ is proper and bounded from below. Then $f \Box g$ is real-valued and uniformly continuous on bounded sets.*

PROOF. See T. Strömberg [80, Proposition 2]. □

For more on infimal convolution of functions that are uniformly continuous on bounded sets, see [21, 33, 69, 80].

We have already presented some sets of functions that are closed under the operation of infimal convolution. A few more sets of functions with this property are given in our final statement in this chapter.

Theorem 2.11 *Let X be a Hausdorff topological additive group. The following subsets of $\bar{\mathbb{R}}^X$ are closed under infimal convolution:*

(i) *The set of all functions $f: X \to \bar{\mathbb{R}}$ having connected strict sublevel sets (that is, $\{x \in X \mid f(x) < \alpha\}$ is connected for each $\alpha \in \mathbb{R} \cup \{+\infty\}$);*

(ii) *The set of all proper lower semicontinuous functions $X \to \mathbb{R} \cup \{+\infty\}$ having compact connected sublevel sets;*

(iii) *The set of all proper lower semicontinuous functions $X \to \mathbb{R} \cup \{+\infty\}$ having compact connected sublevel sets and for which every local minimizer is global.*

In (i) *and* (ii), *the adjective "connected" may be replaced by "path-connected."*

PROOF. See D. H. Martin [55]. □

Chapter 3

The convex case

In this chapter we focus our attention on the basic semicontinuity, continuity, and differentiability properties of the infimal convolute of two convex functions. In particular, we present dual formulations for infimal convolution. Moreover, we exhibit smoothing properties of the operation, and point out that the regularization feature is in general up to some limit.

3.1 Basic results

We have collected the most elementary properties of infimal convolution of convex or concave functions in our next proposition. They are of course well-known, see J.-J. Moreau [62, Section 3f], [63, Section 9d].

Theorem 3.1 *Let X be a vector space and $f, g\colon X \to \bar{\mathbb{R}}$.*

(a) $f \Box f = 2 \bullet f \Leftrightarrow f$ *is midpoint-convex:* $f(\frac{1}{2}y + \frac{1}{2}z) \leq \frac{1}{2}f(y) \dot{+} \frac{1}{2}f(z)$ *for all* $y, z \in X$.

(b) $(s \bullet f) \Box (t \bullet f) = (s + t) \bullet f$ *for all* $s, t \in \mathbb{P} \Leftrightarrow f$ *convex.*

(c) *If both f and g are convex, then so is $f \Box g$.*

(d) *If f or g is concave, then $f \Box g$ is concave.*

PROOF. Statements (a) and (b) are selfevident [62, Section 3f], [63, Section 9d].

(c) The strict epigraph of $f \Box g$ is convex since it is the sum of the convex strict epigraphs of f and g.

(d) If g is concave, then $x \mapsto f(y) \dot{+} g(x - y)$ is a concave function for each $y \in X$. Consequently, $f \Box g$ is concave since it is the meet of a set of concave functions. $\Box$

Let X be a locally convex topological vector space. The infimal convolute of two convex functions is always convex, but we should be aware of the fact that $\Gamma(X)$ is not stable under the operation of infimal convolution because $f \Box g$, with $f, g \in \Gamma(X)$, need not be lower semicontinuous or even proper.

EXAMPLE. The following hold:

(a) If x^* and y^* are two distinct elements of $X^* \subseteq \Gamma(X)$, then $x^* \Box y^*$ is the constant $-\infty$.

(b) Let A and B be the epigraphs of the zero function and the exponential function $x \mapsto e^x$ on $\mathbb{R}$, respectively. Then $\delta_A, \delta_B \in \Gamma(\mathbb{R}^2)$ but, nevertheless, $\delta_A \Box \delta_B = \delta_{A+B}$ is not lower semicontinuous because $A + B$ is the open upper half-plane.

If f and g share a common affine minorant, then $f \Box g$ is proper. In particular, the condition $\operatorname{dom} f^* \cap \operatorname{dom} g^* \neq \emptyset$ forces $f \Box g$ to be proper.

Theorem 3.2 *If X is a locally convex topological vector space, then*

$$(f \Box g)^* = f^* \dotplus g^* \text{ for any } f, g: X \to \bar{\mathbb{R}}.$$

PROOF. See Theorem 2.4. □

This means that infimal convolution is the dual operation of (lower) addition. This is important in the duality theory of convex optimization.

An immediate consequence of Theorem 3.2 is the inequality

$$(f^* \dotplus g^*)^* = (f \Box g)^{**} \leq f \Box g.$$

A main problem is to find sufficient conditions on convex proper functions f and g to guarantee $f \Box g$ to be lower semicontinuous. The two functions $f \Box g$ and $(f^* \dotplus g^*)^*$, are equal if and only if $f \Box g$ is lower semicontinuous. Similarly, if $f, g \in \Gamma(X)$, then $(f + g)^* = f^* \Box g^*$ if and only if $f^* \Box g^*$ is weak* lower semicontinuous.

The following two theorems state basic semicontinuity or continuity properties of $f \Box g$.

Theorem 3.3 *Let X be a locally convex Hausdorff topological vector space and let f and g be be convex proper functions $X \to \mathbb{R} \cup \{+\infty\}$ with $\operatorname{dom} f^* \cap \operatorname{dom} g^* \neq \emptyset$. Suppose that f or g is real-valued and continuous at some point. Then the interior of $\operatorname{dom} f + \operatorname{dom} g$ is nonempty, and $f \Box g$ is a proper function which is continuous on $\operatorname{int}(\operatorname{dom} f + \operatorname{dom} g)$. In particular, if $\operatorname{dom} f + \operatorname{dom} g = X$, then $f \Box g$ is real-valued and continuous. Moreover, if x belongs to the interior of $\operatorname{dom} f + \operatorname{dom} g$, then*

$$(f \Box g)(x) = \max_{x^* \in X^*} (\langle x^*, x \rangle - f^*(x^*) - g^*(x^*)), \tag{3.1}$$

and $\partial(f \Box g)(x)$ is equal to the (nonempty) set of maximizers of (3.1).

PROOF. For a complete proof, see [81, Proposition 1]. Compare with [62, Section 9b]. □

Theorem 3.4 *Assume that X is a locally convex Hausdorff topological vector space. Let $f, g: X \to \mathbb{R} \cup \{+\infty\}$ be proper convex functions.*

(a) *If* $\operatorname{dom} f \cap \operatorname{dom} g$ *contains a point at which f or g is continuous, then*

$$(f+g)^* = f^* \square g^*$$

with exact infimal convolution.

(b) *Suppose that X is reflexive. Suppose f and g belong to $\Gamma(X)$. If* $\operatorname{dom} f^* \cap \operatorname{dom} g^*$ *contains a point at which f^* or g^* is continuous, then*

$$f \square g = (f^* + g^*)^*$$

with exact infimal convolution. Consequently, $f \square g \in \Gamma(X)$.

PROOF. See R. T. Rockafellar's paper [71]. □

REMARK. The conclusion of part (b) may fail in a nonreflexive space. The following counter example is due to H. Attouch and H. Brezis [6]. Let $(X, \|\cdot\|)$ be a normed space. Suppose for the moment that X is nonreflexive. Then there exists a closed hyperplane H in X such that $1 = \inf_{x \in H} \|x\|$ is not achieved. Put $f = \delta_U$ and $g = \delta_H$. Then f^* is real-valued and continuous and nevertheless $f \square g \neq (f^* + g^*)^*$. Indeed, $f \square g = \delta_{U+H}$ is not lower semicontinuous because $U + H$ is not closed: $0 \notin U + H$ but $0 \in \operatorname{cl}(U + H)$.

Recently, Rockafellar's theorem has been formulated for cs-convex functions defined on a Fréchet space.

Theorem 3.5 *Let X be a Fréchet space. Suppose f and g are proper cs-convex functions $X \to \mathbb{R} \cup \{+\infty\}$ such that*

$$\operatorname{cone}(\operatorname{dom} f - \operatorname{dom} g) \text{ is a closed linear subspace of } X.$$

Then the infimal convolution $f^* \square g^*$ is exact and

$$(f+g)^* = f^* \square g^*.$$

REMARK. The above result was proved for the case where X is a Banach space and the functions f, g elements of $\Gamma(X)$ by H. Attouch and H. Brezis in [6]. It was generalized by B. Rodrigues and S. Simons to the case where X is an arbitrary Fréchet space in [73, 74], and to the form stated above by S. Simons [77]. See also M. Volle [82, 85].

Corollary 3.1 *Suppose X to be a reflexive Banach space and $f, g \in \Gamma(X)$. If*

$$\operatorname{cone}(\operatorname{dom} f^* - \operatorname{dom} g^*) \text{ is a closed linear subspace of } X^*,$$

then the infimal convolution $f \square g$ is exact and $f \square g$ belongs to $\Gamma(X)$.

3.2 Differential calculus, and first order differentiability

We start with an elementary example.

EXAMPLE. Let f and g be the quadratic convex functions $x \mapsto ax^2/2$ and $x \mapsto bx^2/2$ on the real line, where $a, b \in \mathbb{P}$. Then $f \square g$ is equal to the function $x \mapsto cx^2/2$, where $c = (a^{-1} + b^{-1})^{-1}$. The number $(a^{-1} + b^{-1})^{-1}$ is sometimes called the "parallel sum" of a and b. This terminology is motivated by the parallel connection of resistors in electrical networks, see W. N. Andersson and R. J. Duffin [1]. Identifying $f' \leftrightarrow a$, and similarly $g' \leftrightarrow b$ and $(f \square g)' \leftrightarrow c$, we may say that $(f \square g)'$ is the parallel sum of f' and g'.

If S and T are subsets of $X \times X^*$, considered as multivalued mappings from X into X^*, then we write $S//T$ for their *parallel sum* defined by

$$S//T = (S^{-1} + T^{-1})^{-1}.$$

Put differently

$$S//T = \{(x, x^*) \in X \times X^* \mid \exists y, z \in X\colon y + z = x, \text{ and } x^* \in S(y) \cap T(z)\}.$$

See G. B. Passty [68] for properties of parallel addition of multivalued mappings. Parallel addition has also been studied in [1, 30, 38, 46, 56, 65, 76].

Theorem 3.6 *Let X be a locally convex Hausdorff topological vector space. Let f, g be proper functions $X \to \mathbb{R} \cup \{+\infty\}$, bounded from below by the same affine function. Let $x \in \operatorname{dom} f + \operatorname{dom} g$.*

(a) *Then*

$$\partial(f \square g)(x) = \bigcap_{\varepsilon \in \mathbb{P}} \bigcup_{y \in X} (\partial_\varepsilon f(y) \cap \partial_\varepsilon g(x - y)).$$

More precisely, for any $\varepsilon \in \mathbb{P}$,

$$\partial(f \square g)(x) \subseteq \bigcup_{y \in X} (\partial_\varepsilon f(y) \cap \partial_\varepsilon g(x - y)) \subseteq \partial_{2\varepsilon}(f \square g)(x).$$

(b) *If the infimal convolution $f \square g$ is exact at x, say $(f \square g)(x)$ is equal to $f(y) + g(x - y)$ for some $y \in X$, then $\partial(f \square g)(x) = \partial f(y) \cap \partial g(x - y)$.*

(c) *If $\partial f(y) \cap \partial g(x - y)$ is nonempty for some $y \in X$, then $f \square g$ is exact at x with $(f \square g)(x) = f(y) + g(x - y)$ and $\partial(f \square g)(x) = \partial f(y) \cap \partial g(x - y)$.*

(d) *Suppose that $f \square g$ is exact. Then $\partial(f \square g) = \partial f // \partial g$.*

(e) *If* $(f \Box g)(x) = (f^* + g^*)^*(x)$ *and* $\operatorname{dom} f^* \cap \operatorname{dom} g^* \neq \emptyset$, *then*

$$\partial(f \Box g)(x) = \arg\max\{\langle x^*, x\rangle - f^*(x^*) - g^*(x^*) | \, x^* \in X^*\}.$$

PROOF. For a proof of (a) see the recent article [39] by J.-B. Hiriart-Urruty and R. R. Phelps; for the remaining statements see [49, 68, 79]. □

REMARK. The set $\partial f(y) \cap \partial g(x-y)$ in (b) may be empty.

EXAMPLE. Let X be the real line and let f and g be the exponential function $x \mapsto e^x$ and the zero function, respectively. Then $f \Box g$ is the zero function. Let x be an arbitrary point in $\mathbb{R}$. Then $\partial(f \Box g)(x) = \{0\}$, while

$$\partial f(y) \cap \partial g(x-y) = \emptyset \text{ for all } y \in \mathbb{R}.$$

This is not in conflict with statement (b) of the previous theorem since $f \Box g$ is not exact at x.

As mentioned earlier, the operation of infimal convolution acts smoothingly in several situations.

Theorem 3.7 *Let X be a normed space. Let $f \in \Gamma(X)$, and let $g: X \to \mathbb{R}$ be convex continuous with* $\operatorname{dom} g^* = X^*$.

(a) *If g^* is strictly convex, then $f \Box g$ is Gâteaux differentiable.*

(b) *Suppose that g^* is locally uniformly convex in the sense that*

$$\liminf_{n\to\infty}(g^*(\tfrac{1}{2}x^* + \tfrac{1}{2}x_n^*) - \tfrac{1}{2}g^*(x^*) - \tfrac{1}{2}g^*(x_n^*)) \geq 0 \textit{ implies } \lim_{n\to\infty} \|x^* - x_n^*\|_* = 0$$

for any sequence $\langle x_n^*\rangle$ in X^* and any $x^* \in X^*$. Then $f \Box g$ is Fréchet differentiable.

(c) *Let X be a reflexive Banach space. Suppose that $g: X \to \mathbb{R}$ is convex and Fréchet (respectively, Gâteaux) differentiable. Let $f \in \Gamma(X)$. Then $f \Box g$ is Fréchet (respectively, Gâteaux) differentiable with*

$$D(f \Box g) = \partial f // Dg \textit{ (respectively, } \nabla(f \Box g) = \partial f // \nabla g).$$

PROOF. First of all we note (Theorem 3.3) that $f \Box g$ is continuous real-valued with

$$\partial(f \Box g)(x) = \arg\max\{\langle x^*, x\rangle - f^*(x^*) - g^*(x^*) | \, x^* \in X^*\} \neq \emptyset \text{ for each } x \in X.$$

(a) Since g^* is strictly convex, $\partial(f \Box g)(x)$ is a singleton for each $x \in X$ and, hence, $f \Box g$ is Gâteaux differentiable.

For a proof of (b), see [81, Theorem 2].

(c) Let $x \in X$. There exists $y \in X$ such that

$$\partial(f \Box g)(x) = \partial f(y) \cap \partial g(x-y).$$

Since $\partial g(x-y)$ is a singleton it yields that $\partial(f \Box g)(x)$ is also a singleton. The choice of x being arbitrary, we may conclude that $f \Box g$ is Gâteaux differentiable. If g is Fréchet differentiable, then so is $f \Box g$, see [81, Theorem 1]. □

Results of this type are particularly important for regularization of convex functions.

3.3 Higher order differentiability, and the regularity of the sum of convex sets

We have already observed that infimal convolution has a regularizing effect in several situations, at least in the convex case. But we shall see in this section that we cannot in general expect that this applies to higher than C^1-regularity, even when the dimension of the underlying space X is finite-dimensional. Several quite surprising and remarkable results in this direction have been obtained by J. Boman [19, 20] and C. O. Kiselman [43, 44] in their investigations of the regularity of the sum of smooth convex subsets of low-dimensional Eucledian spaces. With the fact in mind that infimal convolution is the operation of addition of (strict) epigraphs, it should come as no surprise that the study of the sum $A+B$ of two closed convex sets $A, B \subseteq \mathbb{R}^2$ can be reduced to the investigation of infimal convolution of convex functions defined on $\mathbb{R}$. A precise statement reads as follows:

Theorem 3.8 *Let A and B be two closed convex sets in $\mathbb{R}^2$ with C^1 boundaries, and consider a point c on the boundary of $A+B$. If c happens to belong to $A+B$, say $c = a+b$ with $a \in A$, $b \in B$, then there exists an affine coordinate system and convex functions f and g such that A, B and $A+B$ agree with* $\operatorname{epi} f$, $\operatorname{epi} g$ *and* $\operatorname{epi} f \Box g$ *near a, b and c, respectively. If, on the other hand, $c \notin A+B$, then the boundary of $A+B$ contains an entire straight line through c, so that $A+B$ is either a half-plane or a strip.*

PROOF. See C. O. Kiselman [44, Proposition 2.1]. □

Theorem 3.9 *The following two assertions hold:*

(a) *Suppose that $f, g\colon \mathbb{R} \to \mathbb{R}$ are convex germs with $f'(0) = 0 = g'(0)$. If $k \in \{1,2,3,4\}$ and $f, g \in C^k$, then $f \Box g \in C^k$.*

(b) *There exist two strictly convex C^∞ germs f and g with $f'(0) = 0 = g'(0)$ such that $f \Box g \notin C^5$.*

PROOF. A proof can be found in J. Boman [19, Propositions 1 & 2]. □

REMARK. Part (a) is stated, without a proof, in C. O. Kiselman [44].

In particular, part (b) of implies that there exist two bounded strictly convex subsets A and B of the plane with C^∞ boundaries such that the boundary of the sum $A + B$ is not C^5.

In higher dimensions, the regularity may drop considerably.

Theorem 3.10 *Let f and g be real analytic strictly convex germs whose values at each point (u, v, w) near the origin in $\mathbb{R}^3$ are*

$$\begin{aligned} f(u,v,w) &= u^2 + v^2 + w^2(u^2 + v^2 + w^2/6), \\ g(u,v,w) &= f(u,v,w) + uvw. \end{aligned}$$

Then the second derivative of $f \square g$ is discontinuous at the origin.

PROOF. See J. Boman [20, Proposition 1]. □

EXAMPLE. (Compare with C. O. Kiselman [44, Example 2.2].) Let f and g be the two functions $x \mapsto x^4/4$ and $x \mapsto x^6/6$, respectively, on the real line. Then, according to Theorem 3.6,

$$(f \square g)'^{-1}(y) = f'^{-1}(y) + g'^{-1}(y) = y^{1/3} + y^{1/5} \text{ for all } y \in \mathbb{R}.$$

An expansion yields

$$(f \square g)'(x) = x^5 - 5x^{17/3} + \frac{55}{3}x^{19/3} + r(x), \text{ as } x \to 0,$$

where $r \in C^7$. We conclude that $(f \square g)'$ belongs to the Hölder class $C^{17/3}$ (near the origin), but to no smaller Hölder class. Therefore, $f \square g \in C^{20/3}$ but $f \square g$ fails to belong to C^p for any $p > 20/3$.

Following C. O. Kiselman [43], if p and $\alpha_1, \ldots, \alpha_m$ are non-negative, rational numbers with odd terminators, and $k \in \mathbb{N} \cup \{\infty\}$, then we write $C_p^k(\alpha_1, \ldots, \alpha_m)$ for the set of all germs of functions f defined on some neighborhood U of the origin in $\mathbb{R}$ and such that

$$f(x) = x^p F(x^{\alpha_1}, \ldots, x^{\alpha_m}) \quad (x \in U)$$

for some function F which is C^k in a neighborhood of the origin in $\mathbb{R}^m$ and satisfying $F(0, \ldots, 0) > 0$.

Theorem 3.11 *The following assertions hold:*

(a) *Let* f *and* g *be germs of differentiable convex functions such that* $f' \in \mathcal{C}_p^k(\alpha_1, \ldots, \alpha_m)$ *and* $g' \in \mathcal{C}_q^k(\beta_1, \ldots, \beta_n)$, *where* $k \in \{1, 2, \ldots, \infty\}$ *and* $\alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_n$ *and* $p, 1/p, q, 1/q$, $p \geq q > 0$, *are all rational numbers with odd denominators. Then*

$$(f \Box g)' \in \mathcal{C}_p^k(\alpha_1, \ldots, \alpha_m, p\beta_1/q, \ldots, p\beta_n/q, (p/q) - 1).$$

(b) *Let* $f_1, f_2, \ldots, f_m$ *be a finite number of* C^∞ *convex functions on* $\mathbb{R}$ *all with the property that, at every point, some derivative of order two or higher does not vanish. The infimal convolute* $f = f_1 \Box f_2 \Box \cdots \Box f_m$ *may be the constant* $-\infty$ *or an affine function. Otherwise, the germ of* f *at any point consists of an affine function plus a germ from one of*

$$\mathcal{C}_p := \mathcal{C}_{p+1}^\infty(1, 2/M(p)),\ p \in \{1, 3, 5, \ldots\},$$

where $M(1) = 1$, *and* $M(p)$, $p \geq 3$, *is the smallest common multiple of the numbers* $1, 3, 5, 7, \ldots, p-2$. *Moreover,* $\mathcal{C}_1 \subseteq C^\infty$, $\mathcal{C}_3 \subseteq C^\infty$, $\mathcal{C}_p \subseteq C^{p+1+2/M(p)}$ *for any* p *in* $\{5, 7, 9, \ldots\}$; *in particular* $\bigcap\{\mathcal{C}_p | \, p \in \{1, 3, 5, \ldots\}\} \subseteq C^{5+1+2/3} = C^{20/3}$.

PROOF. See C. O. Kiselman [44, Theorem 4.3, Corollary 4.4]. □

In particular, if A and B are two convex subsets of the plane with real-analytic boundaries, then their sum $A + B$ must be of class $C^{20/3}$, but no better in general.

The reader may study C. O. Kiselman [45] for other types of regularity classes.

3.4 The occasional distributivity of composition over infimal convolution

The following theorem on distributivity of composition over infimal convolution is due to S. Simons [77].

Theorem 3.12 *The following hold:*

(a) *Let* X *and* Z *be vector spaces,* $X^\#$ *and* $Z^\#$ *their respective algebraic duals. Let* $x^\# \in X^\#$, $B: X \to Z$ *be linear,* $\varphi: Z \to \mathbb{R} \cup \{+\infty\}$ *be convex and proper, and* $((\varphi \circ B) \Box x^\#)(0) \in \mathbb{R}$. *Then there exists* $z^\# \in Z^\#$ *for which*

$$z^\# \circ B = x^\#, \text{ and } (\varphi \Box z^\#) \circ B = (\varphi \circ B) \Box (z^\# \circ B) = (\varphi \circ B) \Box x^\#.$$

(b) *Let* X *and* Z *be metrizable locally convex topological vector spaces. Let* $B: X \to Z$ *be linear and the graph of* B *be a complete subspace of* $X \times Z$. *Suppose that* $\varphi: Z \to \mathbb{R} \cup \{+\infty\}$ *is proper, cs-convex, and that*

$$\operatorname{cone}(\operatorname{dom} \varphi + B[X]) \text{ is a barrelled linear subspace of } Z.$$

Moreover, suppose that $x^* \in X^*$ and $((\varphi \circ B)\Box x^*)(0) \in \mathrm{I\!R}$. Then there exists $z^* \in Z^*$ such that

$$z^* \circ B = x^*, \text{ and } (\varphi \Box z^*) \circ B = (\varphi \circ B)\Box(z^* \circ B) = (\varphi \circ B)\Box x^*.$$

As a corollary we get a version of the celebrated duality theorem of Fenchel.

Corollary 3.2 *Let X and Y be Fréchet spaces and $A: X \to Y$ be linear and continuous. Let $f: X \to \mathrm{I\!R} \cup \{+\infty\}$ and $g: Y \to \mathrm{I\!R} \cup \{+\infty\}$ be proper, cs-convex, and*

$$\operatorname{cone}(\operatorname{dom} f - A[\operatorname{dom} g]) \text{ be a closed subspace of } Y.$$

Then

$$(f + g \circ A)^*(x^*) = \min\{f^*(x^* - y^* \circ A) + g^*(y^*) \mid y^* \in X^*\}$$

for all $x^* \in X^*$.

PROOF. Let $Z = X \times Y$, $B: X \to Z$, $\varphi: Z \to \mathrm{I\!R} \cup \{+\infty\}$ be defined by $Bx = (x, Ax)$ and $\varphi(x,y) = f(x) + g(y)$. Then $\varphi \circ B = f + g \circ A$, and

$$\bigcup\{\lambda(\operatorname{dom}\varphi + B[X]) \mid \lambda \in \mathrm{I\!R}_+\} = L^{-1}\left(\bigcup\{\lambda(\operatorname{dom} f - A[\operatorname{dom} g]) \mid \lambda \in \mathrm{I\!R}_+\}\right),$$

where L denotes the linear continuous map $(x,y) \mapsto y - Ax$ of Z into Y, and hence is a closed subspace. If $x^* \in \operatorname{dom}(\varphi \circ B)^*$, then $((\varphi \circ B)\Box x^*)(0) = -(\varphi \circ B)^*(x^*) \in \mathrm{I\!R}$. We can now apply Theorem 3.12 to get

$$\max\{((\varphi \Box z^*) \circ B)(0) \mid z^* \in Z^*, \text{ and } z^* \circ B = x^*\} = ((\varphi \circ B)\Box x^*)(0),$$

which means, since $((\varphi \Box z^*) \circ B)(0) = (\varphi \Box z^*)(0) = -\varphi^*(z^*)$, that

$$\min\{f^*(x^* - y^* \circ A) + g^*(y^*) \mid y^* \in X^*\} = (f + g \circ A)^*(x^*)$$

for all $x^* \in X^*$. $\Box$

REMARK. If we apply Corollary 3.2 with $Y = X$, and A the identity mapping on X, then we obtain Theorem 3.5.

Chapter 4

Continuity of the operation of infimal convolution

As a general reference in this chapter we recommend G. Beer's recent monograph on hyperspace topologies [15]. For $(X, \|\cdot\|)$ a normed space, we adopt the following notations:

$\mathbf{C}(X)$—the nonempty, closed and convex subsets of X;
$\mathbf{CB}(X)$—the nonempty, closed, convex and bounded subsets of X.

Let A and B be two nonempty subsets of X. The *gap* $D(A, B)$ between A and B is defined by the formula

$$D(A,B) = \inf\{\|a-b\| \mid a \in A, \text{ and } b \in B\} = \inf\{\varepsilon \in \mathbb{P} \mid A \cap (B + \varepsilon U) \neq \emptyset\}.$$

The *excess* of A over B is the number

$$e(A,B) = \sup\{d(a,B) \mid a \in A\} = \inf\{\varepsilon \in \mathbb{P} \mid A + \varepsilon U \subseteq B\}.$$

A popular approach is to represent certain topologies on $\mathbf{C}(X)$ in "hit-and-miss" form, which we proceed to define. With any nonempty subset B of X we associate the following subsets of $\mathbf{C}(X)$:

$$\begin{aligned} B^- &= \{A \in \mathbf{C}(X) \mid A \cap B \neq \emptyset\}; \\ B^+ &= \{A \in \mathbf{C}(X) \mid A \subseteq B\}; \\ B^{++} &= \{A \in \mathbf{C}(X) \mid D(A, B^c) > 0\}. \end{aligned}$$

A set in B^- "hits" B, whereas a set in B^+ "misses" B^c. A set in B^{++} misses B^c in a definite way, in that the set and B^c cannot be asymptotic. Suppose now that Δ is a set whose elements are closed nonempty subsets of X. By the *hit-and-miss topology* determined by Δ, we will mean the topology on $\mathbf{C}(X)$ having as a subbase all sets of the form V^- where V is an open subset of X, and all sets of the form $(B^c)^+$ where $B \in \Delta$. By the *proximal hit-and-miss topology* determined by Δ, we will mean the topology having as a subspace all sets of the form V^- where V is an open subset of X, and all sets of the form $(B^c)^{++}$ where $B \in \Delta$.

Often it is convenient to "split" a hit-and-miss topology, or a proximal hit-and-miss topology, into two parts. Therefore, we define the topology τ^- on $\mathrm{C}(X)$ as the topology generated by

$$\{V^- \mid V \text{ nonempty open subset of } X\}.$$

A proper lower semicontinuous convex function defined on X, identified with its epigraph, is a nonempty closed convex subset of $X \times \mathbb{R}$ that recedes in the vertical direction and contains no vertical lines. Hyperspace topologies on $\mathrm{C}(X \times \mathbb{R})$ naturally induce topologies on $\Gamma(X)$. In this chapter we consider continuity of the operation of infimal convolution on $\Gamma(X)$ with respect to such topologies.

4.1 Epi-convergence

First of all, let us recall the notion of convergence in the sense of Kuratowski and Painlevé of sequences of subsets of X. Given a sequence $\langle A_n \rangle_{n \in \mathbb{N}}$ of subsets of X, then $\liminf_{n\to\infty} A_n$ (respectively, $\limsup_{n\to\infty} A_n$) is, by definition, the set of all $x \in X$ for which every neighborhood of x meets eventually (respectively, frequently) the sets $A_0, A_1, A_2, \ldots$. The sequence $\langle A_n \rangle_{n \in \mathbb{N}}$ is declared convergent to the closed set A in the sense of Kuratowski–Painlevé if

$$\limsup_{n\to\infty} A_n \subseteq A \subseteq \liminf_{n\to\infty} A_n. \tag{4.1}$$

For our purposes we need concepts of convergence of functions: if $\langle f_n \rangle_{n \in \mathbb{N}}$ is a sequence of extended real-valued functions on X, then we say that $\langle f_n \rangle$ *epi-converges* to a lower semicontinuous function f provided that the sequence of epigraphs $\langle \operatorname{epi} f_n \rangle$ converges in the Kuratowski–Painlevé sense to the epigraph of f. Locally, this amounts, at each $x \in X$, to the conjunction of the following two conditions:

(i) There exists a sequence $\langle x_n \rangle$ convergent to x with $\limsup_{n\to\infty} f_n(x_n) \le f(x)$;

(ii) For each sequence $\langle x_n \rangle$ convergent to x, $\liminf_{n\to\infty} f_n(x_n) \ge f(x)$.

For a proof, see [15, Theorem 5.3.5]. Even for continuous functions $[0,1] \to \mathbb{R}$ this mode of convergence is neither stronger nor weaker than pointwise convergence, see for instance [15, page 156]. The notion of epi-convergence is particularly important in the theory of optimization, study the monograph [2] by H. Attouch.

We remark that it is well-known that the Moreau–Yosida approximation scheme, discussed in the final chapter, enables us to reduce the epi-convergence of a sequence of functions to pointwise convergence of the corresponding sequence of regularized functions, consult for instance [2, 21, 86].

The "epi-distance" topology of H. Attouch and R. J.-B. Wets provides in the case where X is finite-dimensional a topology compatible with epi-convergence, even without convexity assumptions [7, 15]. The topology is then first countable, so the following two results for epi-convergence of sequences of infimal convolutes are true continuity results with respect to the Attouch–Wets topology, to which we return in the final section of this chapter.

Theorem 4.1 *Let X be a Eucledian space. Let $\langle f_n\rangle$ and $\langle g_n\rangle$ be two sequences of extended real-valued functions defined on X, epi-convergent to f and g, respectively. Suppose also that*

$$\bigcup\{\operatorname{slv}(f_n;\alpha)\mid n\in\mathbb{N}\} \text{ is bounded for each } \alpha\in\mathbb{R},$$

and that

$$\inf\{g_n(x)\mid (n,x)\in\mathbb{N}\times X\} > -\infty.$$

Then, $\langle f_n\Box g_n\rangle$ epi-converges to $f\Box g$.

PROOF. See H. Attouch and R. J.-B. Wets [7, Proposition 4.2]. □

The following continuity result for infimal convolution of convex functions is due to L. McLinden and R. C. Bergstrom [58, Theorems 1 & 2].

Theorem 4.2 *Let X be a Eucledian space. Suppose that $f, f_0, f_1, \ldots$ and $g, g_0, g_1, \ldots$ are elements of $\Gamma(X)$, such that $\langle f_n\rangle$ epi-converges to f, $\langle g_n\rangle$ epi-converges to g, and $0\in\operatorname{int}(\operatorname{dom} f^* - \operatorname{dom} g^*)$. Then, $f\Box g\in\Gamma(X)$, $f_n\Box g_n\in\Gamma(X)$ for all sufficiently large n, and $\langle f_n\Box g_n\rangle$ epi-converges to $f\Box g$.*

4.2 Mosco convergence

A celebrated notion of convergence of sequences of subsets of X is that of Mosco convergence [66]: a sequence $\langle A_n\rangle_{n\in\mathbb{N}}$ of subsets of a normed X is declared *Mosco convergent* to A provided $\langle A_n\rangle_{n\in\mathbb{N}}$ converges to A in the Kuratowski–Painlevé sense with respect to both the strong and the weak topology of X. By the identification of extended real-valued functions with their epigraphs, we arrive at the concept of Mosco convergence of sequences of extended real-valued functions. Explicitly, a sequence $\langle f_n\rangle$ of extended real-valued functions on X Mosco converges to f, in symbols $\text{M-}\lim_{n\to\infty} f_n = f$, if and only if, at each $x\in X$, both of the following conditions are satisfied:

(i) There exists a sequence $\langle x_n\rangle$ convergent to x with $\limsup_{n\to\infty} f_n(x_n)\le f(x)$;

(ii) For each sequence $\langle x_n\rangle$ weakly convergent to x, $\liminf_{n\to\infty} f_n(x_n)\ge f(x)$.

The main feature of Mosco convergence in *reflexive* spaces is the sequential bicontinuity of the Young–Fenchel transform $f\mapsto f^*$ from $\Gamma(X)$ onto $\Gamma(X^*)$: $\text{M-}\lim_{n\to\infty} f_n = f$ if and only if $\text{M-}\lim_{n\to\infty} f_n^* = f^*$ for any sequence $\langle f_n\rangle$ in $\Gamma(X)$ and any $f\in\Gamma(X)$, a result due to U. Mosco [67].

Theorem 4.3 *Let X be a reflexive Banach space and suppose that the functions f, g and $f_0, f_1, f_2, \ldots$ all belong to $\Gamma(X)$.*

(a) *If g^* is real-valued, then*

$$\text{M-}\lim_{n\to\infty} f_n = f \Rightarrow \text{M-}\lim_{n\to\infty} f_n \Box g = f \Box g.$$

(b) *If g^* is real-valued and continuously differentiable, or real-valued and weakly sequentially continuous, then*

$$\text{M-}\lim_{n\to\infty} f_n = f \Leftrightarrow \text{M-}\lim_{n\to\infty} f_n \Box g = f \Box g.$$

PROOF. Assertion (a) and the part of (b) concerning g^* continuously differentiable are proved in [5]. Suppose that g^* is real-valued and weakly sequentially continuous. It suffices to show that the Mosco convergence of $\langle (f_n \Box g)^* \rangle$ to $(f \Box g)^*$, that is, of $\langle f_n^* + g^* \rangle$ to $f^* + g^*$, implies that of $\langle f_n^* \rangle$ to f^*. This implication is straightforward to verify. $\Box$

4.3 The Mosco topology and the slice topology

G. Beer has introduced, see [11], a topology τ_{M} on $\mathbf{C}(X)$, named the "Mosco topology," such that if X is a Banach space, then the convergent sequences in $(\mathbf{C}(X), \tau_{\mathrm{M}})$ are exactly the Mosco convergent sequences [11, Theorem 3.1] or [15, Theorem 5.4.6]. The topological space $(\mathbf{C}(X), \tau_{\mathrm{M}})$ is Hausdorff if and only if X is reflexive. When X is reflexive, $(\mathbf{C}(X), \tau_{\mathrm{M}})$ is first countable if and only if the underlying space X is separable [11, Lemma 4.1]. Consequently, in the reflexive case, sequences determine the Mosco topology if and only if X is separable. Actually, $(\mathbf{C}(X), \tau_{\mathrm{M}})$ is completely metrizable and separable when X is reflexive and separable [11, Theorem 4.3]. The *Mosco topology* τ_{M} on $\mathbf{C}(X)$ is defined to be the hit-and-miss topology determined by the (nonempty) weakly compact subsets of X. The topologies τ^- and τ_{M}^+ for $\mathbf{C}(X)$ are generated, respectively, by

$$\{V^- \mid V \text{ nonempty open subset of } X\}$$

and

$$\{(K^{\mathrm{c}})^+ \mid K \text{ nonempty weakly compact subset of } X\}.$$

In other words, a subbasic set G^- in τ^-, where G is open, consists of those members of $\mathbf{C}(X)$ that meet G; whereas a subbasic set $(K^{\mathrm{c}})^+$ in τ_{M}^+, where K is nonempty and weakly compact, consists of those elements of $\mathbf{C}(X)$ that are disjoint from K. The Mosco topology τ_{M} on $\mathbf{C}(X)$ is the supremum of the two topologies above: $\tau_{\mathrm{M}} = \tau^- \vee \tau_{\mathrm{M}}^+$. By the weak lower semicontinuity of the norm, we have $(K^{\mathrm{c}})^+ = (K^{\mathrm{c}})^{++}$. Thus, the Mosco topology is a hit-and-miss as well as a proximal hit-and-miss topology. The sequential continuity for the Young–Fenchel transform has been extended to a true continuity result for the Mosco topology: the Young–Fenchel transform is a homeomorphism of $(\Gamma(X), \tau_{\mathrm{M}})$ onto

$(\Gamma(X^*), \tau_M)$ [15]. However, G. Beer and J. M. Borwein have demonstrated that Mosco convergence is a useful notion only when X is reflexive; for instance, whenever X is a nonreflexive space, the Young–Fenchel transform fails to be sequentially continuous for the Mosco topology [16, Theorems 3.2, 3.3].

The *slice topology* τ_s on $\mathbf{C}(X)$ is the weak topology generated by the family of gap functions $\{D(B,\cdot)|\ B \in \mathbb{CB}(X)\}$. Put differently, a net $\langle C_\lambda \rangle$ in $\mathbf{C}(X)$ τ_s-converges to $C \in \mathbf{C}(X)$ if and only if $\lim D(B, C_\lambda) = D(B, C)$ for each $B \in \mathbb{CB}(X)$. A proximal hit-and-miss representation of the slice topology is obtained by taking as a subbase for $(\mathbf{C}(X), \tau_s)$ all sets of the form V^-, where V is norm open, and $(B^c)^{++}$, where $B \in \mathbb{CB}(X)$. In general the slice topology is finer than the Mosco topology and the two topologies agree precisely when X is reflexive. The teminology "slice topology" arises from the fact that it suffices for B to run over all "slices" of balls to generate the slice topology [13, Theorem 3.5].

Many of the nice properties enjoyed by the Mosco topology in a reflexive setting remain valid for the slice topology on $\Gamma(X)$ with X an arbitrary normed space, see [13]. For instance, the Young–Fenchel transform is a homeomorphism of $(\Gamma(X), \tau_s)$ onto $(\Gamma^*(X^*), \tau_s^*)$, where τ_s^* denotes the "dual slice topology" [15, Theorem 8.2.3].

The following two lemmas will be useful.

Lemma 4.1 *Let X be a normed space. Suppose $\langle f_\lambda \rangle_{\lambda\in\Lambda}$ and $\langle g_\lambda \rangle_{\lambda\in\Lambda}$ are nets in $\Gamma(X)$ on a common directed set, τ^--convergent to $f \in \Gamma(X)$ and $g \in \Gamma(X)$, respectively, where* $\operatorname{dom} f^* \cap \operatorname{dom} g^* \neq \emptyset$. *Then $\langle \operatorname{cl}(f_\lambda \Box g_\lambda) \rangle$ τ^--converges to* $\operatorname{cl}(f \Box g)$.

PROOF. According to [10, Lemma 2.1] we may use a somewhat smaller subbase for τ^- than mentioned above. A subbase for τ^- consists of all sets of the form

$$\Gamma(X) \cap (G\times\,]-\infty, \mu[)^-, \text{ where } G \text{ is an open subset of } X \text{ and } \mu \in \mathbb{R}.$$

Suppose $\operatorname{cl}(f\Box g) \in (G\times\,]-\infty,\mu[)^-$, where G is an open subset of X and $\mu \in \mathbb{R}$; so that $(f\Box g)(x) < \mu$ for some $x \in G$. (First select $x' \in G$ with $\operatorname{cl}(f\Box g)(x') < \mu$. Then there exists $U \in \mathcal{N}(x')$, $U \subseteq G$, with $\inf_{\xi\in U}(f\Box g)(\xi) < \mu$.) By the definition of infimal convolution, there exist y and z in X satisfying $y+z = x$ and $f(y)+g(z) < \mu$. Select $\alpha, \beta \in \mathbb{R}$ such that $f(y) < \alpha$, $g(z) < \beta$ and $\alpha + \beta \le \mu$. Let N_y and N_z be open neighborhoods of y and z, respectively, with $N_y + N_z \subseteq G$. Then, eventually, $f_\lambda \in (N_y\times\,]-\infty,\alpha[)^-$ and $g_\lambda \in (N_z\times\,]-\infty,\beta[)^-$; hence, $f_\lambda(y_\lambda) < \alpha$ and $g_\lambda(z_\lambda) < \beta$ for some $y_\lambda \in N_y$ and $z_\lambda \in N_z$. It follows that, eventually,

$$\operatorname{cl}(f_\lambda\Box g_\lambda)(y_\lambda + z_\lambda) \le f_\lambda(y_\lambda) + g_\lambda(z_\lambda) < \alpha + \beta \le \mu,$$

which together with the fact that $y_\lambda + z_\lambda \in G$ show that $\operatorname{cl}(f_\lambda\Box g_\lambda) \in (G\times\,]-\infty,\mu[)^-$. $\Box$

Lemma 4.2 *Let X be a normed vector space, $\langle f_\lambda \rangle$ a net in $\Gamma(X)$ and $f \in \Gamma(X)$. Then $f = \tau_s$-$\lim f_\lambda$ if and only if both of the following conditions are met:*

(i) *For each open subset V of X and each real α, the condition $f \in (V\times\,]-\infty,\alpha[)^-$ implies that $f_\lambda \in (V\times\,]-\infty,\alpha[)^-$ eventually;*

(ii) *For each open subset W of X^* and each real β, the condition $f^* \in (W\times\,]-\infty,\beta[)^-$ implies that $f^*_\lambda \in (W\times\,]-\infty,\beta[)^-$ eventually.*

PROOF. See G. Beer [13, Proposition 2.4], [14, Theorem 4.1]. □

Theorem 4.4 *Let X be a normed space and $g \in \Gamma(X)$ with g^* real-valued.*

(a) *The map $f \mapsto \mathrm{cl}(f\Box g)$ is a continuous operation on $(\Gamma(X), \tau_{\mathrm{s}})$.*

(b) *Suppose that X is reflexive. Then $f \mapsto f\Box g$ is a continuous operation on $(\Gamma(X), \tau_{\mathrm{M}})$.*

PROOF. (a) Suppose $\langle f_\lambda\rangle_{\lambda\in\Lambda}$ is a net in $\Gamma(X)$, slice-convergent to $f \in \Gamma(X)$. Condition (i) of Lemma 4.2 is fulfilled (with f replaced by $f\Box g$ and f_λ by $f_\lambda\Box g$), according to Lemma 4.1. It remains to verify condition (ii). To this end, assume that $f^* + g^* = [\mathrm{cl}(f\Box g)]^*$ belongs to $(W\times\,]-\infty,\mu[)^-$, where W is an open subset of X^*, μ real. Select a $\xi^* \in W$ with $f^*(\xi^*) + g^*(\xi^*) < \mu$. Take real numbers α and β such that $f^*(\xi^*) < \alpha$, $g^*(\xi^*) < \beta$, and $\alpha + \beta \leq \mu$. Since g^* is continuous, there exists an open neighborhood N of ξ^*, $N \subseteq W$, such that $g^*(x^*) \leq \beta$ whenever $x^* \in N$. As $f^* \in (N\times\,]-\infty,\alpha[)^-$, according to Lemma 4.2, there exists $\lambda_0 \in \Lambda$ such that $f^*_\lambda \in (N\times\,]-\infty,\alpha[)^-$ (put differently, there exists $x^*_\lambda \in N$ with $f^*_\lambda(x^*_\lambda) < \alpha$) whenever $\lambda \geq \lambda_0$. Hence,

$$[\mathrm{cl}(f_\lambda\Box g)]^*(x^*_\lambda) = f^*_\lambda(x^*_\lambda) + g^*(x^*_\lambda) < \alpha + \beta \leq \mu$$

whenever $\lambda \geq \lambda_0$. Therefore, $[\mathrm{cl}(f_\lambda\Box g)]^*$ belongs to $(W\times\,]-\infty,\mu[)^-$ eventually.

(b) If X is reflexive, then the slice topology reduces to the Mosco topology and all functions involved are lower semicontinuous (Theorem 3.4), and the proof follows. □

In the case of sequences we have the following results due to J. Lahrache:

Theorem 4.5 *Let X be a normed space, let $\langle f_n\rangle$ and $\langle g_n\rangle$ be sequences in $\Gamma(X)$ τ_{s}-convergent to $f \in \Gamma(X)$ and $g \in \Gamma(X)$, respectively. Suppose there exist $x^*_0 \in \mathrm{dom}\, f^*$ and $\delta \in \mathbb{P}$ such that*

$$\sup\{g^*_n(x^*) \mid (n, x^*) \in \mathbb{N} \times B(x^*_0, \delta)\} < +\infty.$$

Then $f\Box g$ is proper, $f_n\Box g_n$ is proper for all n sufficiently large, and

$$\tau_{\mathrm{s}}\text{-}\lim_{n\to\infty} \mathrm{cl}(f_n\Box g_n) = \mathrm{cl}(f\Box g).$$

PROOF. See J. Lahrache [47, Theorem 3.6]. □

Theorem 4.6 *Suppose that $f, g \in \Gamma(X)$, and that $\langle f_n\rangle$ is a sequence in $\Gamma(X)$, where X is a Banach space. Suppose moreover that g^* is continuously differentiable. Then,*

$$\tau_{\mathrm{s}}\text{-}\lim_{n\to\infty} f_n = f \Leftrightarrow \tau_{\mathrm{s}}\text{-}\lim_{n\to\infty} f_n\Box g = f\Box g.$$

PROOF. See J. Lahrache [47, Theorem 5.2]. □

4.4 The affine topology

In the paper [12] G. Beer introduced a proximal hit-and-miss topology τ_a on $\Gamma(X)$, where X stands for an arbitrary normed space, called the *affine topology*. It is, by definition, generated by all sets of the form $\Gamma(X) \cap G^-$, where G is a norm open subset of $X \times \mathbb{R}$, and $\Gamma(X) \cap (\operatorname{epi} a)^{++}$, where $a: X \to \mathbb{R}$ is continuous and affine. The affine topology τ_a is the weakest topology τ on $\Gamma(X)$ that enjoys, firstly, that $f \mapsto \inf f$ is a continuous map $(\Gamma(X), \tau) \to \mathbb{R} \cup \{-\infty\}$, and, secondly, that for each continuous convex function $g: X \to \mathbb{R}$, $f \mapsto f + g$ is a continuous operation on $(\Gamma(X), \tau)$ [12, Theorem 3.3].

The affine topology τ_a is stronger than the slice topology [12, Lemma 2.3]. If $f \in \Gamma(X)$ and $\langle f_\lambda \rangle$ is a net in $\Gamma(X)$, then $f = \tau_a\text{-}\lim f_\lambda$ if and only if $f = \tau_s\text{-}\lim f_\lambda$ and the associated net of conjugate functions $\langle f_\lambda^* \rangle$ is pointwise convergent to f^*, see [12, Theorem 4.4].

Theorem 4.7 *Let X be a normed space, $g: X \to \mathbb{R}$ be convex continuous, and $x \in X$. Then $f \mapsto (f \Box g)(x)$ is a continuous map from $(\Gamma(X), \tau_a)$ into $\mathbb{R} \cup \{-\infty\}$.*

PROOF. The map under consideration may be written as the composition

$$f \mapsto f(\cdot) + g(x - \cdot) \mapsto (f \Box g)(x),$$

the first map continuous from $(\Gamma(X), \tau_a)$ into itself, the second continuous from $(\Gamma(X), \tau_a)$ into $\mathbb{R} \cup \{-\infty\}$. $\Box$

Following [12] we write $\Omega(x^*; \alpha)$ for $\Gamma(X) \cap (\operatorname{epi} a)^{++}$ when a is the continuous and affine function $x \mapsto \langle x^*, x \rangle - \alpha$, where $x^* \in X^*$ and $\alpha \in \mathbb{R}$. Put differently, $\Omega(x^*; \alpha)$ consists of all $f \in \Gamma(X)$ for which

$$\inf\{f(x) - (\langle x^*, x \rangle - \alpha) \mid x \in X\} > 0,$$

or, equivalently, $f^*(x^*) < \alpha$.

Theorem 4.8 *Let X be a normed space. Suppose that $\langle f_\lambda \rangle_{\lambda \in \Lambda}$ and $\langle g_\lambda \rangle_{\lambda \in \Lambda}$ are two nets in $\Gamma(X)$, τ_a-convergent to $f \in \Gamma(X)$ and $g \in \Gamma(X)$, respectively, where* $\operatorname{dom} f^* \cap \operatorname{dom} g^* \neq \emptyset$. *Then $f \Box g$ is proper, $f_\lambda \Box g_\lambda$ is eventually proper, and* $\operatorname{cl}(f \Box g) = \tau_a\text{-}\lim \operatorname{cl}(f_\lambda \Box g_\lambda)$.

PROOF. The τ^--part is clear according to Lemma 4.1. Suppose that $\operatorname{cl}(f \Box g) \in \Omega(x^*; \mu)$, where $x^* \in X^*$ and μ real. In other words, $[\operatorname{cl}(f \Box g)]^*(x^*) < \mu$. Hence, $f^*(x^*) + g^*(x^*) < \mu$. Let α, β be real numbers such that $f^*(x^*) < \alpha$, $g^*(x^*) < \beta$ and $\alpha + \beta \le \mu$. Since $f \in \Omega(x^*; \alpha)$ and $g \in \Omega(x^*; \beta)$ we have eventually $f_\lambda \in \Omega(x^*; \alpha)$ and $g_\lambda \in \Omega(x^*; \beta)$. Thus, eventually $x^* \in \operatorname{dom} f_\lambda^* \cap \operatorname{dom} g_\lambda^*$ and $(f_\lambda^* + g_\lambda^*)(x^*) < \alpha + \beta \le \mu$, which shows that $f_\lambda \Box g_\lambda$ is proper and $\operatorname{cl}(f_\lambda \Box g_\lambda) \in \Omega(x^*; \mu)$. $\Box$

4.5 The Attouch–Wets topology

The *Attouch-Wets topology* τ_{AW} on $\mathbf{C}(X)$ is the topology that $\mathbf{C}(X)$ inherits from the space of continuous real functionals on X, equipped with the metrizable topology of uniform convergence on bounded subsets of X, under the identification $C \leftrightarrow d(\cdot, C)$. This means that a sequence $\langle C_n \rangle_{n \in \mathbb{N}}$ in $\mathbf{C}(X)$ τ_{AW}-converges to $C \in \mathbf{C}(X)$ if and only if

$$\lim_{n \to \infty} \sup_{x \in B} |d(x, C) - d(x, C_n)| = 0$$

for each bounded subset B of X. The space $(\mathbf{C}(X), \tau_{\mathrm{AW}})$ is completely metrizable when X is complete [15, Theorem 3.1.3]. The Attouch–Wets topology admits another presentation. For each $\rho \in \mathbb{P}$ and $A, B \in \mathbf{C}(X)$, write

$$e_\rho(A, B) = e(\rho U \cap A, B), \text{ and } \operatorname{haus}_\rho(A, B) = \max\{e_\rho(A, B), e_\rho(B, A)\}.$$

Put differently,

$$\operatorname{haus}_\rho(A, B) = \inf\{\varepsilon \in \mathbb{P} |\, \rho U \cap A \subseteq B + \varepsilon U \text{ and } \rho U \cap B \subseteq A + \varepsilon U\}.$$

It is noteworthy to see that τ_{AW}-convergence is very close to Hausdorff metric convergence of bounded sets: C is the τ_{AW}-limit of $\langle C_n \rangle$ if and only if, for each $\rho \in \mathbb{P}$, $\operatorname{haus}_\rho(C, C_n) \to 0$ as n tends to ∞. The topology τ_{AW} reduces to the Hausdorff metric topology on $\mathbb{CB}(X)$. On $\mathbf{C}(X)$, the Attouch–Wets topology is stronger than the slice topology, and the two topologies coincide if $\dim X < \infty$.

The Attouch–Wets topology on $\Gamma(X)$, where $X \times \mathbb{R}$ is supplied with the box norm, is stable with respect to duality: the Young–Fenchel transform is a homeomorphim between $(\Gamma(X), \tau_{\mathrm{AW}})$ and $(\Gamma^*(X^*), \tau_{\mathrm{AW}})$ [15, Theorem 7.2.11]. The Attouch–Wets topology is a highly potent convergence concept applicable to optimization and approximation problems. For quantitative results, see H. Attouch and R. J.-B. Wets [8].

Theorem 4.9 *Let X be a normed space. Suppose $\langle f_n \rangle$ and $\langle g_n \rangle$ are two sequences in $\Gamma(X)$ τ_{AW}-convergent to $f \in \Gamma(X)$ and $g \in \Gamma(X)$, respectively, and that f^* is real-valued and continuous at some point of $\operatorname{dom} g^*$. Then $f \Box g$ is proper, $f_n \Box g_n$ is proper for all n sufficiently large, and $\tau_{\mathrm{AW}}\text{-}\lim_{n\to\infty} \operatorname{cl}(f_n \Box g_n) = \operatorname{cl}(f \Box g)$.*

PROOF. See G. Beer & R. Lucchetti [17], G. Beer [14, Proposition 4.2]. □

Theorem 4.10 *Suppose that $f, g \in \Gamma(X)$ and that $\langle f_n \rangle$ is a sequence in $\Gamma(X)$, where X is a normed space. Suppose moreover that g^* is Lipschitz continuous on norm bounded sets. Then,*

$$\tau_{\mathrm{AW}}\text{-}\lim_{n \to \infty} f_n = f \Leftrightarrow \tau_{\mathrm{AW}}\text{-}\lim_{n \to \infty} f_n \Box g = f \Box g.$$

PROOF. See H. Attouch, D. Azé and G. Beer [5, Theorem 3.4].

Theorem 4.11 *Let X be a normed space, and let $\{g_t | \, t \in \mathbb{P}\}$ be elements of $\Gamma(X)$ such that*

(i) *for each $t \in \mathbb{P}$, g_t is nonnegative;*

(ii) *for each $t \in \mathbb{P}$, $g_t(0) = 0$ and g_t is continuous at the origin;*

(iii) *$\langle g_t \rangle$ is Attouch–Wets convergent to $\delta_{\{0\}}$ as $t \to 0$.*

Suppose that f belongs to $\Gamma(X)$. Then, $\langle \mathrm{cl}(f \Box g_t) \rangle$ converges pointwise and in the Attouch–Wets topology to f as $t \to 0$.

PROOF. See G. Beer [15, Theorem 7.3.8]. $\Box$

Chapter 5

Regularization and approximation

Assume that $(X, \|\cdot\|)$ is a normed space. The most celebrated approximation process based on infimal convolution is *Moreau–Yosida approximation*. This method, initiated by J.-J. Moreau [61], is nowadays well understood and classical, at least in the case of convex functions on a Hilbert space; it is widely displayed, even in textbooks [2, 9, 22], and has been applied to a vareity of problems in convex analysis, optimization and variational analysis. With any extended real-valued function f on X we associate its *Moreau–Yosida approximates* $\langle f_t \rangle_{t \in \mathbb{P}}$, defined by $f_t = f \Box \frac{1}{2t}\|\cdot\|^2$ for each $t \in \mathbb{P}$.

5.1 Rudimentary properties of the Moreau–Yosida approximation

Some of the most elementary properties of the Moreau–Yosida approximates in the setting of a general normed space are collected in the following two statements.

Theorem 5.1 *Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a proper function defined on a normed space X. Then*

(a) *The function $t \mapsto f_t(x)$ is nonincreasing with $\lim_{t \to +\infty} f_t(x) = \inf f$ for each $x \in X$;*

(b) $f_t = (\operatorname{cl} f)_t \leq \operatorname{cl} f$ *for each* $t \in \mathbb{P}$;

(c) $\inf f_t = \inf f$ *and* $\operatorname{arg\,min} f_t = \operatorname{arg\,min} \operatorname{cl} f \supseteq \operatorname{arg\,min} f$ *for each* $t \in \mathbb{P}$;

(d) $(f_s)_t = f_{s+t}$ *for all* $s, t \in \mathbb{P}$.

PROOF. Part (a) is clear. (b) The equality $f_t = (\operatorname{cl} f)_t$ is a special case of Theorem 2.6, and the inequality $(\operatorname{cl} f)_t \leq \operatorname{cl} f$ is immediate. The identites $\inf f_t = \inf f$ and $\operatorname{arg\,min} f_t = \operatorname{arg\,min} \operatorname{cl} f$ in (c) follow from (b) and Theorem 2.3, and a proof of the inclusion $\operatorname{arg\,min} \operatorname{cl} f \supseteq \operatorname{arg\,min} f$ is straightforward. Let us verify the "semi-group property" (d). If $g \in \overline{\mathbb{R}}^X$ is convex, then $(s \bullet g) \Box (t \bullet g) = (s + t) \bullet g$ for all $s, t \in \mathbb{P}$, see Theorem

3.1. Hence, $(f\square(s\bullet g))\square(t\bullet g)$ coincides with $f\square((s+t)\bullet g)$ for any $f\in\bar{\mathbb{R}}^X$ and $s,t\in\mathbb{P}$, and the particular case $g=\frac{1}{2}\|\cdot\|^2$ gives (d). □

A "quadratic minorization" condition on f guarantees that the Moreau–Yosida approximates $\langle f_t\rangle_{t\in\mathbb{P}}$ be real-valued for all sufficiently small values of the index t. In that case we have the following regularization and approximation properties.

Theorem 5.2 *Let $(X,\|\cdot\|)$ be a normed space. Suppose that $f\colon X\to\mathbb{R}\cup\{+\infty\}$ is proper with $f+\frac{1}{2T}\|\cdot\|^2$ bounded from below for some $T\in\mathbb{P}$.*

(a) *Provided $t\in\,]0,T[$, f_t is real-valued and Lipschitz continuous on bounded sets;*

(b) *At each $x\in X$, $f_t(x)$ increases to $(\operatorname{cl} f)(x)$ as t decreases to zero.*

PROOF. See for instance H. Attouch [2, Theorem 2.64]. □

We close this section with a result for the Moreau–Yosida approximates of a convex function.

Theorem 5.3 *Let $(X,\|\cdot\|)$ be a normed space and $f\in\Gamma(X)$.*

(a) *The following dual formulation for $\langle f_t\rangle$ holds:*

$$f_t(x)=\max\left\{\langle x^*,x\rangle-f^*(x^*)-\frac{t}{2}\|x^*\|_*^2\,\middle|\,x^*\in\operatorname{dom} f^*\right\}\ \text{for each } (t,x)\in\mathbb{P}\times X.$$

Moreover, $(t,x)\mapsto f_t(x)$ is a convex and continuous function $\mathbb{P}\times X\to\mathbb{R}$.

(b) *$\langle f_t\rangle$ converges pointwise and in the Attouch–Wets topology to f.*

(c) *If the dual space X^* is rotund, then f_t is Gâteaux differentiable for each $t\in\mathbb{P}$.*

(d) *If the dual space X^* is locally uniformly rotund, then f_t is Fréchet differentiable for each $t\in\mathbb{P}$.*

(e) *If X is reflexive, then X admits an equivalent renorm, still denoted $\|\cdot\|$, such that simultaneously X and its dual X^* are locally uniformly rotund and Fréchet smooth (off zero). Then there exists exactly one minimizer, denoted $R_t^f x$, of $f(\cdot)+\frac{1}{2t}\|x-\cdot\|^2$ and f_t is Fréchet differentiable with $Df_t(x)=J(x-R_t^f x)/t$. Here, J is the "duality mapping" $J=D(\frac{1}{2}\|\cdot\|^2)$.*

PROOF. For (a) and (b), apply Theorem 3.3 and Theorem 4.11, respectively.

The additional assumption in (c) (respectively, (d)) implies that $\|\cdot\|_*^2$ is strictly convex (respectively, locally uniformly convex) and, hence, that f_t is Gâteaux differentiable (respectively, Fréchet differentiable), see Theorem 3.7 and [87, Theorem 4.1].

(e) See H. Attouch [2, Theorem 3.24]. □

5.2 Approximation of functions on a Hilbert space

In this section we review some of the most important properties of the Moreau–Yosida regularization and approximation process of a function defined on a Hilbert space. Let us first consider the Moreau–Yosida approximates $\langle f_t \rangle_{t\in\mathbb{P}}$ of a given $f \in \Gamma(X)$. Recall that

$$f_t(x) = \inf_{y\in X}\left(f(y) + \frac{1}{2t}\|x-y\|^2\right) \text{ for each } (t,x) \in \mathbb{P}\times X.$$

Due to the nice properties of the quadratic kernel function in a Hilbert space and the convex structure, the Moreau–Yosida approximates enjoy peculiar properties, well-known in the field of convex analysis. The above infimum is achieved at a unique point, denoted $R_t^f x$, characterized by the equation

$$\frac{1}{t}(x - R_t^f x) \in \partial f(R_t^f x) \text{ or, equivalently, } R_t^f x = (I + t\,\partial f)^{-1}x.$$

This means that $x \mapsto R_t^f x$ is the resolvent mapping of index t of the maximal monotone operator ∂f. The mappings $R_t^f : x \mapsto R_t^f x$ and $I - R_t^f$ share the same monotonicity and Lipschitz properties: for all $t \in \mathbb{P}$, x and y in X,

$$\begin{aligned}(R_t^f x - R_t^f y \,|\, x - y) &\geq \|R_t^f x - R_t^f y\|^2,\\ ((x - R_t^f x) - (y - R_t^f y) \,|\, x - y) &\geq \|(x - R_t^f x) - (y - R_t^f y)\|^2;\end{aligned}$$

consequently, by Schwarz inequality,

$$\begin{aligned}\|R_t^f x - R_t^f y\| &\leq \|x - y\|,\\ \|(x - R_t^f x) - (y - R_t^f y)\| &\leq \|x - y\|.\end{aligned}$$

Theorem 5.4 *Let X be a Hilbert space and $f \in \Gamma(X)$. Then, for each $t \in \mathbb{P}$, the Moreau–Yosida approximate f_t of f enjoys the following properties:*

(a) *To each $t \in \mathbb{P}$ and $x \in X$ there exists exactly one minimizer, written $R_t^f x$, of the function $f + \frac{1}{2t}\|x - \cdot\|^2$, and $R_t^f x = (I + t\,\partial f)^{-1}x$.*

(b) *For each $t \in \mathbb{P}$, f_t is real-valued, convex, Fréchet-differentiable with a derivative Lipschitz continuous of rate $1/t$. More precisely,*

$$Df_t = \frac{1}{t}(I - R_t^f)^{-1} = ((\partial f)^{-1} + tI)^{-1}.$$

(c) *As t decreases to zero, f_t increases to f, while the graph of Df_t converges in the Kuratowski-Painlevé set convergence sense to ∂f. Moreover, at each $x \in \mathcal{D}(\partial f)$, $Df_t(x)$ norm-converges to the element of minimal norm of $\partial f(x)$; in particular, $\|Df_t(x)\| \to 0$ if and only if $x \in \arg\min f$.*

PROOF. For a complete proof we refer to H. Attouch's monograph [2, Theorem 3.24, Proposition 3.56]. □

Next we present an interesting connection to a Hamilton–Jacobi equation.

Theorem 5.5 *If X is a Hilbert space and $f \in \Gamma(X)$, then $(t,x) \mapsto f_t(x)$ is a solution to the Hamilton–Jacobi equation*

$$\begin{aligned} \dot{u}(x,t) + \tfrac{1}{2}\|Du(x,t)\|^2 &= 0, \quad (x,t) \in X \times \mathbb{P}, \\ \lim_{t\to 0} u(x,t) &= f(x), \quad x \in X. \end{aligned}$$

PROOF. See H. Attouch [2, Lemma 3.27, Remark 3.32]. □

The nice properties enjoyed by the Moreau–Yosida approximation process of convex functions have been extended to functions "convex up to a square," see [4, 21]. We say that $f: X \to \mathbb{R}\cup\{+\infty\}$ is *convex up to a square* if there exists $c \in \mathbb{R}_+$ such that $f + \frac{1}{2}c\|\cdot\|^2$ is a convex function, that is,

$$f(tx + (1-t)y) \le tf(x) + (1-t)f(y) + t(1-t)\frac{c}{2}\|x-y\|^2$$

for all $x, y \in X$ and $t \in [0,1]$. Similarly, $g: X \to \mathbb{R}\cup\{+\infty\}$ is *concave up to a square* if there exists $c \in \mathbb{R}_+$ such that $g - \frac{1}{2}c\|\cdot\|^2$ is a concave function.

If f is an arbitrary extended real-valued function on X, then

$$\begin{aligned} f_t(x) &= \inf_{y\in X}\left[f(y) + \frac{1}{2t}(\|x\|^2 - 2(x|\,y) + \|y\|^2)\right] \\ &= \frac{1}{2t}\|x\|^2 - \sup_{y\in X}\left[(x/t|\,y) - \left(f(y) + \frac{1}{2t}\|y\|^2\right)\right] \\ &= \frac{1}{2t}\|x\|^2 - \left(f + \frac{1}{2t}\|\cdot\|^2\right)^*(x/t) \qquad (5.1) \end{aligned}$$

for all $x \in X$ and $t \in \mathbb{P}$. Consequently, the following statement holds (see also [4, Proposition 3.1]):

Theorem 5.6 *Let X be a Hilbert space and let f be a proper extended real-valued function on X with $f + \frac{1}{2T}\|\cdot\|^2$ bounded from below for some $T \in \mathbb{P}$. Then, for every $t \in]0,T[$, f_t is real-valued and concave up to a square: $f_t - \frac{1}{2t}\|\cdot\|^2$ is concave.*

Moreover, if $f + \frac{1}{2T}\|\cdot\|^2$ belongs to $\Gamma(X)$ for some $T \in \mathbb{P}$, then it is not hard to see that f_t is continuously differentiable whenever $t \in]0,T[$. Indeed, by duality (see Theorem 3.4 (a)),

$$\begin{aligned} (f + (2t)^{-1}\|\cdot\|^2)^* &= ((f + (2T)^{-1}\|\cdot\|^2) + ((2t)^{-1} - (2T)^{-1})\|\cdot\|^2)^* \\ &= (f + (2T)^{-1}\|\cdot\|^2)^* \square (t^{-1} - T^{-1})^{-1}\tfrac{1}{2}\|\cdot\|^2, \end{aligned}$$

which is a continuously differentiable function according to Theorem 5.4 (b). More precisely the following hold:

Theorem 5.7 *Let X be a Hilbert space and let f be a proper extended real-valued function on X with $f + \frac{1}{2T}\|\cdot\|^2$ in $\Gamma(X)$ for some $T \in \mathbb{P}$.*

(a) *Suppose $t \in]0,T[$. Then*

$$\begin{aligned} f_t(x) &= \frac{1}{2t}\|x\|^2 - \left(f + \frac{1}{2T}\|\cdot\|^2\right)_{t^{-1}-T^{-1}}(x/t) \\ &= -\frac{1}{2(T-t)}\|x\|^2 + \left(f + \frac{1}{2T}\|\cdot\|^2\right)_{(t^{-1}-T^{-1})^{-1}}(Tx/(T-t)) \end{aligned}$$

for all $x \in X$. The function $f_t - \|\cdot\|^2/(2t)$ is concave, and $f_t + \|\cdot\|^2/(2(T-t))$ is convex. Moreover, f_t is Fréchet differentiable and Df_t is $\max\{1/t, T/(T-t)\}$*-Lipschitz continuous.*

(b) *The graph of Df_t converges in the Kuratowski–Painlevé set convergence sense to ∂f.*

PROOF. See H. Attouch and D. Azé [4, Theorem 3.1].

If f is a proper extended real-valued function on X with $f + \frac{1}{2T}\|\cdot\|^2$ bounded from below for some $T \in \mathbb{P}$, then, according to Theorem 5.6, $-f_t + \frac{1}{2t}\|\cdot\|^2$ is convex continuous whenever $t \in]0,T[$. The assumptions of Theorem 5.7 are then satisfied with $-f_t$ in place of f. Thus, the *Lasry-Lions approximant* $f_{(s,t)} := -(-f_t)_s$ (see J.-M Lasry and P.-L. Lions [48]) is continuously differentiable provided $0 < s < t < T$. Moreover, it is readily verified that $f_t \leq f_{(s,t)} \leq \operatorname{cl} f$, which, in view of Theorem 5.2, forces pointwise convergence of $\langle f_{(s,t)} \rangle$ to $\operatorname{cl} f$ as $0 < s < t \to 0$. The "Lasry-Lions method" has been studied in [48, 4, 21, 83, 84].

Theorem 5.8 *Let X be a Hilbert space. Suppose $f\colon X \to \mathbb{R} \cup \{+\infty\}$ to be proper and quadratically minorized: $f + \frac{1}{2T}\|\cdot\|^2$ bounded from below for some $T \in \mathbb{P}$.*

(a) *Provided $0 < s < t < T$, $f_{(s,t)}$ is real-valued, convex, and Fréchet differentiable with $Df_{(s,t)}$* $\max\{1/s, 1/(t-s)\}$*-Lipschitz continuous.*

(b) *Provided $0 < s < t < T$, $f_t \leq f_{(s,t)} \leq \operatorname{cl} f$. Moreover, $f_{(s,t)}(x) \to (\operatorname{cl} f)(x)$ as $0 < s < t \to 0$, for each $x \in X$.*

(c) *If $f \in \Gamma(X)$, then $f_{(s,t)} = f_{t-s}$ when $0 < s < t < T$.*

PROOF. For a detailed proof, see H. Attouch and D. Azé [4, Theorems 4.1 & 4.2]. □

REMARK. Statement (c) means, roughly speaking, that the Lasry–Lions method and the classical Moreau–Yosida method coincide for elements of $\Gamma(X)$.

5.3 Local properties of the Moreau–Yosida approximates

In this section we make constant use of the following lemma, which is of independent interest:

Lemma 5.1 *Let $(X, \|\cdot\|)$ be a normed space. Suppose that $f\colon X \to \mathbb{R} \cup \{+\infty\}$ is proper with $f + \frac{1}{2T}\|\cdot\|^2 + \alpha \geq 0$ for some $T \in \mathbb{P}$ and $\alpha \in \mathbb{R}$.*

(a) *If $x \in \operatorname{dom} \operatorname{cl} f$, $t \in]0, \frac{1}{2}T[$ and $\rho > \bar{\rho}(x, t, T)$, then*

$$f_t(x) = \inf\left\{ f(y) + \frac{1}{2t}\|x - y\|^2 \,\middle|\, y \in B(x, \rho) \right\},$$

where

$$\bar{\rho}(x, t, T) = \left[t\frac{T((\operatorname{cl} f)(x) + \alpha) + \|x\|^2}{\frac{1}{2}T - t} \right]^{1/2}.$$

(b) *If f is bounded on a bounded set B, $x \in B$, $t \in]0, \frac{1}{2}T[$, and*

$$\rho > \hat{\rho}(B, t, T) := \left[t\frac{T(\sup_{x\in B}(\operatorname{cl} f)(x) + \alpha) + \sup_{x\in B}\|x\|^2}{\frac{1}{2}T - t} \right]^{1/2},$$

then

$$f_t(x) = \inf\left\{ f(y) + \frac{1}{2t}\|x - y\|^2 \,\middle|\, y \in B(x, \rho) \right\}.$$

PROOF. For a proof of part (a), see H. Attouch and D. Azé [4, Proposition 1.2]. Part (b) is, obviously, implied by (a). □

REMARK. Note that $\hat{\rho}(B, t, T) = \mathcal{O}(\sqrt{t})$ as $t \to 0$ if f is bounded on B.

Theorem 5.9 *Let X be a Hilbert space. Assume that $f\colon X \to \mathbb{R} \cup \{+\infty\}$ is a proper function and that there exist $T, \varepsilon \in \mathbb{P}$ and a bounded convex closed subset C of X such that*

(i) *$f + \frac{1}{2T}\|\cdot\|^2$ is bounded from below;*

(ii) *$f + \frac{1}{2T}\|\cdot\|^2$ is convex, lower semicontinuous and majorized on $C_\varepsilon := C + \varepsilon U$.*

Then for all sufficiently small values of t, $f_t|_C = [f + \delta_{C_\varepsilon}]_t|_C$ and $[f + \delta_{C_\varepsilon}]_t$ has a Lipschitz continuous derivative.

PROOF. According to Lemma 5.1 (b) there exists $T' \in]0, \frac{1}{2}T[$ such that

$$f_t(x) = \inf \left\{ f(y) + \frac{1}{2t}\|x-y\|^2 \middle| \ y \in x + \varepsilon U \right\}.$$

whenever $(t, x) \in]0, T'[\times C$. (There exists $T' \in]0, \frac{1}{2}T[$ such that $\hat{\rho}(C, t, T) \leq \varepsilon$ for all $t \in]0, T'[$.) As $C + \varepsilon U = C_\varepsilon$,

$$f_t(x) = \inf \left\{ f(y) + \frac{1}{2t}\|x-y\|^2 \middle| \ y \in C_\varepsilon \right\},$$

for all $(t, x) \in]0, T'[\times C$. Hence,

$$f_t(x) = [f + \delta_{C_\varepsilon}]_t(x)$$

whenever $(t, x) \in]0, T'[\times C$. The remaining assertions follows from Theorem 5.7 (b) and the fact that $f + \frac{1}{2T}\|\cdot\|^2 + \delta_{C_\varepsilon} \in \Gamma(X)$. □

Theorem 5.10 *Let $(X, \|\cdot\|)$ be a normed space. Let $f\colon X \to \mathbb{R} \cup \{+\infty\}$ be proper with $f + \frac{1}{2T}\|\cdot\|^2$ bounded from below for some $T \in \mathbb{P}$.*

(a) *Suppose that f is uniformly continuous on $B_\varepsilon := B + \varepsilon U$, where B is a bounded subset of X and $\varepsilon \in \mathbb{P}$. Moreover, suppose that f has modulus of uniform continuity m on B_ε, that is, $|f(x) - f(y)| \leq m(\|x-y\|)$ for all $x, y \in B_\varepsilon$, where $m\colon \mathbb{R}_+ \to \mathbb{R}_+$ is continuous, nondecreasing and subadditive with $m(0) = 0$. Then for all sufficiently small t*

$$\sup_{x\in B}(f(x) - f_t(x)) \leq \delta(t),$$

where

$$\delta(t) := \sup_{r\in\mathbb{R}_+}(m(r) - r^2/(2t))$$

goes to zero with t. In particular, if f is Lipschitz continuous on B_ε with constant k, then

$$\sup_{x\in B}(f(x) - f_t(x)) \leq \frac{k^2 t}{2}.$$

for all sufficiently small t.

(b) *If f is uniformly continuous on bounded sets, then $\langle f_t \rangle$ converges in the topology of uniform convergence on bounded sets and in the Attouch–Wets topology to f.*

REMARK. In the previous chapter we presented the Attouch–Wets topology in a convex setting. In the above theorem, the stated Attouch–Wets convergence of $\langle f_t \rangle$ to f means that $\langle d(\cdot, \operatorname{epi} f_t) \rangle$ converges uniformly to $d(\cdot, \operatorname{epi} f)$ on bounded subsets of $X \times \mathbb{R}$ ($X \times \mathbb{R}$ endowed with its box norm).

PROOF. (a) According to Lemma 5.1 there exists $T' \in]0, \frac{1}{2}T[$ such that

$$f_t(x) = \inf \left\{ f(y) + \frac{1}{2t}\|x-y\|^2 \middle| \; y \in B(x,\varepsilon) \right\}.$$

whenever $(t,x) \in]0,T'[\times B$. As $B + \varepsilon U = B_\varepsilon$,

$$f_t(x) = \inf \left\{ f(y) + \frac{1}{2t}\|x-y\|^2 \middle| \; y \in B_\varepsilon \right\}.$$

for all $(t,x) \in]0,T'[\times B$. It follows that for $t \in]0,T'[$

$$\begin{aligned}
\sup\{f(x) - f_t(x) \,|\, x \in B\} &= \sup \left\{ f(x) - f(y) - \frac{1}{2t}\|x-y\|^2 \middle| \, (x,y) \in B \times B_\varepsilon \right\} \\
&\leq \sup \left\{ m(\|x-y\|) - \frac{1}{2t}\|x-y\|^2 \middle| \, (x,y) \in B \times B_\varepsilon \right\} \\
&\leq \sup\{m(r) - r^2/(2t) \,|\, r \in \mathbb{R}_+\} = \delta(t).
\end{aligned}$$

It is an easy exercise to verify that $\delta(t)$ goes to zero with t.

(b) According to (a), $\langle f_t \rangle$ converges to f uniformly on bounded sets. This forces convergence in the Attouch–Wets topology, see [15, Lemma 7.1.2]. □

Corollary 5.1 *Let $(X, \|\cdot\|)$ be a normed space. Let $f: X \to \mathbb{R} \cup \{+\infty\}$ be proper with $f + \frac{1}{2T}\|\cdot\|^2$ bounded from below for some $T \in \mathbb{P}$, and let $x_0 \in X$. Suppose that f is convex continuous on some neighborhood of x_0. Then there exist $\delta \in \mathbb{P}$ and $c \in \mathbb{R}_+$ such that*

$$\sup_{x \in B(x_0,\delta)} (f(x) - f_t(x)) \leq ct \tag{5.2}$$

for all sufficiently small t.

PROOF. Under these assumptions f is locally Lipschitz at x_0, see [29]. Apply Theorem 5.10. □

REMARK. The previous theorem has, of course, a counterpart for the Lasry–Lions approximants; simply note that uniform convergence of $\langle f_t \rangle$ to f implies that of $\langle f_{(s,t)} \rangle$ to f since $f_t \leq f_{(s,t)} \leq f$ provided $0 < s < t < T$.

Since we are dealing with locally Lipschitz continuous functions, we shall recall the useful notion of Clarke's subdifferential [24]. If $f: X \to \mathbb{R} \cup \{+\infty\}$ is finite and satisfies a Lipschitz condition near a given point $x \in X$, then we put

$$f^\circ(x; v) = \limsup_{(\xi,\lambda) \to (x,0+)} \frac{f(\xi + \lambda v) - f(\xi)}{\lambda},$$

the *generalized directional derivative* of f at x in the direction v, and

$$\partial f(x) = \{x^* \in X^* | \langle x^*, v\rangle \leq f^\circ(x; v) \text{ for all } v \in X\},$$

the *generalized gradient* at x, which is a nonempty, convex, weak* compact subset of X^*. When f is convex, and Lipschitz near x, then $\partial f(x)$ coincides with the subdifferential at x in the sense of convex analysis, see F. H. Clarke [24, Proposition 2.2.7]. This justifies the notation "$\partial f(x)$."

Next we present a result on the weak* closed upper limit of the generalized gradients of the Moreau–Yosida approximates.

Theorem 5.11 *Let $(X, \|\cdot\|)$ be a normed space. Let $f: X \to \mathbb{R} \cup \{+\infty\}$ be proper with $f + \frac{1}{2T}\|\cdot\|^2$ bounded from below for some $T \in \mathbb{P}$. Suppose f is locally Lipschitz continuous at $x \in X$. Then,*

$$w^*\text{-}\limsup_{(\xi,t)\to(x,0)} \partial f_t(\xi) \subseteq \partial f(x).$$

PROOF. Suppose $x^* \notin \partial f(x)$, which means that there exist $v \in X$ and $\varepsilon \in \mathbb{P}$ with

$$\langle x^*, v\rangle \geq f^\circ(x; v) + 2\varepsilon.$$

Fix a δ in $\mathbb{P}$ in such a way that f is Lipschitz continuous on $B(x, 3\delta)$ and

$$\frac{f(y + \lambda v) - f(y)}{\lambda} \leq f^\circ(x; v) + \varepsilon \text{ whenever } (y, \lambda) \in B(x, 3\delta)\times\,]0, \delta[.$$

The two inequalities above jointly imply that

$$\frac{f(y + \lambda v) - f(y)}{\lambda} \leq \langle x^*, v\rangle - \varepsilon \text{ whenever } (y, \lambda) \in B(x, 3\delta)\times\,]0, \delta[.$$

Moreover, take $T' \in\,]0, \frac{1}{2}T[$ (see Lemma 5.1 (b)) such that

$$f_t(\xi) = \inf_{y\in B(x,3\delta)} \left(f(y) + \frac{1}{2t}\|\xi - y\|^2\right) \text{ for any } (\xi, t) \in B(x, 2\delta)\times\,]0, T'[.$$

If $(\xi, y, t, \lambda) \in B(x, 2\delta) \times B(x, 3\delta)\times\,]0, T'[\, \times\,]0, \delta[$, then

$$\begin{aligned} & f_t(\xi + \lambda v) - \left(f(y) + \frac{1}{2t}\|\xi - y\|^2\right) \\ \leq\; & \left(f(y + \lambda v) + \frac{1}{2t}\|(\xi + \lambda v) - (y + \lambda v)\|^2\right) - \left(f(y) + \frac{1}{2t}\|\xi - y\|^2\right) \\ =\; & f(y + \lambda v) - f(y) \leq \lambda(\langle x^*, v\rangle - \varepsilon); \end{aligned}$$

so that dividing by λ and taking supremum over y in $B(x, 3\delta)$, we obtain that

$$\frac{f_t(\xi + \lambda v) - f_t(\xi)}{\lambda} \leq \langle x^*, v\rangle - \varepsilon \text{ whenever } (\xi, t, \lambda) \in B(x, 2\delta)\times\,]0, T'[\, \times\,]0, \delta[.$$

Consequently, $(f_t)^\circ(\xi; v) \leq \langle x^*, v\rangle - \varepsilon$ for all $(\xi, t) \in B(x, \delta)\times\,]0, T'[$. If $y^* \in X^*$ satisfies $\langle x^* - y^*, v\rangle < \varepsilon$, then we have, for each $(\xi, t) \in B(x, \delta)\times\,]0, T'[$, $\langle y^*, v\rangle > (f_t)^\circ(\xi; v)$ and, hence, $y^* \notin \partial f_t(\xi)$. We conclude that $\{y^* \in X^* |\, \langle x^* - y^*, v\rangle < \varepsilon\}$ is a weak* neighborhood of x^* which does not meet $\bigcup\{\partial f_t(\xi) |\, (\xi, t) \in B(x, \delta)\times\,]0, T'[\}$. Thus, x^* does not belong to $w^*\text{-}\limsup_{(\xi,t)\to(x,0)} \partial f_t(\xi)$. Since x^* was arbitrarily chosen outside $\partial f(x)$ we must have

$$w^*\text{-}\limsup_{(\xi,t)\to(x,0)} \partial f_t(\xi) \subseteq \partial f(x),$$

as asserted. □

5.4 Generalized Moreau–Yosida approximation

We close this text with a result concerning more general kernel functions than quadratic ones. In particular, we complement a result due to H. Attouch. We rely our arguments on the following lemma, the proof of which is based on the variational properties of epi-convergence:

Lemma 5.2 *Let $\varphi: Y \to \mathbb{R} \cup \{+\infty\}$ be a proper function which is bounded from below, and let $\mathcal{H}: Y \to \mathbb{R}_+$ (called the "viscosity function"). Suppose that* $\arg\min(\varphi + t\mathcal{H})$ *is nonempty for each $t \in \mathbb{P}$, and that there exists a topology τ on Y such that both of the following conditions are met:*

(i) *Whenever, for each $t \in \mathbb{P}$, y_t is selected in* $\arg\min(\varphi + t\mathcal{H})$, *then $\langle y_t\rangle$ is τ-relatively compact as $t \to 0$;*

(ii) *φ and $\mathcal{H}$ are τ-lower semicontinuous.*

Then,

(a) *Every τ-cluster point $\bar{y}$ of $\langle y_t\rangle$ minimizes φ on Y and satisfies the "viscosity selection principle"*

$$\mathcal{H}(\bar{y}) \leq \mathcal{H}(y) \text{ for every } y \in \arg\min \varphi.$$

(b) $\lim_{t\to 0} \varphi(y_t) = \inf \varphi$; *more precisely,*

$$\lim_{t\to 0}(\varphi(y_t) - \inf \varphi)/t = 0.$$

(c) *$\langle \mathcal{H}(y_t)\rangle$ converges according to*

$$\lim_{t\to 0} \mathcal{H}(y_t) = \min\{\mathcal{H}(y) |\, y \in \arg\min \varphi\}.$$

PROOF. See H. Attouch [3, Theorem 2.1]. □

Our final theorem complements statements due to H. Attouch [3, Proposition 5.3], and G. Beer and R. Lucchetti [17, Theorem 3.8].

Theorem 5.12 *Let X be a normed space. Suppose that $\mathcal{L}: X \to \mathbb{R}_+$ is convex continuous with $\mathcal{L}(0) = 0$. Moreover, suppose that $\mathcal{H} := \mathcal{L}^*$ is real-valued, coercive, bounded on norm bounded sets, and locally uniformly convex in the sense that*

$$\liminf_{t\to 0}(\mathcal{H}(\tfrac{1}{2}x^* + \tfrac{1}{2}x_t^*) - \tfrac{1}{2}\mathcal{H}(x^*) - \tfrac{1}{2}\mathcal{H}(x_t^*)) \geq 0 \text{ implies } \lim_{t\to 0}\|x^* - x_t^*\|_* = 0$$

whenever $x^* \in X^*$ and $\langle x_t^*\rangle_{t\in\mathbb{P}}$ is a net in X^*. Let $f \in \Gamma(X)$ and put $f_t = f\Box(t \bullet \mathcal{L})$ for each $t \in \mathbb{P}$. Then the following assertions hold:

(a) *$(f_s)_t = f_{s+t}$ for all $s, t \in \mathbb{P}$. (Of course, $(f_s)_t = f\Box(s \bullet \mathcal{L})\Box(t \bullet \mathcal{L})$.)*

(b) *$(x,t) \mapsto f_t(x): X \times \mathbb{P} \to \mathbb{R}$ is a convex continuous function, and*

$$f_t(x) = -\min_{x^*\in X^*}(f^*(x^*) - \langle x^*, x\rangle + t\mathcal{H}(x^*)) \tag{5.3}$$

for all $(x,t) \in X \times \mathbb{P}$.

(c) *$\langle f_t\rangle$ converges pointwise and in the Attouch-Wets topology to f as t goes to zero.*

(d) *The function f_t is Fréchet differentiable for every $t \in \mathbb{P}$, and $Df_t(x)$ is equal to the unique minimizer of (5.3). Moreover, $\sup_{t\in\mathbb{P}}\|Df_t(x)\|_*$ is a finite number if and only if $\partial f(x) \neq \emptyset$. In that case*

$$\mathcal{H}(Df_t(x)) \leq \min\{\mathcal{H}(x^*) \mid x^* \in \partial f(x)\}$$

for all $t \in \mathbb{P}$, and

$$\lim_{t\to 0}\mathcal{H}(Df_t(x)) = \min\{\mathcal{H}(x^*) \mid x^* \in \partial f(x)\}.$$

(e) *Suppose $\partial f(x) \neq \emptyset$. Then, as $t \to 0$, $\langle Df_t(x)\rangle$ norm converges to the unique minimizer, denoted $[\partial f(x)]_{\mathcal{H}}$, of*

$$\min\{\mathcal{H}(x^*) \mid x^* \in \partial f(x)\}.$$

(f) *Provided $\partial f(x) \neq \emptyset$,*

$$\lim_{t\to 0}\frac{f_t(x) - f(x)}{t} = -\min\{\mathcal{H}(x^*) \mid x^* \in \partial f(x)\}.$$

(g) *For any $(x,t) \in X \times \mathbb{P}$,*

$$\lim_{h\to 0+}\frac{f_{t+h}(x) - f_t(x)}{h} = -\mathcal{H}(Df_t(x)).$$

PROOF. For a proof of assertion (a), see the proof of Theorem 5.1 (d).

(b) According to Theorem 3.3, since $(t \bullet \mathcal{L})^* = t\mathcal{H}$, we have the dual formulation

$$f_t(x) = -\min_{x^* \in X^*}(f^*(x^*) - \langle x^*, x\rangle + t\mathcal{H}(x^*))$$

for all $(x,t) \in X \times \mathbb{P}$. In particular, $(x,t) \mapsto f_t(x)$ is continuous.

(c) We claim that $\tau_{\mathrm{AW}}\text{-}\lim_{t\to 0} t \bullet \mathcal{L} = \delta_{\{0\}}$. We have that $\langle (t \bullet \mathcal{L})^* = t\mathcal{H}\rangle$ converges uniformly to the zero function on bounded sets as $t \to 0$, since $\mathcal{H}$ is bounded on bounded sets. Hence $\langle (t \bullet \mathcal{L})^*\rangle$ Attouch–Wets converges to 0 (see [15, Lemma 7.1.2]) which means, in view of the fact that the Legendre–Fenchel transform is a homeomorphism between $(\Gamma(X), \tau_{\mathrm{AW}})$ and $(\Gamma^*(X^*), \tau_{\mathrm{AW}})$, that $\langle t \bullet \mathcal{L}\rangle$ Attouch–Wets converges to $\delta_{\{0\}}$, as claimed. Now, the stated convergence of $\langle f_t\rangle$ to f follows from Theorem 4.11.

(d) Since $\mathcal{H}$ is locally uniformly convex, f_t is Fréchet differentiable, see Theorem 3.7.

In the remaining part of the proof of the theorem, $x \in X$ is fixed, and $\varphi: X^* \to \mathbb{R} \cup \{+\infty\}$ is the function whose value at each $x^* \in X^*$ is

$$\varphi(x^*) = f^*(x^*) - \langle x^*, x\rangle,$$

so that $\arg\min \varphi = \partial f(x)$ and $\arg\min(\varphi + t\mathcal{H}) = \{Df_t(x)\}$. We set $x_t^* = Df_t(x)$.

Suppose $\arg\min \varphi \neq \emptyset$. Obviously, for every $x^* \in \arg\min \varphi$,

$$\varphi(x_t^*) + t\mathcal{H}(x_t^*) \leq \varphi(x^*) + t\mathcal{H}(x^*) = \inf \varphi + t\mathcal{H}(x^*),$$

from which it follows that

$$\mathcal{H}(x_t^*) \leq \min\{\mathcal{H}(x^*) \mid x^* \in \arg\min \varphi\}.$$

This inequality together with the coerciveness of $\mathcal{H}$ force $\{x_t^* | t \in \mathbb{P}\}$ to be norm bounded. Conversely, suppose that $\{x_t^* | t \in \mathbb{P}\}$ is norm bounded and, hence, by the Banach–Alaoglu theorem, weak* relatively compact. Part (a) of Lemma 5.2 may be applied with X^* supplied with its weak* topology as the topological space (Y, τ), the conjugate functions appearing in (5.3) being weak* lower semicontinuous. Thus every weak* cluster point of $\langle x_t^*\rangle$—there exists at least one—minimizes φ over X^*, see Lemma 5.2 (a). Moreover, by Lemma 5.2 (c),

$$\lim_{t\to 0} \mathcal{H}(x_t^*) = \min\{\mathcal{H}(x^*) \mid x^* \in \arg\min \varphi\}.$$

(e) Write $x^* = [\partial f(x)]_{\mathcal{H}}$. According to the proof of (d), $\langle x_t^*\rangle$ is weak* relatively compact, and each weak* cluster point of $\langle x_t^*\rangle$ minimizes $\mathcal{H}$ over $\arg\min \varphi$, that is, is equal to x^*. Hence, the whole net $\langle x_t^*\rangle$ weak* converges to x^*. We claim that $\langle x_t^*\rangle$ converges to x^* in norm. The weak* lower semicontinuity of $\mathcal{H}$, and $\lim_{t\to 0} \mathcal{H}(x_t^*) = \mathcal{H}(x^*)$ jointly imply the string of inequalities

$$\begin{aligned}
&\liminf_{t\to 0}(\mathcal{H}(\tfrac{1}{2}x^* + \tfrac{1}{2}x_t^*) - \tfrac{1}{2}\mathcal{H}(x^*) - \tfrac{1}{2}\mathcal{H}(x_t^*)) \\
&\geq \liminf_{t\to 0} \mathcal{H}(\tfrac{1}{2}x^* + \tfrac{1}{2}x_t^*) - \tfrac{1}{2}\mathcal{H}(x^*) - \tfrac{1}{2}\limsup_{t\to 0} \mathcal{H}(x_t^*) \\
&\geq \mathcal{H}(x^*) - \tfrac{1}{2}\mathcal{H}(x^*) - \tfrac{1}{2}\mathcal{H}(x^*) = 0,
\end{aligned}$$

and it follows, recall that $\mathcal{H}$ is assumed to be locally uniformly convex, that $\|x^* - x_t^*\|_*$ goes to zero with t, as claimed.

(f) If $\arg\min \varphi \neq \emptyset$, then, according to Lemma 5.2 (b),

$$\begin{aligned} 0 &= \lim_{t\to 0} \frac{\varphi(x_t^*) - \inf \varphi}{t} = \lim_{t\to 0} \frac{-f_t(x) - t\mathcal{H}(x_t^*) + f(x)}{t} \\ &= -\lim_{t\to 0} \frac{f_t(x) - f(x)}{t} - \min\{\mathcal{H}(x^*) |\, x^* \in \partial f(x)\}. \end{aligned}$$

(g) Part (a), and part (f) with f replaced by f_t yield

$$\lim_{h\to 0+} \frac{f_{t+h}(x) - f_t(x)}{h} = \lim_{h\to 0+} \frac{(f_t)_h(x) - f_t(x)}{h} = -\mathcal{H}(Df_t(x)).$$

The proof is complete. □

REMARK. In particular, we may take $\mathcal{L} = \frac{1}{2}\|\cdot\|^2$ if $(X, \|\cdot\|)$ has a locally uniformly rotund dual. Indeed, $\mathcal{H} = \frac{1}{2}\|\cdot\|_*^2$ is then a locally uniformly convex function. We then obtain a list of results for the Moreau–Yosida approximates.

Bibliography

[1] W. N. ANDERSON, AND R. J. DUFFIN, Series and parallel addition of matrices, *J. Math. Anal. Appl.* **26** (1969), 576–594.

[2] H. ATTOUCH, "Variational Convergence for Functions and Operators," Applicable Math. Series, Pitman, London 1984.

[3] H. ATTOUCH, Viscosity solutions of minimization problems. Epi-convergence and scaling, *Sém. Anal. Convexe Montpellier* **22** (1992), 8.1–8.46.

[4] H. ATTOUCH AND D. AZÉ, Approximation and regularization of arbitrary functions in Hilbert spaces by the Lasry–Lions method, *Sém. Anal. Convexe Montpellier* **20** (1990), 12.1–12.23. To appear in *Ann. Inst. H. Poincaré Anal. Non Linéaire.*

[5] H. ATTOUCH, D. AZÉ AND G. BEER, On some inverse stability problems for the epigraphical sum, *Nonl. Anal. T. M. A.* **16** (1991), 241–254.

[6] H. ATTOUCH AND H. BREZIS, Duality for the sum of convex functions in general Banach spaces, *in* "Aspects of Mathematics and its Applications" (J. A. Barroso, Ed.) 125–133, North-Holland, Amsterdam, 1986.

[7] H. ATTOUCH AND R. J.-B. WETS, Epigraphical analysis, *in* "Analyse non Linéaire" (H. Attouch, J.-P Aubin, F. H. Clarke, I. Ekeland, Eds.) 72–100, Gauthiers-Villars Paris et C.R.M. Montreal, 1989.

[8] H. ATTOUCH AND R. J.-B. WETS, Quantitative stability of variational systems: I. The epigraphical distance, *Trans. Amer. Math. Soc.* **328** (1991), 695–729.

[9] J.-P. Aubin, "L'analyse non linéaire et ses motivations économiques," Masson, Paris, 1984.

[10] G. BEER, On the Young–Fenchel transform for convex functions, *Proc. Amer. Math. Soc.* **104** (1988), 1115–1123.

[11] G. BEER, On Mosco convergence of convex sets, *Bull. Austral. Math. Soc.* **38** (1988), 239–253.

[12] G. BEER, Infima of convex functions, *Trans. Amer. Math. Soc.* **315** (1989), 849–859.

[13] G. BEER, The slice topology: a viable alternative to Mosco convergence in nonreflexive space, *Nonl. Anal. T. M. A.* **19** (1992), 271–290.

[14] G. BEER, Lipschitz regularization and the convergence of convex functions, preprint.

[15] G. BEER, "Topologies on Closed and Closed Convex Sets," Kluwer, Dordrecht, Boston, London, 1993.

[16] G. GEER AND J. M. BORWEIN, Mosco convergence and reflexivity, *Proc. Amer. Math. Soc.* **109** (1990), 427–436.

[17] G. BEER AND R. LUCCHETTI, The epi-distance topology: continuity and stability results with applications to convex optimization problems, *Math. Oper. Res.* **17** (1992), 715–726.

[18] J. BERGH AND L. LÖFSTRÖM, "Interpolation Spaces," Springer Verlag, Berlin, Heidelberg, New York, 1976.

[19] J. BOMAN, The sum of two plane convex C^∞ sets is not always C^5, *Math. Scand.* **66** (1990), 216–224.

[20] J. BOMAN, Smoothness of sums of convex sets with real analytic boundaries, *Math. Scand.* **66** (1990), 225–230.

[21] M. BOUGEARD, J.-P. PENOT AND A. POMMELLET, Towards minimal assumptions for the infimal convolution regularization, *J. Approx. Theory* **64** (1991), 245–270.

[22] H. BRÉZIS, "Operateurs Maximaux Monotones et Semi-Groupes de Contractions dans les Espaces de Hilbert," North-Holland, Amsterdam, 1973.

[23] C. BYLKA AND W. ORLICZ, On some generalizations of the Young inequality, *Bull. Acad. Polon. Sér. Sci. Math.* **26** (1978), 115–123.

[24] F. H. CLARKE, "Optimization and Nonsmooth Analysis," John Wiley & Sons, New York, 1983.

[25] M. M. DAY, "Normed Linear Spaces," Third Edition, Springer-Verlag, 1973.

[26] A. L. DONTCHEV AND T. ZOLEZZI, "Well-Posed Optimization Problems," Lecture Notes in Mathematics **1543**, Springer-Verlag, Berlin, Heidelberg, 1993.

[27] I. EKELAND, On the variational principle, *J. Math. Anal. Appl.* **47** (1974), 324–353.

[28] I. EKELAND, Nonconvex minimization problems, *Bull. Amer. Math. Soc.* **1** (1979), 443–474.

[29] I. EKELAND AND R. TEMAM, "Convex Analysis and Variational Problems," North-Holland, Amsterdam, 1976.

[30] S.-L. ERIKSSON AND H. LEUTWILER, A potential-theoretic approach to parallel addition, *Math. Ann.* **274** (1986), 301–317.

[31] W. FENCHEL, "Convex Cones, Sets, and Functions," Lecture Notes, Princeton University, Princeton, 1953.

[32] S. FITZPATRICK AND R. R. PHELPS, Bounded approximants to monotone operators on Banach spaces, *Ann. Inst. H. Poincaré Anal. Non Linéaire* **9** (1992), 573–595.

[33] A. FOUGERES AND A. TRUFFERT, Régularisation S.C.I. et Γ-convergence: approximations inf-convolutives associées à un référentiel, *Ann. Mat. Pura Appl.* **152** (1988), 21–51.

[34] F. HAUSDORFF, Über halbstetige Funktionen und deren Verallgemeinerung, *Math. Zeit.* **5** (1919), 292–309.

[35] J.-B. HIRIART-URRUTY, Extension of Lipschitz functions, *J. Math. Anal. Appl.* **77** (1980), 539–544.

[36] J.-B. HIRIART-URRUTY, Lipschitz r-continuity of the approximative subdifferential of a convex function, *Math. Scand.* **47** (1980), 123–134.

[37] J.-B. HIRIART-URRUTY, A general formula on the conjugate of the difference of functions, *Canad. Math. Bull.* **29** (1986), 482–485.

[38] J.-B. HIRIART-URRUTY AND M.-L. MAZURE, Formulation variationelle de l'addition paralléle et de la soustraction paralléle d'opérateurs semi-définis positifs, *C. R. Acad. Sci. Paris* **302** (1986), 527–530.

[39] J.-B. HIRIART-URRUTY AND R. R. PHELPS, Subdifferential calculus using ε-subdifferentials, *J. Funct. Anal.* **118** (1993), 154–166.

[40] J.-B. HIRIART-URRUTY AND P. PLAZANET, Moreau's decomposition theorem revisited, *Ann. Inst. H. Poincaré Anal. Non Linéaire* **6** (1989), 325–338.

[41] A. D. IOFFE AND V. M. TIHOMIROV, "Theory of Extremal Problems," North-Holland, Amsterdam, 1979.

[42] V. JEYAKUMAR, Duality and infinite dimensional optimization, *Nonl. Anal. T. M. A.* **15** (1990), 1111–1122.

[43] C. O. KISELMAN, How smooth is the shadow of a smooth convex body?, *J. London Math. Soc.* **33** (1986), 101–109.

[44] C. O. KISELMAN, Smoothness of vector sums of plane convex sets, *Math. Scand.* **60** (1987), 239–252.

[45] C. O. KISELMAN, Regularity classes for operations in convexity theory, *Kodai Math. J.* **15** (1992), 354–374.

[46] F. KUBO, Conditional expectations and operations derived from network connections, *J. Math. Anal. Appl.* **80** (1981), 477–489.

[47] J. LAHRACHE, Stabilite et convergence dans les espaces non reflexifs, *Sém. Anal. Convexe Montpellier* **21** (1991), 10.1–10.53.

[48] J.-M. LASRY AND P.-L. LIONS, A remark on regularization in Hilbert spaces, *Israel J. Math.* **55** (1986), 257–266.

[49] P.-J. LAURENT, "Approximation et optimisation," Hermann, Paris 1972.

[50] P. D. LAX, Hyperbolic systems of conservation laws, *Comm. Pure Appl. Math.* **10** (1957), 537–566.

[51] P. L. LIONS, "Generalized Solutions of Hamilton–Jacobi Equations," Pitman, London 1982.

[52] L. MALIGRANDA, "Orlicz Spaces and Interpolation," Sem. Math. 5, Campinas 1989.

[53] L. MALIGRANDA AND L. E. PERSSON, Generalized duality of some Banach function spaces, *Indagationes Math.* **51** (1989), 323–338.

[54] L. MALIGRANDA AND L. E. PERSSON, The E-functional for some pairs of groups, *Results in Mathematics* **20** (1991), 538–553.

[55] D. H. MARTIN, Some function classes closed under infimal convolution, *J. Opt. T. A.* **25** (1978), 579–584.

[56] M.-L. MAZURE, "Analyse variationelle des formes quadratiques convexes," Ph.D. Thesis, Université Paul Sabatier, Toulouse, 1986.

[57] M.-L. MAZURE AND M. VOLLE, Équations inf-convolutives et conjugaison de Moreau-Fenchel, *Annales de la Faculté des Sciences de Toulouse* **12** (1991), n° 1.

[58] L. MCLINDEN AND R. C. BERGSTROM, Preservation of convergence of convex sets and functions in finite dimensions, *Trans. Amer. Math. Soc.* **268** (1981), 127–142.

[59] J.-J. MOREAU, Inf-convolution, *Sém. d'Math. Montpellier* (1963), 3.1–3.48.

[60] J.-J. MOREAU, Inf-convolution des fonctions numériques sur un espace vectoriel, *C. R. Acad. Sci. Paris* **256** (1963), 5047–5049.

[61] J.-J. MOREAU, Proximité et dualité dans un espace Hilbertien, *Bull. Soc. Math. France* **93** (1965), 273–299.

[62] J.-J. MOREAU, "Fonctionnelles convexes," Lecture notes, Collège de France, Paris, 1967.

[63] J.-J. MOREAU, Inf-convolution, sous-additivité, convexité des fonctions numérique, *J. Math. Pures et Appl.* **49** (1970), 109–154.

[64] J.-J. MOREAU, Weak and strong solutions of dual problems, *in* "Contributions to Nonlinear Functional Analysis" (E. Zarantello, Ed.), Academic Press 1971.

[65] T. D. MORLEY, Parallel summation, Maxwell's principle and the infimum of projections, *J. Math. Anal. Appl.* **70** (1979), 33–41.

[66] U. MOSCO, Convergence of convex sets and of solutions of variational inequalities, *Adv. Math.* **3** (1969), 510–585.

[67] U. MOSCO, On the continuity of the Young–Fenchel transform, *J. Math. Anal. Appl.* **35** (1971), 518–535.

[68] G.B. PASSTY, The parallel sum of nonlinear monotone operators, *Nonl. Anal. T. M. A.* **10** (1986), 215–227.

[69] J.-P. PENOT AND M. L. BOUGEARD, Approximation and decomposition properties of some classes of locally d.c. functions, *Math. Prog.* **41** (1988), 195–227.

[70] R. R. PHELPS, "Convex Functions, Monotone Operators and Differentiability," Lecture Notes in Mathematics **1364**, Springer-Verlag, 1989.

[71] R. T. ROCKAFELLAR, Extension of Fenchel's duality theorem, *Duke Math. J.* **33** (1966), 81–89.

[72] R. T. ROCKAFELLAR, "Convex Analysis," Princeton University Press, Princeton, New York 1970.

[73] B. RODRIGUES, The Fenchel duality theorem in Fréchet spaces, *Optimization* **21** (1990), 13–22.

[74] B. RODRIGUES AND S. SIMONS, Conjugate functions and subdifferentials in nonnormed situations for operators with complete graphs, *Nonl. Anal. T. M. A.* **12** (1988), 1069–1078.

[75] W. RUDIN, "Functional Analysis," MacGraw-Hill, New York, 1973.

[76] A. SEEGER, Direct and inverse addition in convex analysis and applications, *J. Math. Anal. Appl.* **148** (1990), 317–343.

[77] S. SIMONS, The occasional distributivity of $\circ$ over $+_e$ and the change of variable formula for conjugate functions, *Nonl. Anal. T. M. A.* **14** (1990), 1111–1120.

[78] T. STRÖMBERG, An operation connected to a Young-type inequality, *Math. Nachr.* **159** (1992), 227–243.

[79] T. STRÖMBERG, Representation formulae for infimal convolution with applications, *in* "Analysis, Algebra, and Computers in Mathematical Research—Proceedings of the Twenty-First Nordic Congress of Mathematicians" (M. Gyllenberg, L. E. Persson, Eds.), Marcel Dekker, to appear in April 1994.

[80] T. STRÖMBERG, On the epigraphical sum of functions uniformly continuous on bounded sets, *Sém. Anal. Convexe Montpellier* **22** (1992), 19.1–19.16.

[81] T. STRÖMBERG, On the operation of infimal convolution and regularization of lower semicontinuous function, *submitted.*

[82] M. VOLLE, Sur quelces formules de dualité convexe et non convexe, *Set-Valued Analysis*, to appear.

[83] M. VOLLE, Complements sur la relation entre la regularisation de Lasry–Lions et l'equation de Hamilton–Jacobi, *Sém. Anal. Convexe Montpellier* **20** (1990), 7.1–7.10.

[84] M. VOLLE, Regularisation des fonctions fortement minorees dans les espaces de Hilbert, *Sém. Anal. Convexe Montpellier* **20** (1990), 8.1–8.8.

[85] M. VOLLE, Some applications of the Attouch–Brezis condition to closedness criterions, optimization, and duality, *Sém. Anal. Convexe Montpellier* **22** (1992), 16.1–16.15.

[86] R. WETS, Convergence of convex functions, variational inequalities and convex optimization problems, *in* "Variational Inequalities and Complementary Problems" (P. Cottle *et al.*, Eds.), Wiley, Chichester, 1980.

[87] C. ZĂLINESCU, On uniformly convex functions, *J. Math. Anal. Appl.* **95** (1983), 344–374.

B

Math. Nachr. **159** (1992) 227–243

An Operation Connected to a YOUNG-Type Inequality

By THOMAS STRÖMBERG*) of Luleå

(Received March 11, 1992)

Abstract. Given two φ-functions F and G we consider the largest φ-function $H = F \oplus G$ such that the YOUNG-*type inequality*

$$H(xy) \leq F(x) + G(y)$$

holds for all $x, y > 0$. We prove an equivalence theorem for $F \oplus G$ with the best constants and, for the special case when F and G are log-convex and satisfy a certain growth condition, a representation formula for $F \oplus G$. Moreover, further properties and examples are presented and the relations to similar results are discussed.

1. Introduction

Throughout this paper $\mathbf{R}$ denotes the set of real numbers, $\mathbf{R}_+ = [0, \infty[$ and $\mathbf{R}_{++} =]0, \infty[$. By $\mathscr{P}$ we understand the set of increasing, right-continuous and unbounded functions from $\mathbf{R}_+$ into $\mathbf{R}_+$ vanishing at and only at the origin. A *continuous* function in $\mathscr{P}$ is called a *φ-function.* For any $p \in \mathscr{P}$ the *right-inverse* $p^{(-1)}\colon \mathbf{R}_+ \to \mathbf{R}_+$ to p is defined by

$$p^{(-1)}(t) = \begin{cases} 0 & \text{if} \quad t = 0, \\ \sup\{s \in \mathbf{R}_+ \mid p(s) \leq t\} & \text{if} \quad t > 0. \end{cases}$$

If p has an inverse function, that is, if p is continuous and strictly increasing, then $p^{(-1)}$ is the inverse function p^{-1} to p. The *inequality of* YOUNG states that

$$xy \leq \int_0^x p(s)\,ds + \int_0^y p^{(-1)}(t)\,dt \tag{1}$$

for any $p \in \mathscr{P}$ and all $x, y \in \mathbf{R}_+$. Moreover, equality in (1) holds if and only if

$$y = p(x) \quad \text{or} \quad x = p^{(-1)}(y).$$

For a proof of these assertions, see for instance [2]. More information about YOUNG's

*) The research reported here was supported in part by *The Royal Swedish Academy of Sciences.*

inequality in its classical form can be found in [3], [4], [5], [7], [9] and [10]. YOUNG's inequality has been generalized in several ways; see for instance PÁLES [16]. In [13] and [16] necessary and sufficient conditions for some YOUNG-type inequalities to hold are given.

In this paper we are concerned with the following problem: Given two φ-functions F and G we wish to determine the largest function $H\colon \mathbf{R}_+ \to \mathbf{R}_+$ such that the *generalized* YOUNG's *inequality*

$$H(xy) \leq F(x) + G(y) \tag{2}$$

holds true for all $x, y \in \mathbf{R}_+$. Inequalities of the type (2) have natural applications to the theory of products of ORLICZ spaces, see for example ANDO [1], DANKERT [8], MALIGRANDA [14] or MALIGRANDA-PERSSON [15]. In connection with inequality (2), it is natural to introduce the function $F \oplus G\colon \mathbf{R}_+ \to \mathbf{R}_+$ defined as

$$(F \oplus G)(u) = \begin{cases} 0 & \text{if} \quad u = 0, \\ \inf\{F(x) + G(y) \mid x, y > 0 \text{ and } xy = u\} & \text{if} \quad u > 0. \end{cases} \tag{3}$$

It turns out that also $F \oplus G$ is a φ-function, see Theorem 1. Thus, $\oplus$ is a binary operation on the set of φ-functions. This means that the largest function H for which (2) is satisfied is the φ-function $F \oplus G$. The operation $\oplus$ is obviously commutative and it is easy to verify that it is associative.

Some results concerning the operation $\oplus$ can be found in our *main reference* [6] and also in the references [14], [15] and [17]. In this paper we illustrate, unify and complement these results in various ways. In particular we prove an equivalence theorem (Theorem 1), a representation formula for $F \oplus G$ (see Theorem 2) and a sharp YOUNG-type inequality (Proposition 5).

This paper is organized in the following way. In section 2 we present some necessary results for log-convex φ-functions which are of independent interest. Section 3 contain a basic YOUNG-type inequality and a calculation of $F_1 \oplus F_2 \oplus \ldots \oplus F_n$ for a special choice of log-convex φ-functions $F_1, F_2, \ldots, F_n$. Section 4 and 5 are devoted to presentation, proof and discussion of the above mensioned equivalence theorem and representation formula, respectively. In section 6 we prove a sharp YOUNG-type inequality and point out also some other consequences of our results. Finally, in section 7 and section 8 we present some concrete calculations of $F \oplus G$ and give some concluding remarks, respectively.

Acknowledgement. I wish to thank LECH MALIGRANDA and LARS ERIK PERSSON for suggesting this problem to me and for stimulating and helpful discussions.

2. Log-convex φ-functions

In what follows the notion of log-convex φ-functions is of importance.

Definition 1. *A real-valued function φ on $\mathbf{R}_+$ is called* log-convex *if*

$$\varphi(x^\lambda y^{1-\lambda}) \leq \lambda\varphi(x) + (1-\lambda)\varphi(y) \tag{4}$$

for every $\lambda \in]0, 1[$ and all x and y in $\mathbf{R}_+$. It is called strictly log-convex *provided that inequality* (4) *is strict for every $\lambda \in]0, 1[$ and all x and y in $\mathbf{R}_+$ with $x \neq y$.*

Every log-convex φ-function F can be written in the form

$$F(x) = \Phi(\ln x) \tag{5}$$

for every $x > 0$, where Φ is a positive convex function on $\mathbf{R}$ satisfying

$$\lim_{x \to \infty} \Phi(x) = \infty \quad \text{and} \quad \lim_{x \to -\infty} \Phi(x) = 0.$$

(It suffices to put $\Phi(x) = F(e^x)$ for every $x \in \mathbf{R}$.) It holds that F is strictly log-convex if and only if Φ is strictly convex. Since Φ possesses a left- and right-derivative at each point, so does F.

Example 1. The functions $x \mapsto x^a$, $a > 0$, and $x \mapsto \ln(x+1)$ on $\mathbf{R}_+$ are strictly log-convex φ-functions.

Some properties follow at once from Definition 1.

Example 2. The following hold:

(a) Let F and G be log-convex φ-functions. Then the function $\mu F + \nu G$ is a log-convex φ-function for any $\mu, \nu > 0$.

(b) Given a log-convex φ-function F and a number $b > 0$, the function $x \mapsto F(x^b)$ is a log-convex φ-function.

(c) If F is a log-convex φ-function and G is a convex φ-function, then the composition $G \circ F$ is a log-convex φ-function.

Proposition 1. *The following assertions hold:*

(a) *Log-convex φ-functions are strictly increasing.*

(b) *Every convex φ-function is strictly log-convex.*

(c) *Let F and G be φ-functions and let $u > 0$. Then the infimum in* (3) *is attained for at least one pair (x, y) of positive numbers with $xy = u$. Especially, if F and G are log-convex φ-functions and F or G is strictly log-convex, then this minimizing pair is unique.*

A proof of the assertion (a) of Proposition 1 can implicitly be found in [6] but for the sake of completeness we include a separate proof.

Proof. (a) Let F be a log-convex φ-function. It suffices to show that the convex function Φ on $\mathbf{R}$ in the representation (5) of F is strictly increasing. Assume that $\Phi(s) = \Phi(t)$ for $s < t$. Take a point $r < s$ such that $\Phi(r) < \Phi(s)$. Since Φ is constant on $[s, t]$ it cannot be convex on $[r, t]$, a contradiction.

(b) Let F be a convex φ-function. If x and y are in $\mathbf{R}_+$, $x \neq y$ and $0 < \lambda < 1$, then, according to the AG-inequality, $x^\lambda y^{1-\lambda} < \lambda x + (1-\lambda) y$ which implies

$$F(x^\lambda y^{1-\lambda}) < F(\lambda x + (1-\lambda) y) \le \lambda F(x) + (1-\lambda) F(y).$$

(c) If F and G are φ-functions, then the function

$$x \mapsto F(x) + G\left(\frac{u}{x}\right)$$

is continuous and tends to infinity as $x \to 0+$ and as $x \to \infty$ for each fixed $u > 0$. Therefore, the infimum

$$(F \oplus G)(u) = \inf_{x > 0} \left\{ F(x) + G\left(\frac{u}{x}\right) \right\}$$

is attained.

Let F and G be log-convex such that either F or G is strictly log-convex. Assume that for some $u > 0$ there exist two *different* pairs (x, y) and (x', y') such that $xy = x'y' = u$ and

$$(F \oplus G)(u) = F(x) + G(y) = F(x') + G(y').$$

Then, with $x'' = \sqrt{xx'}$ and $y'' = \sqrt{yy'}$ we have $x''y'' = u$ and

$$F(x'') + G(y'') < \frac{1}{2}F(x) + \frac{1}{2}F(x') + \frac{1}{2}G(y) + \frac{1}{2}G(y') = (F \oplus G)(u).$$

This contradiction completes the proof. $\square$

3. A Basic Young-type Inequality and a Calculation of $F_1 \oplus F_2 \oplus \ldots \oplus F_n$

By induction it is easily proved that

$$(F_1 \oplus F_2 \oplus \ldots \oplus F_n)(u) = \min\left\{\sum_{i=1}^{n} F_i(x_i) \,\middle|\, x_1, x_2, \ldots, x_n > 0 \text{ and } \prod_{i=1}^{n} x_i = u\right\} \tag{6}$$

for any φ-functions $F_1, F_2, \ldots, F_n$ and all $u > 0$. Moreover, it is clear that the following *basic* Young-*type inequality* holds for all $x_1, x_2, \ldots, x_n$ in $\mathbf{R}_+$:

$$(F_1 \oplus F_2 \oplus \ldots \oplus F_n)(x_1 x_2 \ldots x_n) \le \sum_{i=1}^{n} F_i(x_i).$$

In general, it is not possible to calculate $F_1 \oplus F_2 \oplus \ldots \oplus F_n$ explicitly. However, for a special choice of log-convex φ-functions $F_1, F_2, \ldots, F_n$ we can determine numbers $x_1, x_2, \ldots, x_n$ at which, given $u > 0$, the minimum in equation (6) is attained. This means that in these cases we can calculate $F_1 \oplus F_2 \oplus \ldots \oplus F_n$.

Proposition 2. *Let F be a log-convex φ-function, $\beta_i \in \mathbf{R}_{++}$ for $i = 1, 2, \ldots, n$. Define the log-convex φ-functions $F_1, F_2, \ldots, F_n$ by*

$$F_i(x) = \beta_i F(x^{1/\beta_i})$$

for $i = 1, 2, \ldots, n$ and all $x \in \mathbf{R}_+$. Then

$$(F_1 \oplus F_2 \oplus \ldots \oplus F_n)(u) = \beta F(u^{1/\beta}) \quad \forall u \in \mathbf{R}_+,$$

where

$$\beta = \beta_1 + \beta_2 + \ldots + \beta_n.$$

In particular, $(F \oplus F)(u) = 2F(\sqrt{u})$ for all $u \in \mathbf{R}_+$.

Proof. By letting $x_i = u^{\beta_i/\beta}$ for $i = 1, 2, \ldots, n$ we obtain the inequality

$$(F_1 \oplus F_2 \oplus \ldots \oplus F_n)(u) \le \sum_{i=1}^{n} \beta_i F(u^{1/\beta}) = \beta F(u^{1/\beta}).$$

Moreover, if $\prod_{i=1}^{n} x_i = u$, then the log-convexity of F implies

$$F(u^{1/\beta}) = F\left(\prod_{i=1}^{n} [x_i^{1/\beta_i}]^{\beta_i/\beta}\right) \le \sum_{i=1}^{n} \frac{\beta_i}{\beta} F(x_i^{1/\beta_i})$$

from which it follows that

$$(F_1 \oplus F_2 \oplus \ldots \oplus F_n)(u) \geq \beta F(u^{1/\beta}),$$

and the proposition is proved. □

4. The Equivalence Theorem

In this section we prove that $F \oplus G$ is equivalent to $(F^{(-1)}G^{(-1)})^{(-1)}$ and that our equivalence constants are the best possible. We also point out some further properties of $F \oplus G$. Our *equivalence theorem* reads as follows:

Theorem 1. *Let F and G be φ-functions. Then $F \oplus G$ is a φ-function satisfying, for every $u > 0$, the inequality*

$$(F^{(-1)}G^{(-1)})^{(-1)}(u) < (F \oplus G)(u) \leq 2(F^{(-1)}G^{(-1)})^{(-1)}(u). \tag{7}$$

The function $(F^{(-1)}G^{(-1)})^{(-1)}$ is a φ-function and the constants 1 *and* 2 *in* (7) *are the best possible.*

Proof. Let F and G be arbitrary φ-functions. In order to prove (7) let, for each $u > 0$, $X(u)$ denote the set of solutions x to the equation

$$F(x) = G\left(\frac{u}{x}\right)$$

and define the number $m(u)$ as

$$m(u) = F(x) = G\left(\frac{u}{x}\right) \quad \text{for} \quad x \in X(u).$$

First we claim that $m(u) = (F^{(-1)}G^{(-1)})^{(-1)}(u)$. In order to prove this assertion we need some properties of right-inverses which will be of importance throughout the paper.

Lemma 1. *Let $p \in \mathscr{P}$. Then,*

(i) $p^{(-1)} \in \mathscr{P}$,
(ii) $(p^{(-1)})^{(-1)} = p$,
(iii) $p^{(-1)}(p(s)-) \leq s \leq p^{(-1)}(p(s))$ *for any* $s > 0$.

The proof of Lemma 1 follows from the definition of the right-inverse and some straightforward calculations so we leave out the details; compare with [2] and [6].

Lemma 2. *Let $p \in \mathscr{P}$. The following assertions hold:*

(i) *If $p(s-) < p(s)$ for some $s > 0$, then $p^{(-1)}$ is constant equal to s on $[p(s-), p(s)[$.*
(ii) *If $p^{(-1)}$ is constant equal to s on some interval $[a, b[$, then p is discontinuous at s and*
$$[a, b[\subseteq [p(s-), p(s)[.$$
(iii) *$p^{(-1)}$ is strictly increasing if and only if p is continuous.*

Proof. (i) Assume $p(s-)<p(s)$ for some $s>0$. For any t in the interval $[p(s-), p(s)[$ we have

$$p^{(-1)}(t)=\sup\{\sigma\in\mathbf{R}_+ \mid p(\sigma)\le t\}=\sup[0,s[=s.$$

(ii) Assume that $p^{(-1)}$ is constant equal to s on $[a,b[$. Then,

$$p(s)=(p^{(-1)})^{(-1)}(s)=\sup\{t>0 \mid p^{(-1)}(t)\le s\}\ge b$$

and

$$p(s-)\le a.$$

Hence, $p(s-)<p(s)$. From (i) and the fact that $p^{(-1)}(t)<s$ for $t<p(s-)$ and $p^{(-1)}(t)>s$ for $t>p(s)$, we deduce that $[a,b[\subseteq[p(s-), p(s)[$.

Assertion (iii) is a direct consequence of (i) and (ii). □

By Lemma 1 (iii), for each $u>0$ and all $x\in X(u)$,

$$F^{(-1)}(m(u)-)\,G^{(-1)}(m(u)-)\le \underbrace{x\frac{u}{x}}_{=u}\le F^{(-1)}(m(u))\,G^{(-1)}(m(u)).$$

The continuity of F and G implies that $F^{(-1)}G^{(-1)}$ is strictly increasing (see Lemma 2 (iii)). Hence, $m(u)=(F^{(-1)}G^{(-1)})^{(-1)}(u)$. Since $F^{(-1)}G^{(-1)}$ is a strictly increasing function in $\mathscr{P}$ it follows that $(F^{(-1)}G^{(-1)})^{(-1)}$ is a continuous function in $\mathscr{P}$, by Lemma 1 (i) and Lemma 2 (iii). In other words, $(F^{(-1)}G^{(-1)})^{(-1)}$ is a φ-function.

The inequality

$$F(x)+G\left(\frac{u}{x}\right)>\max\left\{F(x), G\left(\frac{u}{x}\right)\right\}\ge m(u),$$

shows the validity of the left-hand side inequality in (7), since the infimum in the definition of $F\oplus G$ is attained. The right-hand side inequality in (7) follows from

$$(F\oplus G)(u)\le\left(F(x)+G\left(\frac{u}{x}\right)\right)\bigg|_{x\in X(u)}=m(u)+m(u).$$

Let us show that the constants 1 and 2 in the inequality (7) are the best possible. If $F(x)=x$, then, by Proposition 2,

$$(F\oplus F)(u)=2\sqrt{u},$$

which is an example of equality in the right-hand side inequality of (7).

Now, let F_b be defined by $F_b(y)=y^b/b$ for $b>0$. A calculation yields

$$(F^{-1}F_b^{-1})^{-1}(u)=\frac{u^{b/(b+1)}}{b^{1/(b+1)}}$$

and

$$(F\oplus F_b)(u)=\left(1+\frac{1}{b}\right)u^{b/(b+1)}.$$

Thus,

$$\frac{(F\oplus F_b)(u)}{(F^{-1}F_b^{-1})^{-1}(u)}=\left(1+\frac{1}{b}\right)b^{1/(b+1)}\to 1 \quad\text{as}\quad b\to\infty.$$

Hence, the number 1 in the left-hand side inequality in (7) cannot be replaced by any larger number.

Next we prove that $F \oplus G$ is a φ-function. Obviously, $F \oplus G$ is increasing. In order to prove that $F \oplus G$ is left-continuous, it suffices to show that

$$\sup_{v<u}(F \oplus G)(v) \geq (F \oplus G)(u)$$

for each $u > 0$. Let $\{u_n\}_{n=1}^{\infty}$ be an increasing sequence of real numbers such that $0 < u_n < u$ for all n and $u_n \nearrow u$ as $n \to \infty$. Consider

$$\varphi(x,u) = F(x) + G\left(\frac{u}{x}\right)$$

and choose an interval $[a,b] \subset \mathbf{R}_{++}$ in such a manner that

$$\varphi(x,u_1) > (F \oplus G)(u)$$

for every $x \in \mathbf{R}_{++} \backslash [a,b]$. For each index n there exists x_n (see Proposition 1) such that

$$(F \oplus G)(u_n) = \varphi(x_n, u_n).$$

Moreover, each $x_n \in [a,b]$. (If $x \notin [a,b]$ then

$$\varphi(x,u_n) \geq \varphi(x,u_1) > (F \oplus G)(u) \geq (F \oplus G)(u_n).)$$

Thus, $\{x_n\}$ contains a subsequence $(x_{n_k}\}$ which converges to, say x, which implies that

$$\lim_{k\to\infty}(F \oplus G)(u_{n_k}) = \lim_{k\to\infty}\varphi(x_{n_k}, u_{n_k}) = \varphi(x,u) \geq (F \oplus G)(u).$$

Hence, $F \oplus G$ is left-continuous. Moreover, for each $u > 0$,

$$\inf_{v>u}(F \oplus G)(v) = \inf_{v>u}\inf_{xy=v}\{F(x)+G(y)\} = \inf_{xy=u}\{F(x)+G(y)\} = (F \oplus G)(u)$$

from which it follows that $F \oplus G$ is right-continuous on $\mathbf{R}_{++}$. Inequality (7) implies that $F \oplus G$ is continuous at the origin, vanishes only at the origin and the $F \oplus G$ is unbounded. Hence, $F \oplus G$ is a φ-function. The proof is complete. □

We conclude this section by stating some further properties of $F \oplus G$. Of special interest in the sequel is the following subset of the set of log-convex φ-functions.

Definition 2. *The subset of all log-convex φ-functions F satisfying the condition*

$$\lim_{x\to\infty}\frac{F(x)}{\ln x} = \infty$$

is denoted by $\mathscr{L}$.

Example 3. Examples of functions in $\mathscr{L}$ are

(a) $x \mapsto x^b$ on $\mathbf{R}_+$ for any $b > 0$,

(b) any convex φ-function. Indeed, if F is a convex φ-function, then $x \mapsto F(x)/x$ is an increasing function from which it follows that

$$\frac{F(x)}{\ln x} = \frac{F(x)}{x}\,\frac{x}{\ln x} \to \infty \quad \text{as} \quad x \to \infty.$$

The log-convex φ-functions $x \mapsto \ln(x+1)$ does *not* belong to $\mathscr{L}$.

Proposition 3. *The following are true:*

(a) *Let F and G be φ-functions. If F or G is concave, then $F \oplus G$ is concave.*

(b) *The set of log-convex φ-functions and the set $\mathscr{L}$ are closed under the operation $\oplus$.*

Assertion (b) of Proposition 3 is formulated in [6] by BYLKA and ORLICZ. The proof given here is only slightly different from the original one.

Proof. (a) Assume that G is concave. Then the function

$$u \mapsto F(x) + G\left(\frac{u}{x}\right)$$

is concave for each $x > 0$. Hence, as a pointwise infimum of a set of concave functions, $F \oplus G$ is concave on $\mathbf{R}_{++}$ and also on $\mathbf{R}_+$ since it is continuous at the origin.

(b) Assume that F and G are log-convex and choose $u_1, u_2 > 0$ arbitrarily. Then, for each $i \in \{1, 2\}$, there exists a pair (x_i, y_i) of positive numbers such that $x_i y_i = u_i$ and

$$(F \oplus G)(u_i) = F(x_i) + G(y_i).$$

Hence, for every $\lambda \in]0, 1[$,

$$\begin{aligned}(F \oplus G)(u_1^{\lambda} u_2^{1-\lambda}) &\le F(x_1^{\lambda} x_2^{1-\lambda}) + G(y_1^{\lambda} y_2^{1-\lambda}) \\ &\le \lambda(F(x_1) + G(y_1)) + (1-\lambda)(F(x_2) + G(y_2)) \\ &= \lambda(F \oplus G)(u_1) + (1-\lambda)(F \oplus G)(u_2).\end{aligned}$$

We conclude that $F \oplus G$ is log-convex.

Now, let F and G belong to $\mathscr{L}$. In order to show that

$$\lim_{u \to \infty} \frac{(F \oplus G)(u)}{\ln u} = \infty$$

it is sufficient (and necessary) to prove that

$$\lim_{u \to \infty} \frac{(F^{-1} G^{-1})^{-1}(u)}{\ln u} = \infty,$$

see inequality (7). We have

$$\lim_{u \to \infty} \frac{\ln u}{(F^{-1} G^{-1})^{-1}(u)} = \lim_{v \to \infty} \frac{\ln [F^{-1}(v)\, G^{-1}(v)]}{v}$$

$$= \lim_{v \to \infty} \left(\frac{\ln F^{-1}(v)}{v} + \frac{\ln G^{-1}(v)}{v}\right) = \lim_{u \to \infty} \left(\frac{\ln u}{F(u)} + \frac{\ln u}{G(u)}\right) = 0 + 0,$$

which was to prove. $\square$

5. The Representation Theorem

In [12] it is proved that every function F in $\mathscr{L}$ can be represented as an integral

(8) $$F(x) = \int_0^x f(\xi) \frac{d\xi}{\xi} \quad \forall x > 0,$$

where f belongs to $\mathscr{P}$. Converserly, if $f \in \mathscr{P}$ is such that the function F defined by (8) and $F(0)=0$ *is finite*, then F is a member of $\mathscr{L}$. Let us by $\mathscr{P}_0$ denote the following subset of $\mathscr{P}$:

$$\mathscr{P}_0 = \left\{ p \in \mathscr{P} \;\middle|\; \int_0^1 p(s)\frac{ds}{s} < \infty \right\}.$$

Then, the mapping $f \mapsto F$, $\mathscr{P}_0 \to \mathscr{L}$ defined by (8) and $F(0)=0$ is a bijection. We know that the set $\mathscr{L}$ is closed under $\oplus$, see Proposition 3. The following question then naturally arises:

Given $f \in \mathscr{P}_0$ corresponding to $F \in \mathscr{L}$ and $g \in \mathscr{P}_0$ corresponding to $G \in \mathscr{L}$, which function in $\mathscr{P}_0$ corresponds to $F \oplus G \in \mathscr{L}$?

This question is completely answered in this section. Let us first assume that there are given φ-functions F and G in $\mathscr{L}$ such that the corresponding functions $f \in \mathscr{P}_0$ and $g \in \mathscr{P}_0$ in the integral representation (8) are differentiable with positive derivatives on $\mathbf{R}_{++}$. Then, f^{-1} and g^{-1} are the (on $\mathbf{R}_{++}$ differentiable) inverse functions to f and respectively. Put

$$\varphi(x,u) = F(x) + G\left(\frac{u}{x}\right).$$

Since f and g are strictly increasing, F and G are strictly log-convex. Then, $(F \oplus G)(u) = \min_{x>0} \varphi(x,u)$ and the minimum is uniquely attained at the solution $x = \bar{x}(u)$ to the equation

$$\partial_1 \varphi(x,u) = 0.$$

(Here, $\partial_1 \varphi$ denotes the partial derivative of φ with respect to the first variable.) This equation can be written as

$$f(x) = g\left(\frac{u}{x}\right).$$

A comparison with the proof of Theorem 1 yields

$$f(\bar{x}(u)) = g\left(\frac{u}{\bar{x}(u)}\right) = (f^{-1}g^{-1})^{-1}(u).$$

Clearly, $u \mapsto (F \oplus G)(u) = \varphi(\bar{x}(u), u)$ is differentiable on $\mathbf{R}_{++}$ and

$$\begin{aligned}(F \oplus G)'(u) &= \partial_1 \varphi(\bar{x}(u), u)\,\bar{x}'(u) + \partial_2 \varphi(\bar{x}(u), u)\\ &= 0 + \frac{1}{u} g\left(\frac{u}{\bar{x}(u)}\right) = \frac{1}{u}(f^{-1}g^{-1})^{-1}(u)\end{aligned}$$

for all $u > 0$. We have proved the following *representation theorem* for the special case when f and g are differentiable with positive derivatives on $\mathbf{R}_{++}$.

Theorem 2. *Let F and G in $\mathscr{L}$ be represented as*

$$F(x) = \int_0^x f(\xi)\frac{d\xi}{\xi} \quad \forall x > 0,$$

and

$$G(y)=\int_0^y g(\eta)\,\frac{d\eta}{\eta}\quad \forall y>0,$$

where f and g belong to $\mathscr{P}_0$. Then the following representation formula for the function $F\oplus G$ holds:

$$(F\oplus G)(u)=\int_0^u (f^{(-1)}g^{(-1)})^{(-1)}(s)\frac{ds}{s}\quad \forall u>0.$$

Remark. A similar result under more restrictive conditions has been proved by BYLKA-ORLICZ [6]. Our proof is also quite different.

Proof. The existence of sequences $\{f_n\}$ and $\{g_n\}$ of differentiable functions in $\mathscr{P}_0$ having positive derivatives on $\mathbf{R}_{++}$ such that, for every $s>0$, $f_n(s)\searrow f(s)$ and $g_n(s)\searrow g(s)$, follows from the following lemma.

Lemma 3. *For every $p\in\mathscr{P}_0$ there exists a decreasing sequence $\{p_n\}$ of differentiable functions in $\mathscr{P}_0$ having positive derivatives on $\mathbf{R}_{++}$ which satisfies*

$$p_n(s)\searrow p(s)\quad as\quad n\to\infty$$

for each $s>0$.

Proof. Let $p\in\mathscr{P}_0$ and define for each positive integer n the function $\bar p_n$ on $\mathbf{R}_+$ by $\bar p_n(0)=0$ and

$$\bar p_n(s)=\frac{n}{s}\int_s^{s(1+1/n)}\left[\frac{n}{\sigma}\int_\sigma^{\sigma(1+1/n)} p(\rho)\,d\rho\right]d\sigma$$

for all $s>0$. Then, $\{\bar p_n\}$ is a decreasing sequence of continuously differentiable and increasing functions. The inequality

$$p(s)\le \bar p_n(s)\le p\left(s\left(1+\frac{1}{n}\right)^2\right)$$

implies, firstly, that $\bar p_n$ belongs to $\mathscr{P}$ and, secondly, that $\bar p_n(s)\searrow p(s)$ for each $s>0$ since p is right-continuous. Let p_n be defined by

$$p_n(s)=\left(1+\frac{s}{n}\right)\bar p_n(s).$$

Clearly, $\{p_n\}$ is a decreasing sequence of differentiable functions in $\mathscr{P}$ with positive derivatives on $\mathbf{R}_{++}$. Moreover, for all n,

$$\int_0^1 p_n(s)\frac{ds}{s}\le\int_0^1\left(1+\frac{1}{n}\right)p\left(s\left(1+\frac{1}{n}\right)^2\right)\frac{ds}{s}=\int_0^{(1+1/n)^2}\left(1+\frac{1}{n}\right)p(\sigma)\frac{d\sigma}{\sigma}<\infty,$$

and the lemma is proved. □

Let

$$F_n(x) = \int_0^x f_n(\xi)\frac{d\xi}{\xi}$$

and

$$G_n(y) = \int_0^y g_n(\eta)\frac{d\eta}{\eta}.$$

The Monotone Convergence Theorem implies that $F_n \searrow F$ and $G_n \searrow G$ pointwise as $n \to \infty$. Thus, for every $u > 0$,

$$\begin{aligned}\lim_{n\to\infty}(F_n \oplus G_n)(u) &= \inf_n (F_n \oplus G_n)(u) = \inf_n \inf_{xy=u}\{F_n(x) + G_n(y)\}\\ &= \inf_{xy=u}\inf_n\{F_n(x)+G_n(y)\} = \inf_{xy=u}\{F(x)+G(y)\} = (F\oplus G)(u).\end{aligned}$$

We have already proved that

$$(F_n \oplus G_n)(u) = \int_0^u (f_n^{-1}g_n^{-1})^{-1}(s)\frac{ds}{s}.$$

Next we claim that

$$(f_n^{-1}g_n^{-1})^{-1} \searrow (f^{(-1)}g^{(-1)})^{(-1)} \quad \text{as} \quad n\to\infty.$$

In order to prove this assertion we need the following lemma which is also of independent interest.

Lemma 4. *If $\{p_n\}$ is a decreasing sequence of functions in $\mathscr{P}$ converging pointwise to $p \in \mathscr{P}$, then*

$$\lim_{n\to\infty} p_n^{(-1)}(t) = p^{(-1)}(t)$$

for every $t \in \mathbf{R}_{++}$ outside a subset, which is at most countable.

Proof. First we notice that $\{p_n^{(-1)}\}$ is an increasing sequence of functions such that $p_n^{(-1)} \leq p^{(-1)}$. Let E be the set of points $t > 0$ for which there are two different points s and s' such that $t = p(s) = p(s')$. Choose $t \in \mathbf{R}_{++}\setminus E$, let $s = p^{(-1)}(t)$ and consider an arbitrary $\varepsilon \in]0, s[$. Then $p(s-\varepsilon) < t$. Indeed, by Lemma 1, $p(s-\varepsilon) \leq t$ but $p(s-\varepsilon) = t$ is not possible, since in that case we would have $p(\sigma) = t$ for every $\sigma \in [s-\varepsilon, s[$, contradicting $t \notin E$. Thus there exists a positive integer N such that $p_n(s-\varepsilon) \leq t$ for all $n \geq N$, from which it follows that

$$p_n^{(-1)}(t) \geq s - \varepsilon = p^{(-1)}(t) - \varepsilon$$

for all $n \geq N$. Hence, $p_n^{(-1)}(t) \nearrow p^{(-1)}(t)$ as $n \to \infty$.

It remains to show that E is at most countable. For each $t \in E$ let

$$\mathscr{I}(t) = \{s > 0 \mid p(s) = t\}$$

which is an interval. Obviously, $\mathscr{I}(t_1)$ and $\mathscr{I}(t_2)$ are disjoint if $t_1 \neq t_2$. With every point t of E we associate a rational number $r(t) \in \mathscr{I}(t)$. We have thus established a bijection between the set E and a subset of the set of rational numbers. The latter is countable. The lemma is proved. $\square$

Now, define $h_n = (f_n^{-1} g_n^{-1})^{-1}$ and let h be the pointwise limit of the decreasing sequence of functions $\{h_n\}$. Since, $h = \inf_n h_n$ it follows that h is right-continuous and, hence, an element of $\mathscr{P}$. Moreover, $h_n^{-1} = f_n^{-1} g_n^{-1}$ and, according to Lemma 4, $h^{(-1)} = f^{(-1)} g^{(-1)}$ outside a set which is at most countable. But both $h^{(-1)}$ and $f^{(-1)} g^{(-1)}$ are elements of $\mathscr{P}$ and must consequently coincide everywhere. Thus, $h = (f^{(-1)} g^{(-1)})^{(-1)}$ (see assertion (ii) of Lemma 1).

Finally, since $(f_n^{-1} g_n^{-1})^{-1} \searrow (f^{(-1)} g^{(-1)})^{(-1)}$, the Monotone Convergence Theorem implies that

$$(F \oplus G)(u) = \lim_{n\to\infty} (F_n \oplus G_n)(u) = \lim_{n\to\infty} \int_0^u (f_n^{-1} g_n^{-1})^{-1}(s) \frac{ds}{s}$$

$$= \int_0^u (f^{(-1)} g^{(-1)})^{(-1)}(s) \frac{ds}{s}$$

for all $u > 0$. The proof of our representation theorem is complete. $\square$

6. Some Consequences of the Representation Theorem Including a Sharp Young-Type Inequality

From the representation formula in Theorem 2 there follow some results concerning the regularity and convexity of $F \oplus G$. By I we mean the φ-function $x \mapsto x$. Moreover, we say that a function p in $\mathscr{P}$ is *quasi-concave* if the function $s \mapsto p(s)/s$ on $\mathbf{R}_{++}$ is decreasing.

Proposition 4. *Let F and G in $\mathscr{L}$ be represented as*

$$F(x) = \int_0^x f(\xi) \frac{d\xi}{\xi} \quad \forall x > 0,$$

and

$$G(y) = \int_0^y g(\eta) \frac{d\eta}{\eta} \quad \forall y > 0,$$

where f and g belong to $\mathscr{P}_0$. Then the following hold:

(a) *$F \oplus G$ is continuously differentiable on $\mathbf{R}_{++}$ if and only if*

$$f(\mathbf{R}_+) \cup g(\mathbf{R}_+) = \mathbf{R}_+ .$$

In particular, if F or G is continuously differentiable on $\mathbf{R}_{++}$, then $F \oplus G$ is continuously differentiable on $\mathbf{R}_{++}$.

(b) *If f and g are k times continuously differentiable, $k \geq 1$, with f' and g' positive on* $\mathbf{R}_{++}$, *then $F \oplus G$ is $(k+1)$ times continuously differentiable on* $\mathbf{R}_{++}$.
(c) $F \oplus G = I$ *if and only if* $f^{(-1)} g^{(-1)} = I$.
(d) $F \oplus G$ *is convex if and only if* $f^{(-1)} g^{(-1)}$ *is quasi-concave.*

Remark. Assertion (c) of Proposition 4 is also proved in [6]. Part (d) is an answer to a question raised in [15].

Proof. (a) According to Theorem 2, Lemma 2 and some obvious arguments the following statements are equivalent:

- The function $F \oplus G$ is continuously differentiable on $\mathbf{R}_{++}$.
- The function $(f^{(-1)} g^{(-1)})^{(-1)}$ is continuous.
- The function $f^{(-1)} g^{(-1)}$ is strictly increasing.
- There does not exist an interval $[a, b[$ on which both $f^{(-1)}$ and $g^{(-1)}$ are constant.
- There do not exist an interval $[a, b[$ and numbers $s_1, s_2 > 0$ such that

$$[a, b[\ \subseteq [f(s_1 -), f(s_1)[\ \cap [g(s_2 -), g(s_2)[\,.$$

- $f(\mathbf{R}_+) \cup g(\mathbf{R}_+) = \mathbf{R}_+$.

The assertions (b) and (c) are obvious.

(d) Theorem 2 and some elementary arguments imply that the following statements are equivalent:

- The function $F \oplus G$ is convex.
- The function

$$s \mapsto \frac{(f^{(-1)} g^{(-1)})^{(-1)}(s)}{s}$$

is increasing.
- The function

$$t \mapsto \frac{t}{(f^{(-1)} g^{(-1)})(t)}$$

is increasing.
- The function

$$t \mapsto \frac{(f^{(-1)} g^{(-1)})(t)}{t}$$

is decreasing. $\square$

Proposition 5. *The* Young*-type inequality*

$$\int_0^{xy} (f^{(-1)} g^{(-1)})^{(-1)}(s) \frac{ds}{s} \leq \int_0^x f(\xi) \frac{d\xi}{\xi} + \int_0^y g(\eta) \frac{d\eta}{\eta} \tag{9}$$

holds for any f and g in $\mathcal{P}_0$ and all $x, y > 0$. Equality holds if and only if

$$f(x-) - g(y) \leq 0 \leq f(x) - g(y-). \tag{10}$$

Remark. Note that (10) reduces to "$f(x)=g(y)$" when f and g are continuous.

Proof. The functions F and G defined by $F(0)=G(0)=0$,

$$F(x)=\int_0^x f(\xi)\frac{d\xi}{\xi} \quad \forall x>0,$$

and

$$G(y)=\int_0^y g(\eta)\frac{d\eta}{\eta} \quad \forall y>0,$$

are elements of $\mathscr{L}$. The validity of inequality (9) follows from Theorem 2 and the definition of $\oplus$. We prove that (9) holds with equality if and only if (10) holds. First we note that $(F\oplus G)(u)=F(x)+G(u/x)$ if and only if

$$\frac{\partial^-}{\partial x}\left(F(x)+G\left(\frac{u}{x}\right)\right)\le 0\le\frac{\partial^+}{\partial x}\left(F(x)+G\left(\frac{u}{x}\right)\right). \tag{11}$$

By substitution of the identites

$$\frac{\partial^-}{\partial x}\left(F(x)+G\left(\frac{u}{x}\right)\right)=F'_-(x)-\frac{u}{x^2}G'_+\left(\frac{u}{x}\right)=\frac{1}{x}\left(f(x-)-g\left(\frac{u}{x}\right)\right)$$

and

$$\frac{\partial^+}{\partial x}\left(F(x)+G\left(\frac{u}{x}\right)\right)=F'_+(x)-\frac{u}{x^2}G'_-\left(\frac{u}{x}\right)=\frac{1}{x}\left(f(x)-g\left(\frac{u}{x}-\right)\right)$$

into (11), together with the change of variables $u=xy$, we arrive at (10). □

7. Examples

In this section we present two calculations of $F\oplus G$ for special choices of F and G.

Example 4. Let $F(x)=\ln(1+x)$, $x\ge 0$, and $G(y)=2y$, $y\ge 0$. Let us calculate $F\oplus G$. Put

$$\varphi(x,u)=F(x)+G\left(\frac{u}{x}\right).$$

Then, $(F\oplus G)(u)=\min\limits_{x>0}\varphi(x,u)$ and the minimum is attained at stationary points of $x\mapsto\varphi(x,u)$. A calculation yields

$$\partial_1\varphi(x,u)=0\Leftrightarrow\frac{1}{1+x}-\frac{2u}{x^2}=0.$$

The only stationary point is $x=u+\sqrt{u^2+2u}$. Thus

$$(F\oplus G)(u)=\varphi(u+\sqrt{u^2+2u},u)=\ln(1+u+\sqrt{u^2+2u})-u+\sqrt{u^2+2u}.$$

Example 5. Let

$$F(x)=\begin{cases} x & \text{if } 0\le x<1, \\ 2x-1 & \text{if } 1\le x, \end{cases}$$

and

$$G(y)=\begin{cases} y & \text{if } 0\le y<2, \\ 2y-2 & \text{if } 2\le y. \end{cases}$$

Then, f and g are given by

$$f(s)=\begin{cases} s & \text{if } 0\le s<1, \\ 2s & \text{if } 1\le s, \end{cases}$$

and

$$g(s)=\begin{cases} s & \text{if } 0\le s<2, \\ 2s & \text{if } 2\le s. \end{cases}$$

We find that

$$f^{(-1)}(t)=\begin{cases} t & \text{if } 0\le t<1, \\ 1 & \text{if } 1\le t<2, \\ t/2 & \text{if } 2\le t, \end{cases}$$

and

$$g^{(-1)}(t)=\begin{cases} t & \text{if } 0\le t<2, \\ 2 & \text{if } 2\le t<4, \\ t/2 & \text{if } 4\le t. \end{cases}$$

Thus,

$$(f^{(-1)}g^{(-1)})(t)=\begin{cases} t^2 & \text{if } 0\le t<1, \\ t & \text{if } 1\le t<4, \\ t^2/4 & \text{if } 4\le t, \end{cases}$$

$$(f^{(-1)}g^{(-1)})^{(-1)}(s)=\begin{cases} \sqrt{s} & \text{if } 0\le s<1, \\ s & \text{if } 1\le s<4, \\ 2\sqrt{s} & \text{if } 4\le s, \end{cases}$$

which yields, according to Theorem 2,

$$(F\oplus G)(u)=\int_0^u (f^{(-1)}g^{(-1)})^{(-1)}(s)\frac{ds}{s}=\begin{cases} 2\sqrt{u} & \text{if } 0\le u<1, \\ 1+u & \text{if } 1\le u<4, \\ 4\sqrt{u}-3 & \text{if } 4\le u. \end{cases}$$

Remark. In particular, Example 5 shows that $F\oplus G$ can be continuously differentiable though neither F nor G is continuously differentiable. This is, in fact, a consequence of assertion (a) of Proposition 4.

8. Concluding Remarks

The so-called *infimal convolution* $\varphi \square \psi$ of the functions φ and ψ on $\mathbf{R}$, defined as

$$(\varphi \square \psi)(x) = \inf\{\varphi(y) + \psi(z) \mid y, z \in \mathbf{R} \text{ and } y + z = x\}$$

for all real numbers x, is closely connected with the operation $\oplus$. Indeed, if F and G are φ-functions we may define functions $\tilde{F}$ and $\tilde{G}$ on $\mathbf{R}$ by

$$\tilde{F}(x) = F(e^x) \quad \text{and} \quad \tilde{G}(x) = G(e^x)$$

for all real numbers x. Then

$$(\tilde{F} \square \tilde{G})(x) = \inf\{\tilde{F}(y) + \tilde{G}(z) \mid y, z \in \mathbf{R} \text{ and } y + z = x\}$$
$$= \inf\{F(e^y) + G(e^z) \mid y, z \in \mathbf{R} \text{ and } e^y e^z = e^x\} = (F \oplus G)(e^x)$$

for all real numbers x. This means that every result in this paper concerning the operation $\oplus$ can be translated into a corresponding result for infimal convolution. For more on infimal convolution see for instance [11].

Every results in this paper that is valid for two functions can easily be generalized to a result for an arbitrary finite number of functions. As an example, let $F_1, F_2, \ldots, F_n$ be n functions in $\mathscr{L}$ and let f_i in $\mathscr{P}_0$ correspond to F_i for $i = 1, 2, \ldots, n$. By induction on the representation formula in Theorem 2 it follows that the function $F = F_1 \oplus F_2 \oplus \ldots \oplus F_n$ is given by

$$F(u) = \int_0^u f(s) \frac{ds}{s} \quad \forall u > 0,$$

where $f = (f_1^{(-1)} f_2^{(-1)} \ldots f_n^{(-1)})^{(-1)}$.

References

[1] Ando, T., On products of Orlicz spaces, Math. Ann. **140** (1960) 174–186
[2] Birnbaum, Z., and Orlicz, W., Über eine Verallgemeinerung des Begriffes der zueinander konjugierten Potenzen, Studia Math. **3** (1931) 1–67
[3] Boas, R. P., and Marcus, M. B., Inequalities involving a function and its inverse, SIAM J. Math. Anal. Appl. **4** (1973) 585–591
[4] –, –, Generalizations of Young's inequality, J. Math. Anal. Appl. **46** (1974) 36–40
[5] –, –, Inverse functions and integration by parts, Amer. Math. Monthly **81** (1974) 760–761
[6] Bylka, C., and Orlicz, W., On some generalizations of the Young inequality, Bull. Acad. Polon. Sci. Sér. Sci. Math. **26** (1978) 115–123
[7] Cunningham, F., and Grossman, N., On Young's inequality, Amer. Math. Monthly **78** (1971) 781–783
[8] Dankert, G., Über Produkte von Orlicz-Räumen, Arch. Math. **19** (1968) 635–645
[9] Diaz, J. B., and Metcalf, F. T., An analytic proof of Young's inequality, Amer. Math. Monthly **77** (1970) 603–609
[10] Klambauer, G., Integration by parts and inverse functions, Amer. Math. Monthly **85** (1978) 668–699
[11] Laurent, P.-J., Approximation et optimisation, Hermann, Paris 1972
[12] Leśniewicz, R., On Hardy-Orlicz spaces, I, Comm. Math. **15** (1971) 3–56
[13] Losonczi, L., Inequalities of Young-type, Mh. Math. **97** (1984) 125–132
[14] Maligranda, L., Orlicz spaces and interpolation, Seminars in Math. **5** Campinas 1989

[15] Maligranda, L., and Persson, L. E., Generalized duality of some Banach function spaces, Indagationes Math. **51** (1989) 323–338
[16] Páles, Z., On Young-type inequalities, Acta Sci. Math. **54** (1990) 327–338
[17] Zabreiko, P. P., and Rutickiĭ, Ja. B., Some remarks on monotone functions, Funkcional. analiz i teoria funkcii, Kazan' (1976) 114–126

Department of Applied Mathematics
Luleå University of Technology
S-95187 Luleå
Sweden

C

Representation Formulae for Infimal Convolution with Applications

Thomas Strömberg

Department of Mathematics
Luleå University of Technology
S-951 87 Luleå, Sweden

Abstract

This paper is concerned with the operation of *infimal convolution*. In particular, the following formulae for the infimal convolution $f \Box g$ of f and g are proved under the assumptions that f and g are strictly convex, coercive and Gâteaux-differentiable functions defined on a real reflexive Banach space:

$$\begin{aligned} f \Box g &= (f \circ f'^{-1} + g \circ g'^{-1}) \circ (f' // g'), \\ (f \Box g)' &= f' // g', \end{aligned}$$

where $f' // g' = (f'^{-1} + g'^{-1})^{-1}$ (the parallel sum of f' and g').

Some further results are presented including an application to a Hamilton–Jacobi equation and an example from mechanics.

1 Introduction

Let X denote a real vector space. To any function $f: X \to \bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ we associate the *epigraph* of f

$$\operatorname{epi} f = \{(x, \alpha) \in X \times \mathbb{R} | \; f(x) \leq \alpha\},$$

and the *strict epigraph*

$$\operatorname{epi}_s f = \{(x, \alpha) \in X \times \mathbb{R} | \; f(x) < \alpha\}.$$

Many properties of f can in a natural and appealing way be expressed in terms of epigraphs. For instance, f is convex if and only if $\operatorname{epi} f$ (or $\operatorname{epi}_s f$) is a convex set, whereas f lower semicontinuous is equivalent to $\operatorname{epi} f$ closed when X is a topological vector space.

Moreover, differentiability theory for convex and lower semicontinuous functions is most fruitfully developed in terms of supporting hyperplanes to the epigraphs, giving rise to the theory of subdifferentials.

A natural step is to apply set operations on epigraphs. The present paper deals with the Minkowski sum of strict epigraphs; Attouch and Wets [4] use the terminology "epigraphical sum": Given two extended real valued functions f and g defined on X, the addition

$$\mathrm{epi}_s\, f + \mathrm{epi}_s\, g =: \mathrm{epi}_s\, f \Box g$$

gives rise to the strict epigraph of a function that we denote by $f \Box g$ and call the *infimal convolution* or the *epigraphical sum* of f and g. The following functional identity justifies the name "infimal convolution": the function $f \Box g$ assigns to each $x \in X$ the extended real number

$$(f \Box g)(x) = \inf\{f(y) + g(z) \mid y, z \in X \text{ and } y + z = x\}. \tag{1}$$

(By convention, the sums $(+\infty) + (-\infty)$ and $(-\infty) + (+\infty)$ equal $+\infty$.) Infimal convolution $\Box$ is a commutative and associative binary operation on $\bar{\mathbb{R}}^X$ [23]. The function $\mathcal{I}$, defined as

$$\mathcal{I}(x) = \begin{cases} 0 & \text{when } x = 0, \\ +\infty & \text{when } x \neq 0, \end{cases}$$

acts as the identity element for $\Box$. Therefore, $(\bar{\mathbb{R}}^X, \Box)$ is a commutative monoid. Infimal convolution was introduced by Fenchel [11] and is discussed in the books [16], [18], [23] and [25] by Ioffe & Tihomirov, Laurent, Moreau and Rockafellar, respectively. An overview of the basic concepts of epigraphical calculus and analysis can be found in the paper [4] by Attouch and Wets.

Infimal convolution plays a basic role in optimization mainly because of the relation

$$(f \Box g)^* = f^* + g^* \tag{2}$$

which is valid for any proper functions f and g on X, see [16, page 178], X being a topological vector space. Here, $*$ denotes Legendre-Fenchel transformation, see equation (3) for its definition. Another reason for studying infimal convolution has traditionally been its—*sometimes*—regularizing effect. Namely, roughly speaking, if the function g possesses a regularity property, then, often under very mild assumptions on f, $f \Box g$ inherits this property, see [2], [3], [4], [6], [12], [13], [14] or [15] for results in this direction. Especially for convex functions defined on a Hilbert space, the most famous regularization and approximation technique by means of infimal convolution is called Moreau-Yosida approximation, consult for instance [2]. However, the smoothing effect is even in the convex case up to some limit; examples of situations when regularity is lost when performing infimal convolution can be found in [7], [8] and in [17].

Our aim is somewhat different; we focus on *formulae* for $f \Box g$, as well as for its subdifferential $\partial(f \Box g)$, under the assumptions that f and g are proper, convex, lower semicontinuous and coercive functions defined on a reflexive Banach space. One motivation for studying how infimal convolution acts on such "potential energy type" of functions is that the potential energy function of a mechanical system consisting of two springs connected in series is equal to the infimal convolution of the respective potential energy functions of the springs. The subdifferential of the potential energy function of a spring can be interpreted as a relation between force and elongation.

This paper is organized as follows. In section 2 we introduce two sets of functions especially considered: the cone $\mathcal{C}(X)$ [resp., $\mathcal{C}_{\mathrm{w}}(X)$] of all proper lower semicontinuous convex and coercive [resp., weakly coercive] functions defined on a normed vector space X. In particular, assuming X to be a reflexive Banach space, it is proved that $\mathcal{C}(X)$ and $\mathcal{C}_{\mathrm{w}}(X)$ are closed under infimal convolution and, moreover, that the subdifferential $\partial(f \Box g)$ of $f \Box g$ is equal to the parallel sum of ∂f and ∂g, when and f and g belong to $\mathcal{C}_{\mathrm{w}}(X)$. In section 3 a representation formula for $f \Box g$ is presented for the case when f and g are Gâteaux-differentiable, strictly convex and coercive; X is again a reflexive Banach space. Section 4 contains a discussion of the mechanical example mentioned above. As another example, a special Hamilton–Jacobi equation is considered in section 5. In our final section we present a concluding remark concerning the relation between infimal convolution and an operation investigated by the author in the paper [26].

Throughout, an extended real valued function defined on X is called *proper* if it does not take the value $-\infty$ and if it is not identically equal to $+\infty$. By $\operatorname{dom} f$ we denote the *essential domain* of $f: X \to \bar{\mathbb{R}}$, that is,

$$\operatorname{dom} f = \{x \in X \mid f(x) < +\infty\}.$$

ACKNOWLEDGEMENTS. The author thanks Lars Erik Persson for his help and an anonymous referee for references and valuable valuable suggestions which have helped in improving the presentation.

The research reported here was supported in part by *Magnuson's fund* of *The Royal Swedish Academy of Sciences.*

2 General results for the infimal convolution of proper, lower semicontinuous, convex and coercive functions

Let X be a topological vector space with topological dual space X^*. The *subdifferential* $\partial f: X \to \mathcal{P}(X^*)$ of $f: X \to \bar{\mathbb{R}}$ is defined, for each $x \in X$, by setting $\partial f(x) = \emptyset$ if $f(x) = \pm\infty$ and

$$\partial f(x) = \{x^* \in X^* \mid \langle x^*, y - x\rangle \leq f(y) - f(x) \text{ for all } y \in X\}$$

if $f(x)$ is a finite number. The subdifferential generalizes the classical concept of a derivative in the case of proper, convex and lower semicontinuous functions. The subdifferential of a proper, convex and lower semicontinuous function defined on a reflexive Banach space is a maximal monotone mapping. In order to be able to formulate the following proposition we need a definition: for any mappings S and T from X into $\mathcal{P}(X^*)$ (that is, "multimappings" from X into X^*) we define the *parallel sum* $S//T: X \to \mathcal{P}(X^*)$ by

$$S//T = \{(x, x^*) \in X \times X^* | \text{ there exist } y, z \in X \\ \text{such that } y + z = x \text{ and } x^* \in S(y) \cap T(z)\}.$$

We note that

$$(S//T)^{-1} = S^{-1} + T^{-1}.$$

Parallel addition of positive linear operators $H \to H$, with H a Hilbert space, has been studied in [1], [10], [22] and in [24].

We are now in position to state a general result concerning the subdifferential $\partial(f \Box g)$. Compare with [18].

Proposition 1. *Assume that f and g are proper functions defined on a topological vector space X. If the infimum in the formula* (1) *for $(f \Box g)(x)$ is attained for each $x \in X$, then*

$$\partial(f \Box g) = \partial f // \partial g.$$

We recall that if $f : X \to \bar{\mathbb{R}}$, then the *conjugate function* or the *Legendre-Fenchel transform* $f^*: X^* \to \bar{\mathbb{R}}$ of f is defined by

$$f^*(x^*) = \sup\{\langle x^*, x\rangle - f(x) | \, x \in X\} \tag{3}$$

for every $x^* \in X^*$. General properties of conjugate functions can be found in [9], [16], [25] and [28]. If f is a proper function, then

$$\partial f = \{(x, x^*) \in X \times X^* | \, \langle x^*, x\rangle = f(x) + f^*(x^*)\},$$

see [28, page 490].

PROOF. Recall (2). The following equivalences hold for any $(x, x^*) \in X \times X^*$:

$$\begin{gathered} x^* \in (\partial f // \partial g)(x) \\ \Leftrightarrow \\ \exists y, z \in X \quad x = y + z \text{ and } x^* \in \partial f(y) \cap \partial g(z) \\ \Leftrightarrow \\ \exists y, z \in X \quad x = y + z \text{ and } \langle x^*, y\rangle = f(y) + f^*(x^*) \text{ and } \langle x^*, z\rangle = g(z) + g^*(x^*) \\ \Leftrightarrow \\ \exists y, z \in X \quad x = y + z \text{ and } \langle x^*, y + z\rangle = f(y) + g(z) + f^*(x^*) + g^*(x^*) \\ \Leftrightarrow \\ \langle x^*, x\rangle = (f \Box g)(x) + (f \Box g)^*(x^*) \\ \Leftrightarrow \\ x^* \in \partial(f \Box g)(x). \end{gathered}$$

The proof is complete. $\square$

A function $f: X \to \bar{\mathbb{R}}$, where $(X, \|\cdot\|)$ stands for normed vector space, will be called

- *coercive* if
$$\frac{f(x)}{\|x\|} \to +\infty \text{ as } \|x\| \to +\infty,$$
- *weakly coercive* if $f(x) \to +\infty$ as $\|x\| \to +\infty$.

Definition. *We shall denote by $\mathcal{C}(X)$ [resp., by $\mathcal{C}_{\mathrm{w}}(X)$] the cone of all proper lower semicontinuous convex and coercive [resp., weakly coercive] functions defined on X; X being a normed vector space.*

Note that $\mathcal{C}(X) \subset \mathcal{C}_{\mathrm{w}}(X)$. In the sequel we are concerned with the infimal convolution of functions belonging to $\mathcal{C}(X)$ or $\mathcal{C}_{\mathrm{w}}(X)$. In terms of epigraphs, f belongs to $\mathcal{C}_{\mathrm{w}}(X)$ if and only if its epigraph $\operatorname{epi} f$ is a non-empty closed convex subset of $X \times \mathbb{R}$ that contains no vertical lines, and such that the level set $\{x \mid (x, \alpha) \in \operatorname{epi} f\}$ is bounded for each real number α.

When X is a reflexive Banach space, any function in $\mathcal{C}_{\mathrm{w}}(X)$ has a finite minimum (see [9, page 35]), and if $f \in \mathcal{C}(X)$, then $\bigcup_{x \in X} \partial f(x) = X^*$.

Now we are in position to formulate the main result of this section.

Theorem 1. *Let X be a reflexive Banach space.*

(a) *If f and g belong to $\mathcal{C}_{\mathrm{w}}(X)$, then the infimum in the formula* (1) *for $(f \square g)(x)$ is achieved for each $x \in X$ and $\partial(f \square g) = \partial f // \partial g$.*

(b) *The cones $\mathcal{C}(X)$ and $\mathcal{C}_{\mathrm{w}}(X)$ are closed under infimal convolution.*

(c) *$f \square g = (f^* + g^*)^*$ for any elements f and g of $\mathcal{C}_{\mathrm{w}}(X)$.*

Proof. Choose f and g in $\mathcal{C}_{\mathrm{w}}(X)$ arbitrarily. (a) We note that the problem of finding $(f \square g)(x)$ for a fixed $x \in X$ can be reduced to finding the infimum of the function Ψ defined on X as follows:

$$\Psi(y) = f(y) + g(x - y) \text{ for all } y \in X.$$

We observe that Ψ is lower semicontinuous, convex and weakly coercive. From this follows that Φ has a minimum which is a number in $(-\infty, +\infty]$. According to Proposition 1, $\partial(f \square g) = \partial f // \partial g$.

(b) It is not hard to show that $\min f \square g = \min f + \min g$ and, consequently, $\min f \square g \in \mathbb{R}$. Hence, $f \square g$ is proper.

Lower semicontinuity. Let $(x_k)_{k \in \mathbb{N}}$ be a sequence in X that converges to x and let us set $\ell = \liminf_{k \to +\infty} (f \square g)(x_k)$. Obviously, $-\infty < \ell \leq +\infty$. For each $k \in \mathbb{N}$ there exists

y_k in X such that

$$(f\Box g)(x_k) = f(y_k) + g(x_k - y_k).$$

We claim that $(f\Box g)(x) \le \ell$ which is the case if $\ell = +\infty$. Assume $\ell < +\infty$. There exists a subsequence of (x_k), again denoted by (x_k), such that

$$f(y_k) + g(x_k - y_k) \to \ell$$

as $k \to +\infty$. Since f and g are weakly coercive and bounded from below, the sequence (y_k) must be bounded. Hence, since X is reflexive, we can extract a weakly convergent subsequence, for simplicity denoted by (y_k). Let y be its weak limit. Since f and g are convex and lower semicontinuous they are weakly sequentially lower semicontinuous, see [9, page 11]. Therefore,

$$\begin{aligned}
\ell &= \liminf_{k\to+\infty}(f(y_k) + g(x_k - y_k)) \\
&\ge \liminf_{k\to+\infty} f(y_k) + \liminf_{k\to+\infty} g(x_k - y_k) \\
&\ge f(y) + g(x - y) \ge (f\Box g)(x).
\end{aligned}$$

We conclude that $f\Box g$ is lower semicontinuous.

The *convexity* is clear. Indeed, $\mathrm{epi}_s\, f\Box g$ is a convex subset of $X \times \mathbb{R}$ since it is the Minkowski sum of the two convex sets $\mathrm{epi}_s\, f$ and $\mathrm{epi}_s\, g$.

Coerciveness. We show that $f\Box g$ is coercive when f and g are coercive. Take any sequence $(x_k)_{k\in\mathbb{N}}$ in X such that $\|x_k\| \to +\infty$ as k tends to infinity. We shall prove that

$$\lim_{k\to+\infty} \frac{(f\Box g)(x_k)}{\|x_k\|} = +\infty. \tag{4}$$

To each $k \in \mathbb{N}$ there correspond elements y_k and z_k of X such that

$$x_k = y_k + z_k \text{ and } (f\Box g)(x_k) = f(y_k) + g(z_k).$$

Choose, for each natural number k, u_k in $\{y_k, z_k\}$ in such a way that

$$\|u_k\| = \max\{\|y_k\|, \|z_k\|\}.$$

Then

$$\|x_k\| \le \|y_k\| + \|z_k\| \le 2\|u_k\|.$$

Moreover,

$$(f \wedge g)(x) := \min\{f(x), g(x)\} \ge \min\{\min f, \min g\} =: m$$

for all $x \in X$. We note that $\|u_k\| \to +\infty$ as k goes to infinity and that $f \wedge g$ is coercive. Therefore

$$\frac{(f\Box g)(x_k)}{\|x_k\|} \ge \frac{(f \wedge g)(u_k) + m}{2\|u_k\|} \to +\infty$$

as $k \to +\infty$ and (4) is proved.

(c) Since $f \Box g$ is convex and lower semicontinuous, $f \Box g = (f \Box g)^{**}$. Moreover, $(f \Box g)^{**} = (f^* + g^*)^*$, see (2). $\Box$

REMARK. The following subsets of $\bar{\mathbb{R}}^X$, where X stands for a reflexive Banach space, are also closed under $\Box$:

- The convex cone of all finitely valued functions in $\mathcal{C}(X)$.
- The convex cone of all finitely valued functions in $\mathcal{C}_{\mathrm{w}}(X)$.

3 A representation formula

Let X be a normed vector space. We recall that an operator $A: X \to X^*$ is called

- *monotone* if
$$\langle A(x) - A(y), x - y \rangle \geq 0 \tag{5}$$
for all x and y in X, and *strictly monotone* if inequality (5) is strict for all x and y in X with $x \neq y$,
- *coercive* if
$$\frac{\langle A(x), x \rangle}{\|x\|} \to +\infty \text{ as } \|x\| \to +\infty,$$
- *demicontinuous* if $x_n \to x$ as $n \to +\infty$ implies $A(x_n) \rightharpoonup A(x)$ as $n \to +\infty$.

A formula for $f \Box g$ is presented in our next theorem. For its formulation we shall need some results from the theory of monotone operators. Assume that f is a Gâteaux-differentiable, convex and coercive function on X. Then $\langle f'(x), x \rangle \geq f(x) - f(0)$ for every x in X [28, page 247]. Hence,

$$\frac{\langle f'(x), x \rangle}{\|x\|} \geq \frac{f(x)}{\|x\|} - \frac{f(0)}{\|x\|} \to +\infty \text{ as } \|x\| \to +\infty,$$

which implies that f' is a coercive operator. If f is strictly convex and X is a reflexive Banach space, then $f': X \to X^*$ is strictly monotone, coercive and demicontinuous [28, page 247]. By the Browder-Minty theorem [27, page 557], the inverse operator $f'^{-1}: X^* \to X$ exists and is strictly monotone, coercive as well as demicontinuous.

Our *representation theorem* reads as follows.

Theorem 2. *Let X be a reflexive Banach space. Let f and g be Gâteaux-differentiable strictly convex coercive real valued functions on X. Put $P = (f'^{-1} + g'^{-1})^{-1}$. Then the following assertions hold.*

(a) *$P: X \to X^*$ is a strictly monotone, coercive and demicontinuous operator.*

(b) *$f \Box g$ is given by*

$$f \Box g = (f \circ f'^{-1} + g \circ g'^{-1}) \circ P. \tag{6}$$

(c) *$f \Box g$ is Gâteaux-differentiable with $(f \Box g)' = P$. Moreover,*

$$(f \Box g)(x) = (f \Box g)(0) + \int_0^1 \langle P(\lambda x), x \rangle \, \mathrm{d}\lambda \tag{7}$$

for all $x \in X$.

PROOF. We know that the inverse operators f'^{-1} and g'^{-1} from X^* onto X exist and are strictly monotone and demicontinuous. The sum $f'^{-1} + g'^{-1}$ is also strictly monotone, coercive and demicontinuous and (a) holds.

(b) In order to prove (6), we note that $(f \Box g)(x)$, for a fixed $x \in X$, is equal to the minimum of the function Ψ defined on X as

$$\Psi(y) = f(y) + g(x - y) \text{ for all } y \in X.$$

Since f and g are strictly convex, so is Ψ. Therefore, the minimum of Ψ is uniquely attained at, say $\bar{y}$. Hence, $\Psi'(\bar{y}) = 0$. This is eqivalent to

$$f'(\bar{y}) = g'(\bar{z}) =: x^*,$$

where $\bar{z} = x - \bar{y}$. Consequently,

$$f'^{-1}(x^*) + g'^{-1}(x^*) = \bar{y} + \bar{z} = x$$

and thus $x^* = P(x)$. We arrive at

$$(f \Box g)(x) = f(\bar{y}) + g(\bar{z}),$$

where

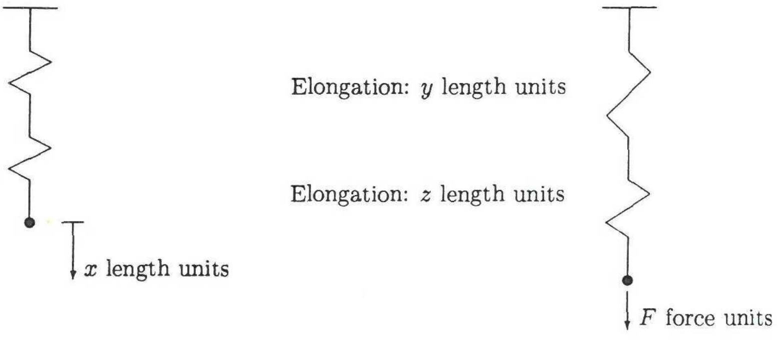
$$\bar{y} = f'^{-1}(P(x)) \text{ and } \bar{z} = g'^{-1}(P(x)).$$

This completes the proof of (b).

(c) It easily verified that $(\partial f // \partial g)(x)$ equals the singleton $\{P(x)\}$ for each $x \in X$. By Theorem 1, $\partial(f \Box g)(x) = \{P(x)\}$ for each $x \in X$ which implies that $f \Box g$ is Gâteaux-differentiable and $(f \Box g)' = P$. Integration yields (7). $\Box$

4 A mechanical example: Two springs connected in series

Let us consider a mechanical system consisting of two springs connected in series. If the system is displaced x length units, then the two springs are elongated y and z length units, respectively, in such a way that $y + z = x$ and the total potential energy of the springs is minimized. Study Figure 1 below.



Unstretched position. *Stretched position.*

FIGURE 1. A mechanical spring system.

Let us denote the potential energy of the first spring due to the displacement y by $U(y)$, and by $V(z)$ the potential energy of the second spring due to the displacement z. The principle of minimal energy states that the potential energy $W(x)$ of the system is given as the solution to the following extremal problem:

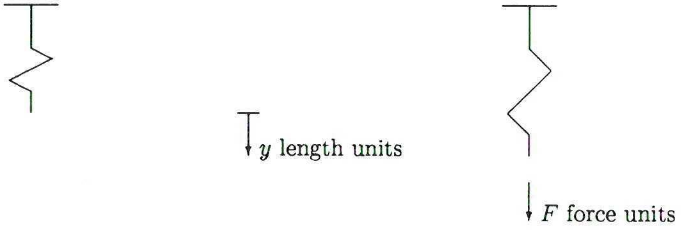
$$\text{Minimize } U(y) + V(z) \text{ subject to } y + z = x.$$

In the terminology of infimal convolution, $W(x) = (U \square V)(x)$, that is we have made the

OBSERVATION. The potential energy function of a system of two springs connected in series is equal to the infimal convolution of the respective potential energy functions of the springs.

Let us briefly discuss the infimal convolution of functions that are reasonable from the mechanical point of view, that is functions that are convex and have certain coerciveness and continuity properties. We introduce the convex cone $\mathcal{E}$ of "admissible potential energy functions" consisting of all convex coercive and lower semicontinuous functions U on $\mathbb{R}$ taking values only in $[0, +\infty]$ and such that $U(0) = 0$. We notice that $\mathcal{E} \subset \mathcal{C}(\mathbb{R})$. It is easily verified that $\mathcal{E}$ is closed under infimal convolution. For a potential energy function

$U \in \mathcal{E}$, the possible elongation of the corresponding spring is any number in the essential domain $\operatorname{dom} U$ and no other number. The set $\operatorname{dom} U$ is convex and contains 0. In the extreme case, $\operatorname{dom} U$ is equal to $\{0\}$ which means that $U = \mathcal{I}$. The spring corresponding to $\mathcal{I}$ admits *no* elongation—it is stiff. If U is a function in $\mathcal{E}$, then the subdifferential $\partial U \colon \mathbb{R} \to \mathcal{P}(\mathbb{R})$ is maximal monotone, $0 \in \partial U(0)$ and $\bigcup_{y \in \mathbb{R}} \partial U(y) = \mathbb{R}$. The relation between the spring force F and the displacement y reads $F \in \partial U(y)$. In particular, if $U(y) = \frac{1}{2}ky^2$, where $k \in (0, +\infty)$, then the mechanical law $F \in \partial U(y)$ takes the form $F = ky$—the case of linear spring theory. See Figure 2 below.



Unstretched position. *Stretched position.*

FIGURE 2. The mechanical law $F \in \partial U(y)$.

Our system in Figure 1 is displaced x length units due to a force of magnitude F force units; thus $F \in \partial(U \square V)(x)$. By an elementary physical argument, we must have $F \in \partial U(y) \cap \partial V(z)$, where y and z $(y + z = x)$ denote the elongations of the first and second spring, respectively. Indeed, the two springs experience the same spring force (of magnitude F) and, therefore, $F \in \partial U(y)$ and $F \in \partial V(z)$. This is in agreement with assertion (a) of our Theorem 1: $\partial(U \square V) = \partial U // \partial V$.

We close this section with an application of Theorem 2.

EXAMPLE. Let U and V be the following two functions in $\mathcal{E}$:

$$U(x) = \frac{|x|^p}{p} \text{ and } V(x) = \frac{|x|^q}{q},$$

where $1 < p < q < +\infty$. It is easily proved that $U \square V$ is even. Denoting $P = (U'^{-1} + V'^{-1})^{-1}$, we have

$$P^{-1}(y) = \left(|y|^{1/(p-1)} + |y|^{1/(q-1)}\right) \operatorname{sgn} y$$

and, by Theorem 2,

$$\begin{aligned}(U \square V)(x) &= U\left(U'^{-1}(P(x))\right) + V\left(V'^{-1}(P(x))\right) \\ &= \frac{1}{p}|P(x)|^{p/(p-1)} + \frac{1}{q}|P(x)|^{q/(q-1)}.\end{aligned}$$

Thus,

$$\lim_{x \to +\infty} \frac{(U \square V)(x)}{x^p} = \lim_{y \to +\infty} \frac{(U \square V)(P^{-1}(y))}{[P^{-1}(y)]^p}$$
$$= \lim_{y \to +\infty} \frac{\frac{1}{p} y^{p/(p-1)} + \frac{1}{q} y^{q/(q-1)}}{[y^{1/(p-1)} + y^{1/(q-1)}]^p} = \frac{1}{p}.$$

Therefore

$$\frac{(U \square V)(x)}{|x|^p/p} \to 1 \text{ as } x \to \pm\infty.$$

In a similar way we get

$$\frac{(U \square V)(x)}{|x|^q/q} \to 1 \text{ as } x \to 0.$$

5 An application to a Hamilton–Jacobi equation

Let $H : \mathbb{R} \to \mathbb{R}$ (the *Hamiltonian*) be convex and coercive. The conjugate function $L = H^*$ (the *Lagrangian*) is then finitely valued, convex and coercive. Let $f : \mathbb{R} \to \mathbb{R}$ be lower semicontinuous and bounded from below. We define a function u on $\mathbb{R} \times [0, +\infty)$ by

$$u(\cdot, t) = \begin{cases} f & \text{when } t = 0, \\ f \square L_t & \text{when } t > 0. \end{cases}$$

Here, $L_t = tL(\cdot/t)$. Note that

$$\operatorname{epi}_s u(\cdot, t) = \operatorname{epi}_s f + t \operatorname{epi}_s L$$

for all $t \in [0, +\infty)$. The function u has the following properties (see Proposition 13.1 of [19] and [5, page 88]):

- $u(x, t) \to f(x)$ as $t \to 0+$ for each real number x.
- u is locally Lipschitzian on $\mathbb{R} \times (0, +\infty)$, hence differentiable almost everywhere.
- At each $(x, t) \in \mathbb{R} \times (0, +\infty)$ where u is differentiable, u satisfies the *Hamilton-Jacobi equation*
$$\dot{u}(x, t) + H(u'(x, t)) = 0. \tag{8}$$
  Here, the prime and the dot denote partial differentiation with respect to the first and second variable, respectively.

Actually, u is the *viscosity solution* of the Hamilton-Jacobi equation (8) with initial condition $u(\cdot, 0) = f$, consult [19].

The following proposition gives properties of and an explicit formula for u in the purely convex case, that is, when also f is convex.

Proposition 2. *Let f be convex. Then*

(a) *u is a convex function and $\partial u(\cdot,t) = ((\partial f)^{-1} + t\partial H)^{-1}$ for each $t \in (0,+\infty)$.*
If in addition f is coercive, then (b) *and* (c) *below hold.*

(b) *$u(\cdot,t)$ is coercive for each $t \in (0,+\infty)$.*

(c) *If f and H are differentiable and strictly convex, then u is a classical solution of* (8) *supplied with the initial condition $u(\cdot,0) = f$, and $u(x,t)$ is given by each of the following three formulae*

$$
\begin{aligned}
u(x,t) &= f(f'^{-1}(\omega(x,t))) + tL(H'(\omega(x,t))), && (9)\\
u(x,t) &= f(x) - \int_0^t H(\omega(x,\tau))\,\mathrm{d}\tau, && (10)\\
u(x,t) &= u(0,t) + \int_0^x \omega(\xi,t)\,\mathrm{d}\xi, && (11)
\end{aligned}
$$

for all $(x,t) \in \mathbb{R} \times (0,+\infty)$. Here, $\omega(x,t) = (f'^{-1} + tH')^{-1}(x)$.

PROOF. First we note that

$$u(x,t) = \min\{f(x-q) + tL(q/t) |\, q \in \mathbb{R}\}$$

for all $(x,t) \in \mathbb{R} \times (0,+\infty)$.

(a) We start by proving that the function $(q,t) \mapsto tL(q/t)$ on $\mathbb{R} \times (0,+\infty)$ is convex. From the definition of conjugate functions

$$tL(q/t) = (tH)^*(q) = \max\{pq - tH(p) |\, p \in \mathbb{R}\}.$$

The fact that $(q,t) \mapsto pq - tH(p)$ is linear for each real number p implies that the function $(q,t) \mapsto tL(q/t)$ is convex since it is a pointwise maximum of linear functions. Hence,

$$(x,t,q) \mapsto f(x-q) + tL(q/t)$$

is a convex function. Then u is convex since it is obtained by taking minimum over one variable in the above formula, see [9, page 50].

The subdifferential of $L_t = (tH)^*$ equals $(t \cdot \partial H)^{-1}$. By Proposition 1, $\partial u(\cdot,t) = ((\partial f)^{-1} + t\partial H)^{-1}$.

Assertion (b) follows directly from assertion (b) of Theorem 1.

(c) According to Proposition 15.1 in [19], u is a classical solution. Let us apply Theorem 2. Since the derivative of the function $L_t = (tH)^*$ equals $(tH')^{-1}$ we get $u' = \omega$

(see point (c) of Theorem 2) from which (11) follows. The formula (9) is a straightforward application of assertion (b) of Theorem 2. Finally,

$$\begin{aligned} u(x,t) &= u(x,0) + \int_0^t \dot{u}(x,\tau)\,\mathrm{d}\tau \\ &= f(x) + \int_0^t (-H(u'(x,\tau)))\,\mathrm{d}\tau \\ &= f(x) - \int_0^t H(\omega(x,\tau))\,\mathrm{d}\tau, \end{aligned}$$

and (10) is proved. □

6 A concluding remark

In [26], an operation $\oplus$ on the set Φ of all continuous and nondecreasing functions $\varphi\colon (0,+\infty) \to (0,+\infty)$ such that $\varphi(t) \to 0$ as $t \to 0+$ and $\varphi(t) \to +\infty$ as $t \to +\infty$, is considered. The operation $\oplus$ is defined by the formula

$$(\varphi \oplus \psi)(s) = \inf\{\varphi(t) + \psi(u) \mid t, u \in (0,+\infty) \text{ and } s = tu\}$$

for any $\varphi, \psi \in \Phi$ and every $s \in (0,+\infty)$. Clearly, given two functions φ and ψ in Φ, the *Young-type inequality*

$$(\varphi \oplus \psi)(tu) \leq \varphi(t) + \psi(u)$$

holds for all $t, u \in (0,+\infty)$. Consult [20] or [21] for possible applications of such inequalities in the theory of Orlicz spaces.

The following relation between $\oplus$ and $\Box$ (with $X = \mathbb{R}$) holds: Given any elements φ and ψ of Φ we can define functions f and g on $\mathbb{R}$ by

$$f(y) = \varphi(e^y) \text{ and } g(z) = \psi(e^z)$$

for all real numbers y and z. Then

$$\begin{aligned} (f \Box g)(x) &= \inf\{f(y) + g(z) \mid y, z \in \mathbb{R} \text{ and } y + z = x\} \\ &= \inf\{\varphi(e^y) + \psi(e^z) \mid y, z \in \mathbb{R} \text{ and } e^y e^z = e^x\} \\ &= (\varphi \oplus \psi)(e^x) \end{aligned}$$

for each real number x. In other words, the assignment $\varphi \mapsto \varphi \circ \exp$ establishes an isomorphism $(\Phi, \oplus) \cong (\mathcal{M}, \Box)$. Here, $\mathcal{M}$ denotes the set of all continuous nondecreasing functions $f : \mathbb{R} \to (0,+\infty)$ satisfying

$$\lim_{x \to -\infty} f(x) = 0 \text{ and } \lim_{x \to +\infty} f(x) = +\infty.$$

We conclude this paper by deducing from Theorem 1 in [26] the following result for the infimal convolution of monotone functions.

Proposition 3. *$\mathcal{M}$ is closed under infimal convolution and*

$$(f//g)(x) < (f\Box g)(x) \leq 2(f//g)(x) \tag{12}$$

for any f and g in $\mathcal{M}$ and all $x \in \mathbb{R}$.

For each $x \in \mathbb{R}$, let $y = \tilde{y}(x)$ be a solution to the equation $f(y) = g(x-y)$. By definition, the parallel sum $f//g$ assigns to x the positive real number $f(\tilde{y}(x)) = g(x - \tilde{y}(x))$. Note that $(f//g)(x)$ is well defined though the equation $f(y) = g(x-y)$ may have more than one root y. If f and g are strictly increasing, then $f//g = (f^{-1} + g^{-1})^{-1}$.

PROOF. We prove only the inequality (12). Choose $x \in \mathbb{R}$ arbitrarily. Clearly,

$$\max\{f(y), g(x-y)\} \geq (f//g)(x)$$

for every $y \in \mathbb{R}$. It is not difficult to verify that

$$(f\Box g)(x) = f(\bar{y}(x)) + g(x - \bar{y}(x))$$

for some $\bar{y}(x) \in \mathbb{R}$. Hence,

$$(f\Box g)(x) = f(\bar{y}(x)) + g(x-\bar{y}(x)) > \max\{f(\bar{y}(x)), g(x-\bar{y}(x))\} \geq (f//g)(x),$$

which is the left-hand side inequality in (12). In order to prove the right-hand side inequality in (12) take $\tilde{y}(x) \in \mathbb{R}$ such that $f(\tilde{y}(x)) = g(x - \tilde{y}(x))$. Then

$$(f\Box g)(x) \leq f(\tilde{y}(x)) + g(x - \tilde{y}(x)) = (f//g)(x) + (f//g)(x),$$

and the proof of inequality (12) is complete. □

References

[1] ANDERSON, W.N. JR. AND DUFFIN, R.J. *Series and Parallel Addition of Matrices*, J. Math. Anal. Appl. 26, 576-594 (1969).

[2] ATTOUCH, H. *Variational Convergence for Functions and Operators*, Applicable Math. Series, Pitman, London 1984.

[3] ATTOUCH, H. AND AZÉ, D. *Approximation and Regularization of Arbitrary Functions in Hilbert Spaces by the Lasry-Lions Method*, Annales de l'Institut H. Poincaré – Analyse non linéaire (to appear).

[4] ATTOUCH, H. AND WETS, R. *Epigraphical Analysis*, in "Analyse non Linéaire", eds. H. Attouch, J.-P Aubin, F.H. Clarke, I. Ekeland, 73-99, Gauthiers-Villars Paris et C.R.M. Montreal 1989.

[5] BENTON, S.H. *The Hamilton–Jacobi Equation: A Global Approach*, Academic Press, New York-San Francisco-London 1977.

[6] BOUGEARD, M., PENOT, J.-P. AND POMMELLET, A. *Towards Minimal Assumptions for the Infimal Convolution Regularization*, J. Approx. Theory 64, 245-270 (1991).

[7] BOMAN, J. *The Sum of Two Plane Convex C^∞ Sets is not Always C^5*, Math. Scand. 66, 216-224 (1990).

[8] BOMAN, J. *Smoothness of Sums of Convex Sets with Real Analytic Boundaries*, Math. Scand. 66, 225-230 (1990).

[9] EKELAND, I. AND TEMAM, R. *Convex Analysis and Variational Problems*, North-Holland, Amsterdam-Oxford 1976.

[10] ERIKSSON, S.-L AND LEUTWILER, H. *A Potential-Theoretic Approach to Parallel Addition*, Math. Ann. 274, 301-317 (1986).

[11] FENCHEL, W. *Convex Cones, Sets, and Functions*, Lecture Notes, Princeton University, Princeton, New Jersey 1953.

[12] FOUGERES, A. AND TRUFFERT, A. *Régularisation S.C.I. et Γ-convergence: approximations inf-convolutives associées à un référentiel*, Annali di Matematica Puri ed Applicati, 21-51 (1986).

[13] HIRIART-URRUTY, J.-B. *Extension of Lipschitz Functions*, J. Math. Anal. Appl. 77, 539-544 (1980).

[14] HIRIART-URRUTY, J.-B. *Lipschitz r-Continuity of the Approximative Subdifferential of a Convex Function*, Math. Scand. 47, 123-134 (1980).

[15] HIRIART-URRUTY, J.-B. *How to Regularize a Difference of Convex Functions*, J. Math. Anal. Appl. 162, 196-209 (1991).

[16] IOFFE, A.D. AND TIHOMIROV, V.M. *Theory of Extremal Problems*, North-Holland, Amsterdam-New York-Oxford 1979.

[17] KISELMAN, C.O. *Smoothness of Vector Sums of Plane Convex Sets*, Math. Scand. 60, 239-252 (1987).

[18] LAURENT, P.-J. *Approximation et optimisation*, Hermann, Paris 1972.

[19] LIONS, P.L. *Generalized Solutions of Hamilton–Jacobi Equations*, Pitman, London 1982.

[20] MALIGRANDA, L. *Orlicz Spaces and Interpolation*, Seminars in Math. 5, Campinas 1989.

[21] MALIGRANDA, L. AND PERSSON, L.E. *Generalized Duality of some Banach Function Spaces*, Indagationes Math. 51, 323-338 (1989).

[22] MAZURE, M.-L. *Analyse variationelle des formes quadratiques convexes*, Thèse de doctorat, Université Paul Sabatier, Toulouse, 1986.

[23] MOREAU, J.J. *Fonctionelles Convexes*, Séminaire sur les équations aux dérivées partielles II, Collège de France 1966-1967.

[24] MORLEY, T.D. *Parallel Summation, Maxwell's Principle and the Infimum of Projections*, J. Math. Anal. Appl. 70, 33-41 (1979).

[25] ROCKAFELLAR, R.T. *Convex Analysis*, Princeton University Press, Princeton, New York 1970.

[26] STRÖMBERG, T. *An Operation Connected to a Young-Type Inequality*, Mathematische Nachrichten (to appear).

[27] ZEIDLER, E. *Nonlinear Functional Analysis and its Applications II*, Springer-Verlag, New York-Berlin-Heidelberg-London-Paris-Tokyo 1990.

[28] ZEIDLER, E. *Nonlinear Functional Analysis and its Applications III*, Springer-Verlag, New York-Berlin-Heidelberg-London-Paris-Tokyo 1985.

D

D

On the Epigraphical Sum of Functions Uniformly Continuous on Bounded Sets *

Thomas Strömberg

Department of Mathematics, Luleå University of Technology, S-951 87 Luleå, Sweden.

Abstract

In this paper we study the properties of the epigraphical sum $f +_e g$ for the case where g is uniformly continuous on bounded sets, with minimal assumptions on the function f. In particular, our results tell to what extent the regularity of g is carried over to $f +_e g$. This is done with the aid of a modulus of uniform continuity. Moreover, we consider approximation procedures of Moreau–Yosida type, with kernel functions uniformly continuous on bounded sets, in a fairly general setting.

1 Preliminaries

The epigraphical analysis is based on the idea to study the effect of set operations and set limits on the epigraphs of functions rather than to consider the corresponding graphs. Recall that with each extended real-valued function $f: X \to \bar{\mathbb{R}}$ — throughout this paper X is taken to be a normed real vector space — we may associate its *epigraph*

$$\operatorname{epi} f = \{(x, \alpha) \in X \times \mathbb{R} |\ f(x) \leq \alpha\}$$

and its *strict epigraph*

$$\operatorname{epi_s} f = \{(x, \alpha) \in X \times \mathbb{R} |\ f(x) < \alpha\}.$$

For instance, Kuratowski–Painlevé set convergence applied to epigraphs gives birth to the concept of epi-convergence, consult [1]. Moreover, the theory of differentiability of convex lower semicontinuous functions may be seen as a theory of supporting hyperplanes to the corresponding epigraphs. Of course, in terms of epigraphs, f lower semicontinuous is equivalent to $\operatorname{epi} f$ closed, whereas f is convex if and only if $\operatorname{epi} f$ (or $\operatorname{epi_s} f$) is a convex set. The epigraphical approach has proved fruitful for example in the study of extremal

*Example 4 in the original version contained an obvious error and has been replaced by a new example.

problems and variational theory. An overview of the basic concepts of epigraphical analysis can be found in the paper [4] by Attouch and Wets. In this paper we limit ourselves to the Minkowski sum of strict epigraphs: Given any two extended real-valued functions f and g defined on X, the addition

$$\mathrm{epi_s}\, f + \mathrm{epi_s}\, g =: \mathrm{epi_s}\,(f +_{\mathrm{e}} g),$$

gives rise to the strict epigraph of a function that we denote by $f +_{\mathrm{e}} g$ and call the *epigraphical sum* or the *infimal convolution* of f and g. The following functional identity justifies the terminology "infimal convolution" (see [13] or [15]): the function $f +_{\mathrm{e}} g$ assigns to each $x \in X$ the extended real number

$$(f +_{\mathrm{e}} g)(x) = \inf\{f(y) + g(z) \mid y, z \in X \text{ and } y + z = x\}.$$

In the above formula, if necessary, the sums $(+\infty) + (-\infty)$ and $(-\infty) + (+\infty)$ are understood to equal $+\infty$.

Epigraphical addition is a commutative and associative binary operation on the set $\bar{\mathbb{R}}^X$ of all extended real-valued functions on X, see [15], with the indicator function $\delta_{\{0\}}$ of $\{0\}$, defined as

$$\delta_{\{0\}}(x) = \begin{cases} 0 & \text{when } x = 0, \\ +\infty & \text{when } x \neq 0, \end{cases}$$

acting as the identity element. Therefore, algebraically, $(\bar{\mathbb{R}}^X, +_{\mathrm{e}})$ is a commutative monoid.

In some cases epigraphical addition has a regularizing effect. For instance, if g satisfies a Lipschitz condition and if f is proper, then $f +_{\mathrm{e}} g$ is either identically $-\infty$ or satisfies a Lipschitz condition with the same constant as does g [11]. This fact may be used if we are interested in approximating a given proper and lower semicontinuous function f with Lipschitz continuous functions. Let us assume the existence of a real number $c \geq 0$ such that $f + c\|\cdot\|$ is bounded from below. Then, if $0 < t < 1/c$, $f +_{\mathrm{e}} (1/t)\|\cdot\|$ is finitely valued and Lipschitz continuous with constant $1/t$. Moreover, $f +_{\mathrm{e}} (1/t)\|\cdot\|$ increases pointwise to f as t decreases to zero. This method was first used by F. Hausdorff [10] who attributes the definition to M. Pasch, see [9] for historical comments. However, the most famous approximation technique by means of epigraphical addition corresponds to the choices $\|\cdot\|^2/(2t)$, $t > 0$, of the kernel functions. The Moreau–Yosida approximates $f_t := f +_{\mathrm{e}} \|\cdot\|^2/(2t)$, $t > 0$, enjoy very peculiar properties especially in the case of proper convex and lower semicontinuous f defined on a Hilbert space — for instance, f_t is finitely valued, convex and Fréchet-differentiable with a Lipschitz continuous derivative. Furthermore, $f_t(x) \nearrow f(x)$ as $t \searrow 0$ for each $x \in X$. Actually, the Moreau–Yosida approximation of convex and lower semicontinuous f defined on a Hilbert space is nowadays well understood and classical. We mention that these results have recently been extended to functions defined on a Hilbert space and only assumed to be quadratically minorized, see [3], [7] and [12]. This is accomplished by first performing an inf-convolution and then a sup-convolution: If f is any proper function on a Hilbert space X with $f + \frac{1}{2}c\|\cdot\|^2$ bounded from below for some $c \geq 0$, then, whenever $0 < s < t < 1/c$,

$$x \mapsto (f_t)^s(x) = \sup_{y \in X} \inf_{u \in X} \left(f(u) + \frac{1}{2t}\|u - y\|^2 - \frac{1}{2s}\|y - x\|^2 \right)$$

is a real-valued differentiable function whose gradient is Lipschitz continuous with constant $\max\{1/s, 1/(t-s)\}$. Moreover, $(f_t)^s$ approaches $\operatorname{cl} f$ — the lower semicontinuous regularization of f — from below as s and t go to zero, $0 < s < t \to 0$. For more on the "Lasry–Lions method", as well as for a survey on Moreau–Yosida approximation, see the recent paper [3] by Attouch and Azé.

In this paper we focus on the properties of the epigraphical sum $f +_e g$ assuming g to be uniformly continuous, at least on each bounded subset of X. We prove that the function $f +_e g$ inherits — at least partly — the properties of g provided that $f(y) + g(x-y)$ tends to $+\infty$ as $\|y\| \to +\infty$ uniformly in x over bounded sets. In the case of g quadratic, the latter coerciveness condition is guaranteed when f is properly quadratically minorized. We also prove results for a Moreau–Yosida type of approximation using kernels assumed to be uniformly continuous on bounded sets.

This paper is organized in the following way: In section 2 we present a minor generalization of some of the results in [11] by Hiriart-Urruty; namely, instead of Lipschitz continuity we consider uniform continuity and introduce the concept of m-uniform continuity; m being a "modulus of uniform continuity". In this setting, Lipschitz continuity corresponds to $m(r) = cr$. This approach enables us treat the case $f +_e g$ where g is uniformly continuous in an easy way. The more general situation of functions uniformly continuous on bounded sets is studied in sections 3 and 4, where section 4 contains results for generalizations of the Moreau–Yosida approximation process.

Throughout, we denote the closed ball with center at the point x and of radius r by $B_r(x)$ and write simply $B_r = B_r(0)$. By $\mathbb{N}$ we understand the set of natural numbers, $\mathbb{R}$ denotes the set of real numbers, and the set of positive real numbers is written $\mathbb{P}$, while $\mathbb{R}_+ = \mathbb{P} \cup \{0\}$, $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$.

ACKNOWLEDGEMENTS. This paper was completed during a visit at Université de Montpellier II. The author would like to thank Laboratoire d'Analyse Convexe, and especially Professor Hedy Attouch, for their hospitality. Moreover, the author wishes to express his gratitude to Professor Attouch, Professor Azé (Perpignan) and Professor Volle (Avignon) for valuable comments and suggestions which have improved the final version of this paper.

The research reported here was partly supported by Magnusons fond of The Royal Swedish Academy of Sciences.

2 Uniformly continuous kernel functions

By $\mathcal{M}$ we denote the set of all nondecreasing and subadditive functions $m: \mathbb{R}_+ \to \mathbb{R}_+$ that are continuous at 0 with $m(0) = 0$. Every function belonging to $\mathcal{M}$ is continuous on

$\mathbb{R}_+$ and is either identically equal to 0 or satisfies $m(r) > 0$ for all $r > 0$. Any concave function $m\colon \mathbb{R}_+ \to \mathbb{R}_+$ which is continuous at 0 and such that $m(0) = 0$, belongs to $\mathcal{M}$. Given a nonempty subset E of X and $m \in \mathcal{M}$, we say that a function $f\colon E \to \mathbb{R}$ is *m-uniformly continuous on E*, in symbols $f \in \mathrm{UC}(E; m)$, iff the following inequality holds for all $x, y \in E$:

$$|f(x) - f(y)| \leq m(\|x - y\|).$$

To say that f is m-uniformly continuous on E, with m not identically equal to 0, is the same as to say that f is Lipschitz continuous on E (with constant 1) with respect to the metric $(x, y) \mapsto m(\|x - y\|)$. In particular, "$m$-uniform continuity" means "Hölder continuity with exponent p and constant 1" — where $0 < p \leq 1$ — in the case where m equals the power function $r \mapsto r^p$.

Obviously, if f is m-uniformly continuous on E for some $m \in \mathcal{M}$, then f is uniformly continuous on E. It turns out that the converse holds true, provided E is a *convex* set.

Lemma 1. *Let E be a nonempty convex subset of X and let $f\colon E \to \mathbb{R}$ be uniformly continuous. Then the assignments $m(0) = 0$ and*

$$\forall r \in \mathbb{P} \quad m(r) = \sup\{|f(x) - f(y)| \mid x, y \in E \text{ and } \|x - y\| \leq r\}$$

together define a m in $\mathcal{M}$ with the property that $f \in \mathrm{UC}(E; m)$.

PROOF. Obviously, m is nondecreasing and continuous at 0. Take any $r, s \in \mathbb{P}$ and $x, z \in E$ with $\|x - z\| \leq r + s$. There exists y on the segment $[x, z]$ such that $\|x - y\| \leq r$ and $\|y - z\| \leq s$. Since E is convex, y belongs to E. Hence,

$$|f(x) - f(z)| \leq |f(x) - f(y)| + |f(y) - f(z)| \leq m(r) + m(s).$$

Thus, $m(r+s) \leq m(r) + m(s)$, and we observe that m cannot take the value $+\infty$; whence $m \in \mathcal{M}$. Obviously,

$$|f(x) - f(y)| \leq m(\|x - y\|)$$

for all elements x and y of E, as asserted. □

EXAMPLE 1. Let $m \in \mathcal{M}$. Then the composite $m \circ \|\cdot\|$ belongs to $\mathrm{UC}(X; m)$. In other words, we claim that

$$|m(\|x\|) - m(\|y\|)| \leq m(\|x - y\|)$$

for all $x, y \in X$. This is in fact just another way of expressing the subadditivity of the function $m \circ \|\cdot\|$.

In terms of epigraphical addition, m-uniform continuity may be expressed in a convenient way. In this section, we make use of the

NOTATIONS. (i) For $f: E \to \mathbb{R}$, where E is a nonempty subset of X, $\tilde{f}$ denotes the extension of f to X defined by setting $\tilde{f}(x) = +\infty$ for all $x \in X \setminus E$.

(ii) For any $m: \mathbb{R}_+ \to \mathbb{R}_+$ we put $\hat{m} = m \circ \|\cdot\|$.

Lemma 2. *Let E be a nonempty subset of X and assume that m belongs to $\mathcal{M}$. Then, for any $f: E \to \mathbb{R}$, f is a member of $\mathrm{UC}(E; m)$ if and only if*

$$\tilde{f} +_{\mathrm{e}} \hat{m} = f \text{ on } E.$$

REMARK. More explicitly,

$$\begin{aligned} (\tilde{f} +_{\mathrm{e}} \hat{m})(x) &= \inf\{\tilde{f}(y) + \hat{m}(x-y) \mid y \in X\} \\ &= \inf\{f(y) + \hat{m}(x-y) \mid y \in E\} \\ &= \inf\{f(y) + m(\|x-y\|) \mid y \in E\} \end{aligned}$$

for all $x \in X$.

PROOF. We observe that $(\tilde{f} +_{\mathrm{e}} \hat{m})(x) \le f(x) + \hat{m}(0) = f(x)$ for every $x \in E$. Therefore, the following statements are equivalent:

- f belongs to $\mathrm{UC}(E; m)$.
- $f(x) - f(y) \le \hat{m}(x-y)$ for all $x, y \in E$.
- $f(x) \le \inf\{f(y) + \hat{m}(x-y) \mid y \in E\} = (\tilde{f} +_{\mathrm{e}} \hat{m})(x)$ for all $x \in E$.
- $f(x) = (\tilde{f} +_{\mathrm{e}} \hat{m})(x)$ for all $x \in E$.

The proof is complete. □

EXAMPLE 2. From Example 1 and Lemma 2 it follows that $\hat{m} +_{\mathrm{e}} \hat{m} = \hat{m}$ for each $m \in \mathcal{M}$. Actually, a proper function φ is subaddive if and only if $\varphi +_{\mathrm{e}} \varphi \ge \varphi$.

Our first result reads

Proposition 1. *Assume that $g: X \to \mathbb{R}$ is uniformly continuous, that is, $g \in \mathrm{UC}(X; m)$ for some $m \in \mathcal{M}$. Let f be a proper function on X. Then $f +_{\mathrm{e}} g$ is either identically $-\infty$ or finitely valued and in $\mathrm{UC}(X; m)$.*

PROOF. An application of Lemma 2, remembering that $+_{\mathrm{e}}$ is associative, yields

$$(f +_{\mathrm{e}} g) +_{\mathrm{e}} \hat{m} = f +_{\mathrm{e}} (g +_{\mathrm{e}} \hat{m}) = f +_{\mathrm{e}} g. \tag{1}$$

Now, assume that $f +_{\mathrm{e}} g$ takes the value $-\infty$. Then $(f +_{\mathrm{e}} g) +_{\mathrm{e}} \hat{m}$ is identically equal to $-\infty$, and we deduce from (1) that also $f +_{\mathrm{e}} g$ is identically $-\infty$. The only other alternative is $f +_e g$ real-valued. In that case, according to equation (1) and Lemma 2, we must have $f +_e g \in \mathrm{UC}(X; m)$. $\square$

REMARK. Under the assumptions of Proposition 1, a necessary condition for $f +_e g$ to be finitely valued is that f is bounded from below on each bounded set. Indeed, if there exists a bounded set $\mathcal{B}$ on which f is not bounded from below, then, for every $x \in X$,

$$(f +_e g)(x) \leq \inf\{f(y) + g(x - y) \mid y \in \mathcal{B}\} = -\infty,$$

since $y \mapsto g(x - y)$ is bounded on $\mathcal{B}$.

Proposition 1 does *not* remain true if "uniformly continuous" is replaced by "continuous" in the hypothesis:

EXAMPLE 3. Let $(X, (\cdot, \cdot))$ be a nontrivial scalar product space and let $f = \|\cdot\|^2$. Then

$$\begin{aligned}
[f +_e (-f)](x) &= \inf\{\|y\|^2 - \|x - y\|^2 \mid y \in X\} \\
&= \inf\{-\|x\|^2 + 2(x, y) \mid y \in X\} \\
&= \begin{cases} 0 & \text{when } x = 0, \\ -\infty & \text{when } x \neq 0. \end{cases}
\end{aligned}$$

EXAMPLE 4. Let us consider the vector space $\mathrm{BUC}(X)$ of all bounded uniformly continuous real-valued functions defined on X endowed with sup-norm. We have that $\mathrm{BUC}(X)$ is closed (as a set) under epigraphical addition, see Proposition 1. Moreover, it is readily verified that $+_e$ acts continuously in $\mathrm{BUC}(X)$.

The main result of this section is the following extension result.

Theorem 1. *Let E be a nonempty subset of X. Assume that $f \in \mathrm{UC}(E; m)$ for some $m \in \mathcal{M}$. Then, $\tilde{f} +_e \hat{m}$ is a finitely valued extension of f to the whole of X and $\tilde{f} +_e \hat{m} \in \mathrm{UC}(X; m)$.*

PROOF. Since $f \in \mathrm{UC}(E; m)$ we have $\tilde{f} +_e \hat{m} = f$ on E, see Lemma 2. By Proposition 1, $\tilde{f} +_e \hat{m}$ is finitely valued (it cannot be identically $-\infty$ since it is finite on E) and $\tilde{f} +_e \hat{m} \in \mathrm{UC}(X; m)$. We have used the fact that $\hat{m} \in \mathrm{UC}(X; m)$, see Example 1. $\square$

In particular, if f is uniformly continuous on a convex set E, then, according to Lemma 1, f belongs to $\mathrm{UC}(E; m)$ for some $m \in \mathcal{M}$ and $F := \tilde{f} +_e \hat{m}$ is a uniformly continuous extension of f to X: $F|_E = f$ and $F \in \mathrm{UC}(X; m)$.

3 Kernel functions that are uniformly continuous on each bounded set

In this section we study the regularization $f \mapsto f +_{\mathrm{e}} g$ for the case where the kernel function g is continuous but not necessarily uniformly continuous on the whole of X. Instead we assume that g is uniformly continuous on each bounded subset of X. Examples of such functions g are $g = \|\cdot\|^p$, where $p > 1$, and any continuous g if X is a finite-dimensional Banach space. In other words, we assume that, for every $\rho \in \mathbb{P}$, there exists $m_\rho \in \mathcal{M}$ such that g is m_ρ-uniformly continuous on the ball B_ρ — recall Lemma 1.

Proposition 2. *Let $g\colon X \to \mathbb{R}$ be uniformly continuous on each bounded subset of X. For each $\rho \in \mathbb{P}$, let $m_\rho(0) = 0$ and*

$$\forall r \in \mathbb{P} \quad m_\rho(r) = \sup\{|g(x) - g(y)| \mid x, y \in B_\rho \text{ and } \|x - y\| \leq r\}.$$

These assignments define a family $(m_\rho)_{\rho \in \mathbb{P}}$ of functions in $\mathcal{M}$ such that g is m_ρ-uniformly continuous on the ball B_ρ for every ρ in $\mathbb{P}$. Let f be a proper function on X with

$$f(y) + g(x - y) \to +\infty \text{ as } \|y\| \to +\infty \tag{2}$$

uniformly in x over each bounded set.

(a) *If there exists a bounded set on which f is not bounded from below, then $f +_{\mathrm{e}} g$ is identically equal to $-\infty$.*

(b) *Provided f is bounded from below on each bounded set, $f +_{\mathrm{e}} g$ is finitely valued and there exists a nondecreasing function $\rho \mapsto \delta(\rho)\colon \mathbb{P} \to \mathbb{R}_+$ such that, for each $\rho \in \mathbb{P}$, $f +_{\mathrm{e}} g$ is $m_{\rho+\delta(\rho)}$-uniformly continuous on the ball B_ρ.*

(c) *If g is Lipschitz continuous on bounded sets, and f is bounded from below on bounded sets, then $f +_{\mathrm{e}} g$ is Lipschitz continuous on bounded sets.*

(d) *Let $(f^t)_{t\in(0,T]}$ be a family of proper functions on X bounded from below on bounded sets and such that*

- *$t \mapsto f^t(y)$ is nonincreasing for each $y \in X$;*
- *There exist $\bar{y} \in X$ and $C < +\infty$ such that $f^t(\bar{y}) \leq C$ for all $t \in (0, T]$;*
- *The uniform convergence (2) holds with f replaced by f^T.*

Then $(f^t +_{\mathrm{e}} g)_{t\in(0,T]}$ is equi-uniformly continuous over each bounded subset of X: There exists a nondecreasing $\rho \mapsto \delta(\rho)\colon \mathbb{P} \to \mathbb{R}_+$ such that, for each $\rho \in \mathbb{P}$ and $t \in (0, T]$, $f^t +_{\mathrm{e}} g$ is $m_{\rho+\delta(\rho)}$-uniformly continuous on the ball B_ρ.

We remark that f is bounded from below on bounded sets for instance when X is a finite-dimensional Banach space and f is lower semicontinuous.

PROOF. (a) See the remark following the proof of Proposition 1. (b) Since f is proper there exists $\bar{y} \in X$ with $f(\bar{y})$ finite. Obviously,

$$(f +_{\mathrm{e}} g)(x) \leq f(\bar{y}) + g(x - \bar{y}) \tag{3}$$

for every $x \in X$. Take an arbitrary $\rho \in \mathbb{P}$. Since g is bounded on bounded sets, we find, with (3) in mind, that $f +_{\mathrm{e}} g$ is bounded from above on bounded sets. Since (2) holds, there exists $\delta(\rho) \in \mathbb{R}_+$ with

$$f(y) + g(x - y) \geq \sup_{B_\rho}(f +_{\mathrm{e}} g) + 1$$

for all $x \in B_\rho$ and $y \in X \setminus B_{\delta(\rho)}$. Therefore, for every $x \in B_\rho$,

$$(f +_{\mathrm{e}} g)(x) = \inf\{f(y) + g(x - y) \mid y \in B_{\delta(\rho)}\}. \tag{4}$$

Take any $x_1, x_2 \in B_\rho$ and an arbitrary $\epsilon \in \mathbb{P}$. There exists $y_2 \in B_{\delta(\rho)}$ with

$$f(y_2) + g(x_2 - y_2) < (f +_{\mathrm{e}} g)(x_2) + \epsilon.$$

Hence,

$$\begin{aligned}
&(f +_{\mathrm{e}} g)(x_1) - (f +_{\mathrm{e}} g)(x_2) \\
&< [f(y_2) + g(x_1 - y_2)] - [f(y_2) + g(x_2 - y_2) - \epsilon] \\
&= g(x_1 - y_2) - g(x_2 - y_2) + \epsilon \leq m_{\rho+\delta(\rho)}(\|x_1 - x_2\|) + \epsilon,
\end{aligned}$$

where we have used the fact that $x_1 - y_2$ and $x_2 - y_2$ are points in $B_{\rho+\delta(\rho)}$. We deduce that

$$|(f +_{\mathrm{e}} g)(x_1) - (f +_{\mathrm{e}} g)(x_2)| \leq m_{\rho+\delta(\rho)}(\|x_1 - x_2\|)$$

for all x_1, x_2 in B_ρ. Obviously, the function $\rho \mapsto \delta(\rho)$ may be assumed to be nondecreasing.

(c) In this case, $m_\rho(r) \leq c_\rho r$ and from part (b) we conclude that $f +_{\mathrm{e}} g$ is Lipschitz continuous on B_ρ with constant $c_{\rho+\delta(\rho)}$.

(d) Choose an arbitrary $\rho \in \mathbb{P}$. In this case we have

$$(f^t +_{\mathrm{e}} g)(x) \leq f^t(\bar{y}) + g(x - \bar{y}) \leq C + g(x - \bar{y}) \leq C'$$

for some $C' < +\infty$, all $t \in \mathbb{P}$ and all $x \in B_\rho$. There exists $\delta(\rho) \in \mathbb{R}_+$ such that

$$f^t(y) + g(x - y) \geq f^T(y) + g(x - y) \geq C' + 1$$

for all $x \in B_\rho$, $y \in X \setminus B_{\delta(\rho)}$ and $t \in (0, T]$. Consequently,

$$(f^t +_{\mathrm{e}} g)(x) = \inf\{f^t(y) + g(x - y) \mid y \in B_{\delta(\rho)}\}$$

for all $x \in B_\rho$ and all $t \in (0, T]$. Let us use the same type of arguments as in the proof of part (b): Choose x_1 and x_2 in B_ρ and $\epsilon \in \mathbb{P}$. For each $t \in (0, T]$ there exists $y_2^t \in B_{\delta(\rho)}$ with

$$f^t(y_2^t) + g(x_2 - y_2^t) < (f^t +_{\mathrm{e}} g)(x_2) + \epsilon.$$

Hence,

$$\begin{aligned}
&(f^t +_e g)(x_1) - (f^t +_e g)(x_2) \\
&\quad < \; [f^t(y_2^t) + g(x_1 - y_2^t)] - [f^t(y_2^t) + g(x_2 - y_2^t) - \epsilon] \\
&\quad = \; g(x_1 - y_2^t) - g(x_2 - y_2^t) + \epsilon \leq m_{\rho+\delta(\rho)}(\|x_1 - x_2\|) + \epsilon.
\end{aligned}$$

We find that

$$|(f^t +_e g)(x_1) - (f^t +_e g)(x_2)| \leq m_{\rho+\delta(\rho)}(\|x_1 - x_2\|)$$

for all x_1, x_2 in B_ρ and all $t \in (0, T]$. The function $\rho \mapsto \delta(\rho)$ may be assumed to be nondecreasing. □

4 Approximation of Moreau–Yosida type

As mentioned in the preliminaries, the Moreau-Yosida approximation of a function f is based upon the performance of epigraphical addition of f and the quadratic kernel function $\|\cdot\|^2/(2t)$, for real $t > 0$. The Moreau–Yosida regularization and approximation process may be naturally generalized by replacing the quadratic kernel function by another smoothing kernel k, more suitable for the problem at hand. This idea has occured to several authors, see for example [7] and [17]. For instance, we may deal with situations where f is a nonconvex function not minorized by any power of the norm, but still be able to approximate f by, say locally Lipschitz continuous functions.

Throughout this section we make the following

HYPOTHESIS. We assume that $k: X \times \mathbb{P} \to \mathbb{R}_+$ is a function with the following three properties:

($\mathcal{K}_1$) $k(0, t) = 0$ for every real $t > 0$;

($\mathcal{K}_2$) $k(z, \cdot)$ is nonincreasing for every $z \in X$;

($\mathcal{K}_3$) As $t \to 0$, $k(\cdot, t) \to +\infty$ uniformly on bounded subsets of $X \setminus B_r$ for any real $r > 0$.

Note that under these assumptions, as $t \searrow 0$, $k(\cdot, t)$ converges increasingly to the identity $\delta_{\{0\}}$ of epigraphical addition, that is

$$k(z, t) \nearrow \begin{cases} 0 & \text{when } z = 0, \\ +\infty & \text{when } z \neq 0. \end{cases}$$

EXAMPLE 4. An example of a function k with the preceeding properties $\mathcal{K}_1$, $\mathcal{K}_2$ and $\mathcal{K}_3$ is obtained by putting $k(\cdot, t) = (1/t)\Phi \circ \|\cdot\|$, where $\Phi: \mathbb{R}_+ \to \mathbb{R}_+$ is continuous, nondecreasing, $\Phi(0) = 0$ and $\Phi(r) > 0$ for all $r > 0$. The special case $\Phi(r) = r$ gives rise to the approximation of lower semicontinuous functions by Lipschitz continuous functions, whereas the case $\Phi(r) = \frac{1}{2}r^2$ leads to the Moreau–Yosida regularization.

We have the following result concerning the convergence and the regularity of $f_t := f +_e k(\cdot, t)$.

Theorem 2. *Let $k: X \times \mathbb{P} \to \mathbb{R}_+$ satisfy $\mathcal{K}_1$, $\mathcal{K}_2$ and $\mathcal{K}_3$ and assume in addition that, for some positive T,*

- *f is a proper function on X which is bounded from below on each bounded set;*
- *For each $t \in (0, T]$, $k(\cdot, t)$ is uniformly continuous [resp., Lipschitz continuous] on each bounded set;*
- *Uniformly in x as long as x remains in any bounded set,*

$$f(y) + k(x - y, T) \to +\infty \text{ as } \|y\| \to +\infty. \tag{5}$$

Then

(a) Provided $t \in (0, T]$, f_t is a finitely valued function which is uniformly continuous [resp., Lipschitz continuous] on each bounded subset of X. Moreover,

$$f_t(x) \nearrow (\mathrm{cl}\, f)(x) \text{ as } t \searrow 0$$

for every $x \in X$.

(b) Assume that f is uniformly continuous on the ball $B_r(x_0)$ and let $\rho \in (0, r)$. Then, f_t converges uniformly to f on the ball $B_\rho(x_0)$.

PROOF. (a) Take any $x \in X$. The assumption $\mathcal{K}_2$ implies that $t \mapsto f_t(x)$ is nonincreasing and from $\mathcal{K}_1$ it follows that

$$f_t(x) \le f(x) + k(0, t) = f(x) + 0.$$

Hence, firstly,

$$\lim_{t \to 0} f_t(x) \text{ exists and equals } \sup_{t \in \mathbb{P}} f_t(x),$$

and, secondly,

$$\sup_{t \in \mathbb{P}} f_t(x) \le f(x). \tag{6}$$

An application of Proposition 2 gives us that, whenever $0 < t \le T$, f_t is uniformly [resp., Lipschitz] continuous on bounded sets. Then $\sup_{t \in \mathbb{P}} f_t(\cdot)$ is lower semicontinuous and, consequently, (6) can be sharpened to

$$\sup_{t \in \mathbb{P}} f_t(x) \le (\mathrm{cl}\, f)(x). \tag{7}$$

We aim to prove the reverse inequality to (7):

$$\sup_{t\in\mathbb{P}} f_t(x) \geq (\mathrm{cl}\, f)(x),$$

which is clear if $S := \sup_{t\in\mathbb{P}} f_t(x) = +\infty$, so let us assume $S < +\infty$. Because of (5) there exists $r \in \mathbb{P}$ such that

$$f(y) + k(x-y,t) \geq f(y) + k(x-y,T) \geq S + 1$$

for all $t \in (0,T]$ and all $y \in X \setminus B_r$. This implies that

$$f_t(x) = \inf\{f(y) + k(x-y,t) |\, y \in B_r\}$$

for all $t \in (0,T]$. To each $t \in (0,T]$ there corresponds y_t in B_r such that

$$f(y_t) + k(x-y_t,t) \leq f_t(x) + t \leq S + t. \tag{8}$$

Since f is bounded from below on B_r, $t \mapsto k(x-y_t,t)$ is bounded from above on $(0,T]$. This fact forces $x - y_t$ to converge to 0 as $t \to 0$, see assumption $\mathcal{K}_3$. From inequality (8) and the positivity of k,

$$(\mathrm{cl}\, f)(x) \leq \liminf_{t\to 0} f(y_t) \leq \liminf_{t\to 0}(S+t) = S,$$

which remained to prove.

(b) We may assume $x_0 = 0$ without loss of generality. We claim that there exists $\tau \in (0,T]$ such that

$$f_t(x) = \inf_{y\in B_r}(f(y) + k(x-y,t)) \text{ for all } x \in B_\rho \text{ and } t \in (0,\tau]. \tag{9}$$

A sufficient condition for (9) is the existence of $\tau \in (0,T]$ with

$$f(y) + k(x-y,\tau) \geq \sup_{B_\rho} f + 1$$

for all $x \in B_\rho$ and $y \in X \setminus B_r$. Because of the uniform convergence (5), we can find a positive R such that

$$f(y) + k(x-y,T) \geq \sup_{B_\rho} f + 1$$

for all $x \in B_\rho$ and $y \in X \setminus B_R$. If $R \leq r$ we have verified (9); if not, thanks to the hypothesis $\mathcal{K}_3$ on the kernel k, there exists $\tau \in (0,T]$ such that

$$\inf_{B_R\setminus B_r} f + k(x-y,\tau) \geq \sup_{B_\rho} f + 1$$

for all $x \in B_\rho$, $y \in B_R \setminus B_r$. Therefore, (9) holds with this choice of τ. Let us show that

$$\epsilon(t) := \sup_{x\in B_\rho}(f(x) - f_t(x))$$

goes to zero with t. Obviously, $\epsilon(\cdot)$ is nondecreasing and positive-valued. Because of (9) we have

$$\epsilon(t) = \sup_{x\in B_\rho} \sup_{y\in B_r} (f(x) - f(y) - k(x-y,t))$$

for all $t \in (0,\tau]$. Since f is uniformly continuous on B_r, there exists $m \in \mathcal{M}$ such that $|f(x) - f(y)| \leq m(\|x-y\|)$ whenever x and y belong to B_r. Hence,

$$\begin{aligned} \epsilon(t) &\leq \sup_{x\in B_\rho} \sup_{y\in B_r} (m(\|x-y\|) - k(x-y,t)) \\ &= \sup_{z\in B_{\rho+r}} (m(\|z\|) - k(z,t)). \end{aligned}$$

If there is a $C > 0$ such that $\epsilon(t) > C$ for every $t > 0$, then, for every $t \in (0,\tau]$, we can find z_t in $B_{\rho+r}$ with

$$m(\|z_t\|) - k(z_t,t) > C. \tag{10}$$

From the inequality (10) and hypothesis $\mathcal{K}_3$ on the kernel k it follows that $\|z_t\|$ goes to zero with t. But (10) tells us that $m(\|z_t\|) > C > 0$ for all $t \in (0,\tau]$ — we have arrived at a contradiction. □

EXAMPLE 5. Let us take $k(\cdot,t) = tL(\cdot/t)$ where $L: X \to \mathbb{R}_+$ is convex, coercive and $L(0) = 0$. For the case $X = \mathbb{R}$, $f: X \to \mathbb{R}$ lower semicontinuous and $f + c|\cdot|$ bounded from below for some $c \in \mathbb{R}_+$, let us define $u: X \times \mathbb{P} \to \mathbb{R}$ by $u(\cdot,t) = f +_e k(\cdot,t)$ for each $t \in \mathbb{P}$. It is not hard to show that k satisfies $\mathcal{K}_1$, $\mathcal{K}_2$ and $\mathcal{K}_3$. According to Theorem 2, $u(\cdot,t)$ converges pointwise to f as $t \to 0$ and the Lipschitz continuity on bounded sets of $k(\cdot,t)$ is transferred to $u(\cdot,t)$. Actually, u has very interesting properties: u is the viscosity solution of the *Hamilton-Jacobi equation*

$$\dot{u}(x,t) + L^*(u'(x,t)) = 0, \quad (x,t) \in X \times \mathbb{P},$$

with initial condition $\lim_{t\to 0} u(x,t) = f(x)$, consult [6] and [14]. Here, the prime and the dot denote partial differentiation with respect to the first and second variable, respectively, and L^* denotes the conjugate function to L. For the case of X a reflexive Banach space, f convex and lower semicontinuous, see [2] and [16].

Assume that $k(\cdot,t) = (1/t)\Phi \circ \|\cdot\|$, where $\Phi: \mathbb{R}_+ \to \mathbb{R}_+$ is nondecreasing, continuous, $\Phi(0) = 0$ and $\Phi(r) > 0$ for all $r > 0$. Obviously, $k(\cdot,t)$ is uniformly continuous on each bounded set. Let f be proper and bounded from below on each bounded subset of X and

$$f(y) + \frac{1}{T}\Phi(\|y\| - r) \to +\infty \text{ as } \|y\| \to +\infty \tag{11}$$

for some $T \in \mathbb{P}$ and each $r \in \mathbb{P}$. Then Theorem 2 applies: provided $t \in (0,T]$, f_t is uniformly continuous on each bounded subset of X and $f_t(x) \nearrow (\operatorname{cl} f)(x)$ as $t \searrow 0$ for every $x \in X$. If Φ is Lipschitz continuous on bounded intervals (for instance if Φ is a convex function), then $k(\cdot, t)$ is Lipschitz continuous on each bounded subset of X and f_t inherits this property for any $t \in (0, T]$.

For example, when $\Phi(r) = e^r - 1$, then (11) is satisfied if there exist $a \in \mathbb{R}_+$ and $b \in [0, 1)$ such that

$$f(y) + a \exp(b\|y\|) \geq 0$$

for all $y \in X$. In the case of Φ a convex power function, we have the following well known result [5].

Proposition 3. *Let f be a proper function defined on X such that, given $p \geq 1$, for some real $\alpha \geq 0$ and β, $f + \alpha\|\cdot\|^p + \beta \geq 0$. Then for any $0 < t < 2^{1-p}/(\alpha p)$,*

$$f_t := f +_{\mathrm{e}} \left(\frac{1}{pt}\|\cdot\|^p\right)$$

satisfies

$$|f_t(x) - f_t(y)| \leq \frac{c}{t}\|x - y\|,$$

where the Lipschitz "constant" c depends continuously on $\|x\|$, $\|x - y\|$, α, t and p; it depends on f only through the value $f(x_f)$ and the norm $\|x_f\|$ at a point x_f at which f is finite. Moreover, $f_t(x) \nearrow (\mathrm{cl}\, f)(x)$ as $t \searrow 0$.

We close this paper with a theorem that states that any lower semicontinuous function which is bounded from below on bounded sets may be approximated with Lipschitz continuous functions by means of a generalized Moreau–Yosida approximation procedure.

Theorem 3. *Let f be a proper function on X which is bounded from below on each bounded subset of X. Then there exists a convex and continuous function $\Phi: \mathbb{R}_+ \to \mathbb{R}_+$ with $\Phi(0) = 0$ and $\Phi(r) > 0$ for all $r > 0$ such that the following hold.*

(a) *Whenever $t \in \mathbb{P}$,*

$$f_t := f +_{\mathrm{e}} \left(\frac{1}{t}\Phi \circ \|\cdot\|\right)$$

is finitely valued and Lipschitz continuous on each bounded subset of X.

(b) *$f_t(x) \nearrow (\mathrm{cl}\, f)(x)$ as $t \searrow 0$ for every $x \in X$.*

(c) *If f is uniformly continuous on the ball $B_r(x_0)$, then, for each $\rho \in (0, r)$, f_t converges uniformly to f on the ball $B_\rho(x_0)$.*

(d) *If f is uniformly continuous on bounded sets, then, for each $T \in \mathbb{P}$, $(f_t)_{t \in (0,T]}$ is equi-uniformly continuous on bounded sets.*

PROOF. We start by showing that there exists a convex continuous $\Phi: \mathbb{R}_+ \to \mathbb{R}_+$ with $\Phi(0) = 0$ and $\Phi(r) > 0$ for all $r > 0$ such that

$$f(y) + \frac{1}{t}\Phi(\|y\| - r) \to +\infty \text{ as } \|y\| \to +\infty$$

for all t and r in $\mathbb{P}$. Let

$$\varphi(\rho) = -\inf\{f(y) \mid y \in B_\rho\}$$

for all $\rho \in \mathbb{R}_+$. We observe that φ is nondecreasing, does not take the value $+\infty$ and that $f(y) \geq -\varphi(\|y\|)$ for every $y \in X$. Take $\rho_0 \in \mathbb{R}_+$ such that $\varphi(\rho_0) > -\infty$ and set

$$\Phi(r) = \frac{r^2}{2} + \int_{\rho_0}^{\rho_0+4r} (\varphi(\rho) - \varphi(\rho_0))\, d\rho$$

for every $r \in \mathbb{R}_+$. Then Φ is convex, continuous at 0, $\Phi(0) = 0$ and $0 < \Phi(r) < +\infty$ for every $r > 0$.

The following inequalities hold for every $r \in \mathbb{R}_+$:

$$\Phi(r) \geq \frac{r^2}{2} + \int_{\rho_0+2r}^{\rho_0+4r} (\varphi(\rho) - \varphi(\rho_0))\, d\rho \geq \frac{r^2}{2} + 2r[\varphi(2r + \rho_0) - \varphi(\rho_0)].$$

Choose $r, t \in \mathbb{P}$ arbitrarily. If $\|y\| \geq 2r$, then

$$\begin{aligned}
&f(y) + \frac{1}{t}\Phi(\|y\| - r) \\
\geq\;& -\varphi(\|y\|) + \frac{1}{t}\left[\frac{1}{2}(\|y\| - r)^2 + 2(\|y\| - r)[\varphi(2(\|y\| - r) + \rho_0) - \varphi(\rho_0)]\right] \\
\geq\;& -\varphi(\|y\|) + \frac{1}{t}\left[\frac{1}{2}(\|y\| - r)^2 + 2(\|y\| - r)[\varphi(\|y\|) - \varphi(\rho_0)]\right] \\
=\;& \left[\frac{2}{t}(\|y\| - r) - 1\right]\varphi(\|y\|) + \frac{1}{t}\left[\frac{1}{2}(\|y\| - r)^2 - 2(\|y\| - r)\varphi(\rho_0)\right].
\end{aligned}$$

Hence

$$f(y) + \frac{1}{t}\Phi(\|y\| - r) \to +\infty \text{ as } \|y\| \to +\infty.$$

The kernel $k(\cdot, t) = (1/t)\Phi \circ \|\cdot\|$ is in this case Lipschitz continuous on bounded sets. By Theorem 2 part (a), $f_t = f +_e k(\cdot, t)$ is Lipschitz continuous on bounded sets and increases to $\operatorname{cl} f$ as t decreases to zero, as asserted in (a) and (b). Assertions (c) and (d) follow from Theorem 2 (b) and Proposition 2 (d), respectively. □

References

[1] ATTOUCH, H. *Variational Convergence for Functions and Operators*, Applicable Math. Series, Pitman, London 1984.

[2] ATTOUCH, H. *Viscosity Solutions of Minimization Problems*, Seminaire d'Analyse Convexe de Montpellier, exposé 8, 1992.

[3] ATTOUCH, H. AND AZÉ, D. *Approximation and Regularization of Arbitrary Functions in Hilbert Spaces by the Lasry-Lions Method*, Annales de l'Institut H. Poincaré - Analyse non linéaire (to appear).

[4] Attouch, H. and Wets, R. *Epigraphical Analysis*, in "Analyse non Linéaire", eds. H. Attouch, J.-P Aubin, F.H. Clarke, I. Ekeland, 73-99, Gauthiers-Villars Paris et C.R.M. Montreal, 1989.

[5] Attouch, H. and Wets, R. *Quantitative Stability of Variational Systems: I. The Epigraphical Distance*, Trans. Amer. Math. Soc. 328, 695-729 (1991).

[6] Benton, S.H. *The Hamilton-Jacobi Equation: A Global Approach*, Academic Press, New York-San Francisco-London 1977.

[7] Bougeard, M., Penot, J.-P. and Pommellet, A. *Towards Minimal Assumptions for the Infimal Convolution Regularization*, J. Approx. Theory 64, 245-270 (1991).

[8] Fenchel, W. *Convex Cones, Sets, and Functions*, Lecture Notes, Princeton University, Princeton, New Jersey 1953.

[9] Fitzpatrick S. and Phelps, R.R. *Bounded Approximants to Monotone Operators on Banach Spaces*, Annales de l'Institut H. Poincaré - Analyse non linéaire Vol. 9, n° 5, 573-595 (1992).

[10] Hausdorff, F. *Über halbstetige Funktionen und deren Verallgemeinerung*, Math. Zeit. 5, 292-309 (1919).

[11] Hiriart-Urruty, J.-B. *Extension of Lipschitz Functions*, J. Math. Anal. Appl. 77, 539-544 (1980).

[12] Lasry, J.-M. and Lions, P.-L. *A Remark on Regularization in Hilbert Spaces*, Israel J. Math. 55, No. 3, 257-266 (1986).

[13] Laurent, P.-J. *Approximation et optimisation*, Hermann, Paris 1972.

[14] Lions, P.L. *Generalized Solutions of Hamilton-Jacobi Equations*, Pitman, London 1982.

[15] Moreau, J.J. *Fonctionelles Convexes*, Séminaire sur les équations aux dérivées partielles II, Collège de France 1966-1967.

[16] Plazanet, P. *Contributions à l'analyse des fonctions convexes ...*, Thèse Université Paul Sabatier, 1990.

[17] Wets, R. *Convergence of Convex Functions, Variational Inequalities and Convex Optimization Problems*, in "Variational Inequalities and Complementary Problems" (P. Cottle *et al.*, Eds.), Wiley, Chichester, 1980.

E

On the Operation of Infimal Convolution and Regularization of Lower Semicontinuous Functions

THOMAS STRÖMBERG

Department of Mathematics, Luleå University of Technology, S-971 87 Luleå, Sweden

Abstract

The continuity and differentiability properties of the infimal convolute of two convex functions are investigated. The results obtained are applied to the regularization of a lower semicontinuous and quadratically minorized function f, not necessarily convex, defined on a Banach space X admitting an equivalent renorm such that simultaneously X and its dual X^* are locally uniformly rotund.

1 Introduction and preliminaries

In this paper we deal with semicontinuity, continuity and differentiability properties of the infimal convolute of two convex functions. Throughout X stands for a real Hausdorff topological vector space. We confine ourselves to the following notations, which are of a common use in the field of convex analysis; below $f: X \to \bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ is an extended real-valued function on X:

- f is termed *proper* if it is nowhere equal to $-\infty$ and somewhere equal to a real number.
- The *essential domain* of f, denoted $\operatorname{dom} f$, is the set of all x in X for which $f(x) < +\infty$.
- The *strict epigraph* $\operatorname{epi}_s f$ is the subset of $X \times \mathbb{R}$ consisting of all $(x, \alpha) \in X \times \mathbb{R}$ satisfying the relation $f(x) < \alpha$.
- The *lower semicontinuous regularization* $\operatorname{cl} f$ of f is defined by

$$(\operatorname{cl} f)(x) = \liminf_{\xi \to x} f(\xi) \text{ at every } x \in X.$$

- By $\Gamma(X)$ we understand the set of all convex, lower semicontinuous and proper functions from X into $\mathbb{R} \cup \{+\infty\}$.

- For any $A \subseteq X$ the *indicator function* δ_A of A is defined by putting $\delta_A(x) = 0$ if $x \in A$ and $\delta_A(x) = +\infty$ if $x \in X \setminus A$; $\delta_A \in \Gamma(X)$ if and only if A is nonempty, closed and convex.

- Suppose that X is locally convex with dual X^*. The *conjugate function* $f^*: X^* \to \bar{\mathbb{R}}$ of f is defined by the assignment

$$f^*(x^*) = \sup_{x \in X}(\langle x^*, x\rangle - f(x)) \text{ for each } x^* \in X^*.$$

Symmetrically, for any $\psi: X^* \to \bar{\mathbb{R}}$, ψ^* is the extended real-valued function on X that to each $x \in X$ assigns the extended real number

$$\psi^*(x) = \sup_{x^* \in X^*} (\langle x^*, x\rangle - \psi(x^*)).$$

The *subdifferential* ∂f of f is the following subset of $X \times X^*$:

$$\partial f = \{(x, x^*) \in X \times X^* | \langle x^*, x\rangle = f(x) + f^*(x^*)\}.$$

- The operation of *upper addition* $\dot{+}$ (respectively, *lower addition* $\underset{\cdot}{+}$) is defined on $\bar{\mathbb{R}}$ as the commutative extension of usual addition for which $(+\infty) \dot{+} (-\infty) = +\infty$ (respectively, $(+\infty) \underset{\cdot}{+} (-\infty) = -\infty$). Since the notions of upper and lower addition coincide when $+\infty$ and $-\infty$ never both occur as values we write $+$ in such a case.

If f and g are extended real-valued functions defined on X, then their *infimal convolute* $f \square g$ is the function defined on X that sends each $x \in X$ to the extended real number

$$(f \square g)(x) = \inf_{y \in X}(f(y) \dot{+} g(x-y)).$$

The infimal convolution of f with g is said to be *exact at* x if the above infimum is achieved, and *exact* if it is exact at each point of X. The reader should observe that the operation of infimal convolution includes Minkowski addition of subsets of X in the sense that—specializing to indicator functions—$\delta_A \square \delta_B$ is equal to δ_{A+B} whenever A and B are subsets of X. Geometrically, performing infimal convolution amounts to the adding of strict epigraphs:

$$\mathrm{epi}_s\, f \square g = \mathrm{epi}_s\, f + \mathrm{epi}_s\, g.$$

The previous identity has led several authors to use the term "epigraphical addition" instead of "infimal convolution" [4]. Another important identity reads

$$(f \square g)^*(x^*) = f^*(x^*) \underset{\cdot}{+} g^*(x^*) \text{ for all } x^* \in X^*,$$

where X is now locally convex with topological dual X^*. We emphasize that a main feature of the operation is that a regularity property of g is frequently, under mild conditions on f, carried over to $f \Box g$. This applies to uniform continuity [9, 12, 25], in the convex case to continuity and, provided X is reflexive, to first order differentiability (see sections 2 and 3). However, performing infimal convolution of convex C^∞ or real analytic functions defined on the real line need not result in an infinitely smooth function, as demonstrated by J. Boman [5, 6] and C. O. Kiselman [14, 15].

The infimal convolute of two convex functions is always convex, but we should be aware of the fact that $\Gamma(X)$ is not stable under the operation of infimal convolution because $f \Box g$, with $f, g \in \Gamma(X)$, need not be lower semicontinuous or even proper. The statement about possible absence of lower semicontinuity should come as no surprise since it is well-known that closedness of the Minkowski sum of two convex closed (noncompact) sets A and B may fail (even when $\dim X = 2$), so that $\delta_A \Box \delta_B \notin \Gamma(X)$.

The outline of this paper is as follows. In sections 2 and 3 we consider, respectively, continuity and differentiability properties of the infimal convolute of two convex functions. In particular, in section 2 we give a partial answer to a question raised by H. Attouch and H. Brezis [2]. Moreover, in section 3 we study conditions ensuring that Fréchet differentiability of g is transferred to $f \Box g$. These results extend or complement existing ones [4, 17, 27]. In the fourth section we derive a procedure for the regularization of a lower semicontinuous and quadratically minorized function f, not necessarily convex, defined on a Banach space X admitting an equivalent renorm such that simultaneously X and its dual X^* are locally uniformly rotund. A complementary example and an additional result are given in our final section.

2 Basic continuity properties of the infimal convolute of convex functions

Suppose that X is locally convex. If $f, g\colon X \to \bar{\mathbb{R}}$, then $(f^* + g^*)^* = (f \Box g)^{**} \leq f \Box g$ with equality if and only if $f \Box g$ is convex and lower semicontinuous. Suppose that f and g are convex proper with $\operatorname{dom} f^* \cap \operatorname{dom} g^* \neq \emptyset$. Then $f \Box g$ is proper and $(f^* + g^*)^* = \operatorname{cl}(f \Box g)$ so that, given $x \in X$, $(f \Box g)(x) = (f^* + g^*)^*(x)$ if and only if $f \Box g$ is lower semicontinuous at x. In this connection we mention the following question which was raised by H. Attouch and H. Brezis in [2].

QUESTION. Assume $f, g \in \Gamma(X)$. When can we conclude that $f \Box g$ is lower semicontinuous?

Our results in this section are on the continuity of $f \Box g$ for the case where f, g are convex functions, enjoying some continuity properties. In particular, some positive answers of the above question are obtained.

Proposition 1 *Let X be a Hausdorff topological vector space and let $f, g\colon X \to \mathbb{R}\cup\{+\infty\}$ be convex proper functions.*

(a) *Suppose that f or g is continuous at some point, or that $\operatorname{epi}_s f + \operatorname{epi}_s g$ has a nonempty interior. Then the interior of $\operatorname{dom} f + \operatorname{dom} g$ is nonempty, and $f \Box g$ is either equal to $-\infty$ on the interior of $\operatorname{dom} f + \operatorname{dom} g$ or is a proper function which is continuous on $\operatorname{int}(\operatorname{dom} f + \operatorname{dom} g)$. In particular, if $\operatorname{dom} f + \operatorname{dom} g = X$ and f and g share a common affine minorant, then $f \Box g$ is real-valued and continuous.*

(b) *If f or g is upper semicontinuous, then $f \Box g$ is upper semicontinuous, $\operatorname{dom} f + \operatorname{dom} g$ is open and nonempty, $f \Box g$ is either identically $-\infty$ on $\operatorname{dom} f + \operatorname{dom} g$, or is a proper function which is continuous on $\operatorname{dom} f + \operatorname{dom} g$.*

Suppose that X is locally convex and $\operatorname{dom} f^* \cap \operatorname{dom} g^* \neq \emptyset$.

(c) *Suppose that f or g is real-valued and continuous at some point and that x belongs to the interior of $\operatorname{dom} f + \operatorname{dom} g$. Then*

$$(f \Box g)(x) = \max_{x^* \in X^*} \left(\langle x^*, x\rangle - f^*(x^*) - g^*(x^*)\right) \in \mathbb{R}, \tag{1}$$

and $\partial(f \Box g)(x)$ is equal to the (nonempty) set of maximizers of (1). Moreover, $f \Box g$ is Gâteaux differentiable at x provided $f^* + g^*$ is strictly convex over $\operatorname{dom} f^* \cap \operatorname{dom} g^*$.

(d) *Let $\mathcal{A}$ denote the affine hull of $\operatorname{dom} f + \operatorname{dom} g$ and endow $\mathcal{A}$ with its subspace topology. Moreover, suppose that $g|_{\mathcal{A}}$ is real-valued and continuous at some point. Then the $\mathcal{A}$-interior of $\operatorname{dom} f + \operatorname{dom} g$ is nonempty, and*

$$(f \Box g)(x) = \max_{x^* \in X^*} \left(\langle x^*, x\rangle - f^*(x^*) - g^*(x^*)\right)$$

for all x in the $\mathcal{A}$-interior of $\operatorname{dom} f + \operatorname{dom} g$. Moreover, if $\operatorname{dom} f + \operatorname{dom} g$ is a closed flat, that is to say a translate of a closed linear subspace, then $f \Box g \in \Gamma(X)$.

For the proof we recall the following basic result, see for instance [13, Theorem 1, page 170].

Lemma 1 *Let X be a Hausdorff topological vector space and $h\colon X \to \overline{\mathbb{R}}$ a convex function. The following are equivalent:*

(i) $\operatorname{epi}_s h$ *has a nonempty interior;*

(ii) *The interior of $\operatorname{dom} h$ is nonempty and, either*

   (a) *h is a proper function which is continuous over the interior of $\operatorname{dom} h$, or*

   (b) *h is equal to $-\infty$ everywhere on the interior of $\operatorname{dom} h$.*

PROOF OF PROPOSITION 1. (a) Suppose that g is continuous at some point. Then

$$\operatorname{int} \operatorname{epi}_s f \Box g = \operatorname{int}(\operatorname{epi}_s f + \operatorname{epi}_s g) \neq \emptyset,$$

since $\operatorname{epi}_s f \neq \emptyset$ and the strict epigraph of g possesses a nonempty interior (see Lemma 1). In virtue of Lemma 1 there are two exclusive alternatives: either $f \Box g$ is proper, and continuous over the (nonempty) interior of $\operatorname{dom} f \Box g$ (that is, over the interior of $\operatorname{dom} f + \operatorname{dom} g$), or $f \Box g$ is equal to $-\infty$ on the interior of $\operatorname{dom} f + \operatorname{dom} g$.

(b) Assume that g is upper semicontinuous, which means that $\operatorname{epi}_s g$ is an open subset of $X \times \mathbb{R}$. Hence, $f \Box g$ possesses an open strict epigraph, since it is the sum of the strict epigraphs of f and g. With Lemma 1 in mind, the statements of part (b) follow.

(c) The assumption $\operatorname{dom} f^* \cap \operatorname{dom} g^* \neq \emptyset$ implies that $f \Box g$ is proper. The convolute $f \Box g$ is continuous on $\operatorname{int}(\operatorname{dom} f + \operatorname{dom} g)$, so that $\partial(f \Box g)(x)$ is nonempty and equal to the set of maximizers of $x^* \mapsto \langle x^*, x\rangle - (f \Box g)^*(x^*)$. Hence,

$$\begin{aligned} \partial(f \Box g)(x) &= \arg\max\{\langle x^*, x\rangle - f^*(x^*) - g^*(x^*) \mid x^* \in \operatorname{dom} f^* \cap \operatorname{dom} g^*\} \text{ and} \\ (f \Box g)(x) &= \max\{\langle x^*, x\rangle - f^*(x^*) - g^*(x^*) \mid x^* \in \operatorname{dom} f^* \cap \operatorname{dom} g^*\}. \end{aligned}$$

If $f^* + g^*$ is strictly convex on $\operatorname{dom} f^* \cap \operatorname{dom} g^*$, then there is ecactly one maximizer, that is, $\partial(f \Box g)(x)$ is a singleton, whence $f \Box g$ is Gâteaux differentiable at x.

(d) We may assume $0 \in \operatorname{dom} f \cap \operatorname{dom} g$ without loss of generality. (If $0 \notin \operatorname{dom} f \cap \operatorname{dom} g$ one may consider $\tilde{f} = f(\cdot + y_0)$ and $\tilde{g} = g(\cdot + z_0)$, where $y_0 \in \operatorname{dom} f$ and $z_0 \in \operatorname{dom} g$.)

Case 1: $\mathcal{A} = X$. In this situation the conclusion follows from (c).

Case 2: $\mathcal{A} \neq X$. Observe that $\operatorname{dom} f$ and $\operatorname{dom} g$ are both subsets of $\operatorname{dom} f + \operatorname{dom} g$ since 0 is an element of $\operatorname{dom} f \cap \operatorname{dom} g$, and that $\mathcal{A}$ is a linear subspace of X, hence a locally convex topological vector space in itself. Choose x in the $\mathcal{A}$-interior of $\operatorname{dom} f + \operatorname{dom} g$. Then there exists, by Case 1, $a^* \in \mathcal{A}^*$ verifying

$$(f|_{\mathcal{A}} \Box g|_{\mathcal{A}})(x) = \langle a^*, x\rangle - (f|_{\mathcal{A}})^*(a^*) - (g|_{\mathcal{A}})^*(a^*).$$

Hence

$$\begin{aligned} (f \Box g)(x) &= \inf\{f(y) + g(z) \mid (y, z) \in \operatorname{dom} f \times \operatorname{dom} g \text{ and } y + z = x\} \\ &= \inf\{f|_{\mathcal{A}}(y) + g|_{\mathcal{A}}(z) \mid (y, z) \in \mathcal{A} \times \mathcal{A} \text{ and } y + z = x\} \\ &= \langle a^*, x\rangle - (f|_{\mathcal{A}})^*(a^*) - (g|_{\mathcal{A}})^*(a^*). \end{aligned}$$

According to the Hahn–Banach theorem, a^* extends to an element x^* of X^*. Now,

$$\begin{aligned} (f|_{\mathcal{A}})^*(a^*) &= \sup_{x \in \mathcal{A}}(\langle a^*, x\rangle - f|_{\mathcal{A}}(x)) \\ &= \sup_{x \in X}(\langle x^*, x\rangle - f(x)) = f^*(x^*), \end{aligned}$$

and similarly $(g|_{\mathcal{A}})^*(a^*) = g^*(x^*)$, so that

$$(f \Box g)(x) = \langle x^*, x\rangle - f^*(x^*) - g^*(x^*),$$

proving the assertion since x was arbitrarily chosen inside the $\mathcal{A}$-interior of $\operatorname{dom} f + \operatorname{dom} g$.

Suppose that $\operatorname{dom} f + \operatorname{dom} g$ is a closed flat. Then $f \Box g$ is lower semicontinuous on $\mathcal{A} = \operatorname{dom} f + \operatorname{dom} g$, and $f \Box g$ is obviously lower semicontinuous on $X \setminus \mathcal{A}$ since $X \setminus \mathcal{A}$ is open and $f \Box g$ is equal to $+\infty$ everywhere on $X \setminus \mathcal{A}$. □

Corollary 1 (R. T. Rockafellar [20].) *Assume that X is a locally convex Hausdorff topological vector space. Let f and g be proper and convex functions $X \to \mathbb{R} \cup \{+\infty\}$.*

(a) *If* $\operatorname{dom} f \cap \operatorname{dom} g$ *contains a point at which f or g is continuous, then*

$$(f+g)^* = f^* \Box g^*$$

with exact infimal convolution.

(b) *Suppose that X is reflexive. Let f and g belong $\Gamma(X)$. If* $\operatorname{dom} f^* \cap \operatorname{dom} g^*$ *contains a point at which f^* or g^* is continuous, then*

$$f \Box g = (f^* + g^*)^*$$

with exact infimal convolution. Consequently, $f \Box g \in \Gamma(X)$.

PROOF. Let $x^* \in X^*$. Let $\breve{g}$ denote the function $x \mapsto g(-x)$. Notice that

$$\operatorname{dom}(f - x^*) + \operatorname{dom} \breve{g} = \operatorname{dom} f - \operatorname{dom} g,$$

and that 0 is a member of the interior of $\operatorname{dom} f - \operatorname{dom} g$. By Proposition 1,

$$[(f - x^*) \Box \breve{g}](0) = \max_{y^* \in X^*} (\langle y^*, 0 \rangle - (f - x^*)^*(y^*) - \breve{g}^*(y^*)),$$

which may be rewritten in the desired form $(f+g)^*(x^*) = (f^* \Box g^*)(x^*)$. For (b) it suffices to notice that if f and g are be replaced by f^* and g^*, respectively, in (a), then

$$(f^* + g^*)^* = f^{**} \Box g^{**} = f \Box g.$$

The proof is complete. □

REMARK. The conclusion of part (b) of the corollary may fail in a nonreflexive space. The following counter example is due to H. Attouch and H. Brezis [2]. Let $(X, \|\cdot\|)$ be a nonreflexive normed space with closed unit ball U. Then there exists a closed hyperplane H in X such that $1 = \inf_{x \in H} \|x\|$ is not achieved. Put $f = \delta_U$ and $g = \delta_H$. Then f^* is real-valued and continuous and nevertheless $f \Box g \neq (f^* + g^*)^*$. Indeed, $f \Box g = \delta_{U+H}$ is not lower semicontinuous because $U + H$ is not closed: $0 \notin U + H$ but $0 \in \operatorname{cl}(U + H)$.

A few years ago the classical Fenchel's duality theorem has been formulated for "cs-convex" functions defined on a Fréchet space. A proper extended real-valued function f on a topological vector space X is called *cs-convex* provided that, for each sequence $\langle x_n \rangle_{n \in \mathbb{N}}$ in X and each sequence $\langle t_n \rangle_{n \in \mathbb{N}}$ of numbers in $[0,1]$ with $\sum_{n=0}^{\infty} t_n = 1$, the following implication holds:

$$\sum_{n=0}^{\infty} t_n x_n \text{ convergent } \Rightarrow f\left(\sum_{n=0}^{\infty} t_n x_n\right) \leq \liminf_{N \to \infty} \sum_{n=0}^{N} t_n f(x_n).$$

Every cs-convex function is convex, and every element of $\Gamma(X)$ is cs-convex. In the following, if C is a nonempty convex subset of X, then we put

$$\operatorname{cone} C = \{\lambda x \mid \lambda \in \mathbb{R}_+ \text{ and } x \in C\}.$$

The *strong quasi relative interior* $\operatorname{sqri} C$ of C is the set of all $x \in C$ for which $\operatorname{cone}(C - x)$ is a closed linear subspace of X, see [7, 8]. The main feature of the strong quasi relative interior is that it will be frequently nonempty even when the topological interior is empty. When X is a Fréchet space we find it fruitful to use the following version, due to S. Simons [22], of Fenchel's duality theorem.

Lemma 2 *Let X be a Fréchet space. Suppose that f and g are proper cs-convex functions $X \to \mathbb{R} \cup \{+\infty\}$ such that $0 \in \operatorname{sqri}(\operatorname{dom} f - \operatorname{dom} g)$. Then*

$$(f+g)^* = f^* \Box g^*$$

with exact infimal convolution.

Proposition 2 *Assume that X is a Fréchet space. Let both $f: X \to \mathbb{R} \cup \{+\infty\}$ and $g: X \to \mathbb{R} \cup \{+\infty\}$ be proper and cs-convex, and $\operatorname{dom} f^* \cap \operatorname{dom} g^* \neq \emptyset$.*

(a) *Suppose that $x \in \operatorname{sqri}(\operatorname{dom} f + \operatorname{dom} g)$. Then*

$$(f \Box g)(x) = \max_{x^* \in X^*} (\langle x^*, x \rangle - f^*(x^*) - g^*(x^*)). \tag{2}$$

Moreover, $\partial(f \Box g)(x)$ is equal to the (nonempty) set of maximizers of (2).

(b) *If $\operatorname{dom} f + \operatorname{dom} g$ is a closed flat, then*

$$(f \Box g)(x) = \max_{x^* \in X^*} (\langle x^*, x \rangle - f^*(x^*) - g^*(x^*))$$

for all $x \in \operatorname{dom} f + \operatorname{dom} g$. Moreover, $f \Box g \in \Gamma(X)$.

(c) *If $\operatorname{dom} f + \operatorname{dom} g = X$, then $f \Box g$ is real-valued, convex and continuous. Moreover,*

$$(f \Box g)(x) = \max_{x^* \in X^*} (\langle x^*, x \rangle - f^*(x^*) - g^*(x^*))$$

for each $x \in X$. If, in addition, $f^* + g^*$ is strictly convex over $\operatorname{dom} f^* \cap \operatorname{dom} g^*$, then $f \Box g$ is Gâteaux differentiable.

(d) *Suppose X to be a reflexive Banach space, and $f, g \in \Gamma(X)$. If 0 belongs to* $\mathrm{sqri}(\mathrm{dom}\, f^* - \mathrm{dom}\, g^*)$, *then the infimal convolution $f \square g$ is exact and $f \square g$ belongs to $\Gamma(X)$.*

PROOF. (a) Define $g_x: X \to \mathbb{R} \cup \{+\infty\}$ by $g_x(y) = g(x-y)$ for all $y \in X$, so that g_x is proper and cs-convex. Clearly, $\mathrm{dom}\, g_x = x - \mathrm{dom}\, g$ and, hence, $0 \in \mathrm{sqri}(\mathrm{dom}\, f - \mathrm{dom}\, g_x)$. According to Lemma 2,

$$\begin{aligned}(f \square g)(x) &= -(f+g_x)^*(0) \\ &= -\min\{f^*(y^*) + (g_x)^*(-y^*) \mid y^* \in X^*\} \\ &= -\min\{f^*(y^*) - \langle y^*, x\rangle + g^*(y^*) \mid y^* \in X^*\} \\ &= \max\{\langle y^*, x\rangle - f^*(y^*) - g^*(y^*) \mid y^* \in X^*\},\end{aligned}$$

establishing the first part of the theorem. Since $(f \square g)(x) = (f^* + g^*)^*(x)$ we have

$$\partial(f \square g)(x) = \arg\max\{\langle x^*, x\rangle - f^*(x^*) - g^*(x^*) \mid x^* \in X^*\},$$

completing the proof.

(b) In this case $\mathrm{sqri}(\mathrm{dom}\, f + \mathrm{dom})$ is equal to $\mathrm{dom}\, f + \mathrm{dom}$. The proof follows by applying (a) and arguing as in the proof of Proposition 1 (d).

A convex and lower semicontinuous functional on a barrelled space is continuous [11, Corollary 2.5, page 13]. Every Fréchet space is barrelled [21, page 60]. Hence the continuity part of (c) is established. The statement about Gâteaux differentiability may be proved in the same way as in Proposition 1 (c).

(d) Since X is a reflexive Banach space, we may directly apply Lemma 2 with f and g relaced by f^* and g^*, respectively. □

Corollary 2 *Suppose that X is a Fréchet space. Let $f: X \to \mathbb{R} \cup \{+\infty\}$ be proper and cs-convex. If $x \in \mathrm{sqri}(\mathrm{dom}\, f)$, then $\partial f(x) \neq \emptyset$.*

PROOF. It suffices to apply Proposition 2 (a) with $g = \delta_{\{0\}}$. □

3 Fréchet differentiability of the infimal convolute of convex functions

Let f and g be convex functions on a Banach space. In this section we investigate under what conditions we can conclude that $f \square g$ is Fréchet differentiable. If S and T are subsets of $X \times X^*$, viewed upon as multivalued mappings from X into X^*, then we write $S//T$ for their *parallel sum* defined by $S//T = (S^{-1} + T^{-1})^{-1}$. If the infimal convolution $f \square g$ is exact, then $\partial(f \square g) = \partial f // \partial g$, see [18, 24]. Our first result in this section reads

Theorem 1 *Let X be a Banach space, $f: X \to \mathbb{R} \cup \{+\infty\}$ be convex and proper, $g: X \to \mathbb{R}$ be convex and continuous with* $\mathrm{dom}\, f^* \cap \mathrm{dom}\, g^* \neq \emptyset$.

(a) *Suppose that $f\Box g$ is exact at some point $x_0 \in X$, so that there exists $y_0 \in X$ such that*

$$(f\Box g)(x_0) = f(y_0) + g(x_0 - y_0).$$

Suppose moreover that g is Fréchet differentiable at $x_0 - y_0$. Then $f\Box g$ is Fréchet differentiable at x_0 with $D(f\Box g)(x_0) = Dg(x_0 - y_0)$.

(b) *Assume that X is a reflexive. Suppose that $f \in \Gamma(X)$, that g is Fréchet differentiable, and that $0 \in \operatorname{sqri}(\operatorname{dom} f^* - \operatorname{dom} g^*)$. Then $f\Box g$ is real-valued, convex and Fréchet differentiable with $D(f\Box g) = \partial f // Dg$.*

(c) *Assume that X is a reflexive. Suppose that g is Fréchet differentiable, and that $0 \in \operatorname{sqri}(\operatorname{dom} f^* - \operatorname{dom} g^*)$. Then $f\Box g$ is real-valued, convex and Fréchet differentiable with $D(f\Box g) = \partial(\operatorname{cl} f)//Dg$.*

PROOF. First of all note that $f\Box g$ is real-valued, convex and continuous (see Proposition 1).

(a) It is well-known that, because of the exactness, $\partial(f\Box g)(x_0) = \{Dg(x_0 - y_0)\}$, see [16]. Moreover, since g if Fréchet differentiable at $x_0 - y_0$ there exists $\delta\colon \mathbb{R}_+ \to \mathbb{R}_+ \cup \{+\infty\}$ with $\delta(0) = 0$ and $\delta(r)/r \to 0$ as $r \to 0$ such that

$$g(x - y_0) - g(x_0 - y_0) - \langle Dg(x_0 - y_0), x - x_0\rangle \leq \delta(\|x - x_0\|) \text{ for all } x \in X.$$

Hence, for every $x \in X$,

$$\begin{aligned} (f\Box g)(x) - (f\Box g)(x_0) &\leq [f(y_0) + g(x - y_0)] - [f(y_0) + g(x_0 - y_0)] \\ &= g(x - y_0) - g(x_0 - y_0) \\ &\leq \langle Dg(x_0 - y_0), x - x_0\rangle + \delta(\|x - x_0\|) \end{aligned}$$

and it follows that $f\Box g$ is Fréchet differentiable at x_0 with $D(f\Box g)(x_0) = Dg(x_0 - y_0)$.

(b) By Proposition 2, $f\Box g$ is exact. Part (a) may be applied and the assertion follows.

(c) We claim that $f\Box g = (\operatorname{cl} f)\Box g$. The inequality $(\operatorname{cl} f)\Box g \leq f\Box g$ is immediate as $\operatorname{cl} f$ minorizes f. Fix an arbitrary $x \in X$. Take $y \in X$ and a sequence $\langle y_n\rangle$ convergent to y such that $f(y_n) \to (\operatorname{cl} f)(y)$. Since g is upper semicontinuous,

$$(\operatorname{cl} f)(y) + g(x - y) \geq \limsup_{n\to\infty}(f(y_n) + g(x - y_n)) \geq (f\Box g)(x).$$

By taking infimum over $y \in X$, it yields that $((\operatorname{cl} f)\Box g)(x) \geq (f\Box g)(x)$. Our claim follows. Because $\operatorname{cl} f \in \Gamma(X)$ and $(\operatorname{cl} f)^* = f^*$ we can apply (b) to conclude the statements of (c).□

REMARK. This theorem complements a result obtained by H. Attouch and R. J.-B Wets [4, Proposition 3.4].

Recall that a Banach space $(X, \|\cdot\|)$ is called *locally uniformly rotund* provided that for any $x \in X$ and any sequence $\langle x_n\rangle$ in X the conditions $\|x_n\| \to 1 = \|x\|$ and $\|x + x_n\| \to 2$

jointly imply $\|x - x_n\| \to 0$. It is a well-known fact that the norm of X is Fréchet differentiable, that is the norm is a Fréchet differentiable function away from zero, provided X^* is locally uniformly rotund. Every reflexive Banach space X admits a renorm such that simultaneously X and X^* are locally uniformly rotund, consult [10, page 160]. Therefore, Theorem 1 implies the following statements:

Corollary 3 *Let X be a reflexive Banach endowed with an equivalent renorm such that simultaneously X and X^* are locally uniformly rotund. Suppose that $\varphi: \mathbb{R}_+ \to \mathbb{R}_+$ is convex, differentiable, $\varphi(0) = 0$ and*

$$\frac{\varphi(r)}{r} \to \begin{cases} 0 & \text{as } r \to 0, \\ +\infty & \text{as } r \to +\infty. \end{cases}$$

Let $f \in \Gamma(X)$. Then $f\Box(\varphi\circ\|\cdot\|)$ is Fréchet differentiable and $D(f\Box(\varphi\circ\|\cdot\|)) = \partial f // J_\varphi$. Here, $J_\varphi = D(\varphi\circ\|\cdot\|)$ is the "duality mapping" associated with φ.

See [3, Theorem 1.20] for a similar result.

REMARK. It is known that there exists a nonreflexive Banach space X with a Fréchet differentiable norm and a two-dimensional linear subspace M such that $d(\cdot, M)$ fails to be Gâteaux differentiable at some point outside M ($d(x, M) = \inf_{y\in M} \|x - y\|$ for each $x \in X$), see [26]. Consequently, the same is true for $d(\cdot, M)^2 = \|\cdot\|^2\Box\delta_M$. This means that with $f = \|\cdot\|^2$ and $g = \delta_M$ we have that f is convex Fréchet differentiable and $g \in \Gamma(X)$ but, nevertheless, $f\Box g$ fails to be globally Gâteaux differentiable.

Theorem 2 *Let X be a Banach space, $f: X \to \mathbb{R} \cup \{+\infty\}$ be convex and proper, and $g: X \to \mathbb{R}$ be convex continuous such that $\operatorname{dom} f^* \cap \operatorname{dom} g^*$ is nonempty.*

(a) *Suppose that g^* is locally uniformly convex in the sense that*

$$\lim_{n\to\infty} (g^*(\tfrac{1}{2}x^* + \tfrac{1}{2}x_n^*) - \tfrac{1}{2}g^*(x^*) - \tfrac{1}{2}g^*(x_n^*)) = 0 \textit{ implies } \lim_{n\to\infty} \|x^* - x_n^*\|_* = 0$$

for any sequence $\langle x_n^* \rangle$ in $\operatorname{dom} g^*$ and any $x^* \in \operatorname{dom} g^*$. Then g and $f\Box g$ are both Fréchet differentiable.

(b) *If g is Fréchet equidifferentiable in the sense that there exists $\delta: \mathbb{R}_+ \to \mathbb{R}_+ \cup \{+\infty\}$ with $\delta(0) = 0$ and $\delta(r)/r \to 0$ as $r \to 0$ such that*

$$g(x + h) - g(x) - \langle Dg(x), h\rangle \le \delta(\|h\|) \textit{ for all } x, h \in X,$$

then $f\Box g$ is Fréchet equidifferentiable and

$$(f\Box g)(x + h) - (f\Box g)(x) - \langle D(f\Box g)(x), h\rangle \le \delta(\|h\|) \textit{ for all } x, h \in X.$$

(c) *If g is Fréchet differentiable with Dg Lipschitz continuous with constant C, then $f\Box g$ is Fréchet differentiable with*

$$(f\Box g)(x+h) - (f\Box g)(x) - \langle D(f\Box g)(x), h\rangle \le C\|h\|^2 \text{ for all } x, h \in X.$$

PROOF. According to Proposition 1, $f\Box g$ is real-valued, convex and continuous.

(a) $(f\Box g)^*$ is locally uniformly convex since $(f\Box g)^*$ is equal to the sum $f^* + g^*$, where g^* is locally uniformly convex and f^* is convex. It is then clear that the statement follows provided that we show that every convex and continuous $h\colon X \to \mathbb{R}$ with h^* locally uniformly convex is Fréchet differentiable. Note that h is Gâteaux differentiable since

$$\partial h(x) = \arg\max\{\langle x^*, x\rangle - h^*(x^*) \mid x^* \in \operatorname{dom} h^*\}$$

is a singleton for every $x \in X$. If we prove that the Gâteaux derivative ∇h is a norm-to-norm continuous map $X \to X^*$, then it follows that h is Fréchet differentiable. To this end, take $x \in X$ and a sequence $\langle x_n\rangle$ convergent to x. Write $x^* := \nabla h(x)$ and $x_n^* := \nabla h(x_n)$. Then

$$\begin{aligned}\langle x^*, x\rangle &= h(x) + h^*(x^*), \text{ and}\\ \langle x_n^*, x_n\rangle &= h(x_n) + h^*(x_n^*) \text{ for each } n \in \mathbb{N}.\end{aligned}$$

Using these identities together with the definition of h^* we obtain

$$\begin{aligned}2h^*(\tfrac{1}{2}(x^* + x_n^*)) - h^*(x^*) - h^*(x_n^*) &\ge 2(\langle \tfrac{1}{2}(x^* + x_n^*), x_n\rangle - h(x_n)) - h^*(x^*) - h^*(x_n^*)\\ &= \langle x^*, x_n\rangle - 2h(x_n) - h^*(x^*) + (\langle x_n^*, x_n\rangle - h^*(x_n^*))\\ &= \langle x^*, x_n\rangle - h(x_n) - h^*(x^*)\\ &\to \langle x^*, x\rangle - h(x) - h^*(x^*) = 0.\end{aligned}$$

Since h^* is locally uniformly convex we must have $\|x^* - x_n^*\|_* \to 0$. Consequently, ∇h is continuous.

(b) The assumption on g implies that

$$g(x+h) + g(x-h) - 2g(x) \le 2\delta(\|h\|) \text{ for all } x, h \in X.$$

Take $x \in X$ and $\varepsilon \in \mathbb{P}$. Choose $y \in X$ with

$$(f\Box g)(x) \ge f(y) + g(x-y) - \varepsilon.$$

Then

$$\begin{aligned}&(f\Box g)(x+h) + (f\Box g)(x-h) - 2(f\Box g)(x)\\ &\le [f(y) + g(x+h-y)] + [f(y) + g(x-h-y)] - 2[f(y) + g(x-y) - \varepsilon]\\ &= g(x+h-y) + g(x-h-y) - 2g(x-y) + 2\varepsilon\\ &\le 2\delta(\|h\|) + 2\varepsilon.\end{aligned}$$

Hence

$$(f\Box g)(x+h)+(f\Box g)(x-h)-2(f\Box g)(x)\leq 2\delta(\|h\|) \text{ for all } x,h\in X.$$

We conclude that $f\Box g$ is Fréchet differentiable and

$$(f\Box g)(x+h)-(f\Box g)(x)-\langle D(f\Box g)(x),h\rangle\leq \delta(\|h\|) \text{ for all } x,h\in X.$$

(c) Let $x,h\in X$. By the Mean Value Theorem, there exist $s,t\in[0,1]$ such that

$$\begin{aligned} g(x+h)+g(x-h)-2g(x) &= (g(x+h)-g(x))+(g(x-h)-g(x)) \\ &= \langle Dg(x+sh),h\rangle+\langle Dg(x-th),-h\rangle \\ &= \langle Dg(x+sh)-Dg(x-th),h\rangle. \end{aligned}$$

Hence,

$$g(x+h)+g(x-h)-2g(x)\leq \|Dg(x+sh)-Dg(x-th)\|_*\|h\|\leq 2C\|h\|^2.$$

Thus, according to (b), $f\Box g$ is Fréchet differentiable with

$$(f\Box g)(x+h)+(f\Box g)(x-h)-2(f\Box g)(x)\leq 2C\|h\|^2 \text{ for all } x,h\in X.$$

Therefore,

$$(f\Box g)(x+h)-(f\Box g)(x)-\langle D(f\Box g)(x),h\rangle\leq C\|h\|^2 \text{ for all } x,h\in X.$$

The proof is complete. □

REMARK. Part (a) complements a result due to C. Zălinescu [27, Theorem 1]. For a result closely related to assertion (c) see [17, Proposition 2.5].

If X is a Banach space with a separable dual or a $c_0(S)$ for some index set S, then X can be supplied with an equivalent norm which makes both X and X^* locally uniformly rotund, consult [10, page 160].

Corollary 4 *Let X be a Banach space if necessary renormed with an equivalent renorm such that X^* is locally uniformly rotund. Suppose that $\varphi\colon \mathbb{R}_+\to\mathbb{R}_+$ is convex, differentiable, $\varphi(0)=0$ and*

$$\frac{\varphi(r)}{r}\to\begin{cases}0 & \text{as } r\to 0,\\ +\infty & \text{as } r\to+\infty.\end{cases}$$

Suppose moreover that $f\colon X\to\mathbb{R}\cup\{+\infty\}$ is proper convex with $\operatorname{dom} f^*\neq\emptyset$. Then $f\Box(\varphi\circ\|\cdot\|)$ is Fréchet differentiable.

PROOF. Let φ^* denote the conjugate function to φ (in the sense of Young):

$$\varphi^*(\rho)=\sup_{r\in\mathbb{R}_+}(\rho r-\varphi(r)) \text{ for each } \rho\in\mathbb{R}_+.$$

It is readily verified that φ^* is real-valued and strictly convex so that $(\varphi\circ\|\cdot\|)^*=\varphi^*\circ\|\cdot\|_*$ is locally uniformly convex since X^* is locally uniformly rotund, see [27, Theorem 4.1]. The infimal convolute $f\Box(\varphi\circ\|\cdot\|)$ is then Fréchet differentiable. □

4 Pointwise approximation of a lower semicontinuous function by smooth functions

Let $(X, \|\cdot\|)$ be a Banach space. Suppose that $f: X \to \mathbb{R} \cup \{+\infty\}$ is a given proper lower semicontinuous function, not necessarily convex. A basic problem discussed in this section is how to approximate f, in the pointwise sense, with smooth functions. The result of this section depends on the classical Moreau–Yosida approximation method combined with a recent result due to R. Poliquin, J. Vanderwerff and V. Zizler [19]. These authors consider a lower semicontinuous and quadratically minorized function f defined on a locally uniformly rotund Banach space $(X, \|\cdot\|)$ and establish the existence of a convex composite representation of f: there exists $F \in \Gamma(X \times \mathbb{R})$ such that $f(x) = F(x, -\frac{1}{2}\|x\|^2)$ for all $x \in X$. Explicitly, $F = (f^\circ)^*$ where f° denotes the *quadratic conjugate function* to f, defined as the function $f^\circ: X^* \times \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ that sends each $(x^*, \alpha) \in X^* \times \mathbb{R}$ to

$$f^\circ(x^*, \alpha) = \begin{cases} \sup_{x \in X}(\langle x^*, x\rangle - f(x) - (\alpha/2)\|x\|^2) & \text{if } \alpha \geq 1, \\ +\infty & \text{if } \alpha < 1. \end{cases}$$

Lemma 3 [19] *Let $(X, \|\cdot\|)$ be a locally uniformly rotund Banach space. Suppose that $f: X \to \mathbb{R} \cup \{+\infty\}$ is proper lower semicontinuous with $f + (c/2)\|\cdot\|^2$ bounded from below for some $c \in \mathbb{R}_+$. Then $f(x) = F(x, -\frac{1}{2}\|x\|^2)$ for all $x \in X$, where $F = (f^\circ)^* \in \Gamma(X \times \mathbb{R})$.*

The lemma means that f can be reconstructed from its quadratic conjugate:

$$f(x) = \sup_{(x^*, \alpha) \in X^* \times \mathbb{R}} (\langle x^*, x\rangle - (\alpha/2)\|x\|^2 - f^\circ(x^*, \alpha)) \text{ for all } x \in X.$$

Recall that the *Moreau–Yosida approximants* $\langle F_\varepsilon \rangle_{\varepsilon \in \mathbb{P}}$ of a given proper function F defined on a normed space $(Y, \|\cdot\|_Y)$ are defined by $F_\varepsilon = F \square (2\varepsilon)^{-1}\|\cdot\|_Y^2$ for each $\varepsilon \in \mathbb{P}$.

Theorem 3 *Let X be a Banach space which admits an equivalent locally uniformly rotund renorm such that X^* is also locally uniformly rotund. Suppose $f: X \to \mathbb{R} \cup \{+\infty\}$ to be proper, lower semicontinuous, and quadratically minorized. Put $F = (f^\circ)^*$ and, for each $\varepsilon \in \mathbb{P}$,*

$$f^\varepsilon(x) = F_\varepsilon(x, -\tfrac{1}{2}\|x\|^2) \textit{ for all } x \in X,$$

where F_ε denotes the Moreau–Yosida approximant of index ε of F:

$$F_\varepsilon(x, \alpha) = \inf_{(y, \beta) \in X \times \mathbb{R}} \left[F(y, \beta) + \frac{1}{2\varepsilon}\left(\|x - y\|^2 + (\alpha - \beta)^2\right)\right] \textit{ for all } (x, \alpha) \in X \times \mathbb{R}.$$

Then f^ε is continuously differentiable for each $\varepsilon \in \mathbb{P}$. Moreover, at each $x \in X$, $f^\varepsilon(x)$ increases to $f(x)$ as ε decreases to zero.

PROOF. Consider $Y = X \times \mathbb{R}$ endowed with the norm $y = (x, \alpha) \mapsto (\|x\|^2 + \alpha^2)^{1/2}$. This makes both Y and Y^* locally uniformly rotund Banach spaces. Then, according to Corollary 4, F_ε is Fréchet differentiable. Since $F_\varepsilon \in C^1(X \times \mathbb{R}; \mathbb{R})$ and $x \mapsto (x, -\frac{1}{2}\|x\|^2)$ is a C^1-mapping $X \to X \times \mathbb{R}$ it follows that $f^\varepsilon \in C^1(X; \mathbb{R})$. The pointwise convergence of $\langle F_\varepsilon \rangle$ to F is well established, see for instance [1]. Hence, at each $x \in X$, according to Lemma 3, it yields that

$$\lim_{\varepsilon \to 0} f^\varepsilon(x) = \lim_{\varepsilon \to 0} F_\varepsilon(x, -\tfrac{1}{2}\|x\|^2) = F(x, -\tfrac{1}{2}\|x\|^2) = f(x),$$

which ends the proof. □

REMARK. The assumptions on X are fulfilled for instance if X is reflexive or X^* is separable or $X = c_0(S)$ for some index set S.

5 A concluding example and result

In order to further motivate our study of the operation of infimal convolution we present an example of a fundamental physical situation involving the operation.

EXAMPLE. Let us consider a mechanical system consisting of two springs connected in series. If the system is displaced x length units, then the two springs are elongated $\bar{y}$ and $\bar{z}$ length units, respectively, in such a way that $\bar{y} + \bar{z} = x$ and the total potential energy of the springs is minimized. Study the figure below.

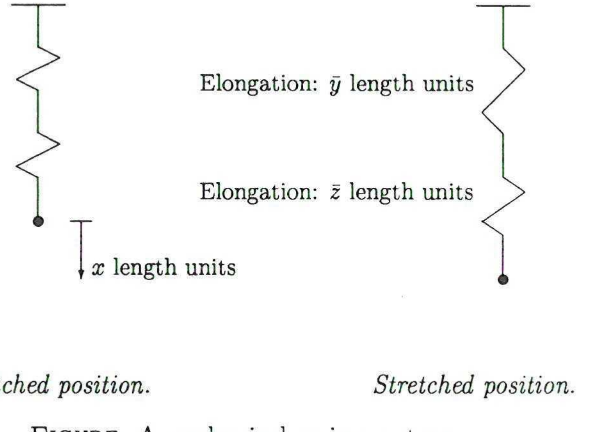


FIGURE. A mechanical spring system.

Let f and g denote the potential energy functions of the two springs so that $f(y)$ (respectively, $g(z)$) is equal to the potential energy of the first (respectively, second) spring due to a elongation y (respectively z) length units. By the energy priciple, the total potential energy $h(x)$ of the system is given as the value of the following extremal problem:

$$\text{Minimize } f(y) + g(z) \text{ subject to } y + z = x.$$

In other words, $h = f \square g$. For a continued discussion of this example, see [24].

We close this paper with a characterization of the differentiability of $f \square g$ for the case where $f, g\colon \mathbb{R} \to \mathbb{R}$ are convex and coercive functions. Recall that f is called *coercive* if $f(x)/|x|$ goes to $+\infty$ as $|x| \to +\infty$. Below f'_+ stands for the right-hand derivative of f.

Proposition 3 *Let $f, g\colon \mathbb{R} \to \mathbb{R}$ be convex and coercive. Then $f \square g$ is convex, coercive and*

$$f \square g \text{ is differentiable} \Leftrightarrow f'_+(\mathbb{R}) \cup g'_+(\mathbb{R}) = \mathbb{R}.$$

In particular, this result shows that $f \square g$ can be differentiable though neither f nor g is differentiable at each point. A proof of the proposition may be constructed using the fact that $\partial(f \square g) = \partial f // \partial g$ and that $f \square g$ is differentiable if and only if $\partial(f \square g)(x)$ is a singleton for each $x \in X$. Compare with [23, Proposition 4, Example 5] and [24, Theorem 1].

ACKNOWLEDGEMENT. The author thanks H. Attouch (Montpellier, France) and L. E. Persson (Luleå, Sweden) for some generous advice concerning questions connected to this paper.

References

[1] H. ATTOUCH AND D. AZÉ, Approximation and regularization of arbitrary functions in Hilbert spaces by the Lasry–Lions method, *Ann. Inst. H. Poincaré Anal. Non Linéaire*, to appear.

[2] H. ATTOUCH AND H. BREZIS, Duality for the sum of convex functions in general Banach spaces, *in* "Aspects of Mathematics and its Applications" (J. A. Barroso, Ed.) 125–133, North Holland, Amsterdam, 1986.

[3] H. ATTOUCH, J. L. NDOUTOUME AND M. THERA, Epigraphical convergence of functions and convergence of their derivatives in Banach spaces, *Sém. Anal. Convexe Montpellier* **20** (1990), 9.1–9.45.

[4] H. ATTOUCH AND R. J.-B. WETS, Epigraphical analysis, *Ann. Inst. H. Poincaré Anal. Non Linéaire* **6** (1989), 73–100.

[5] J. BOMAN, The sum of two plane convex C^∞ sets is not always C^5, *Math. Scand.* **66** (1990), 216–224.

[6] J. BOMAN, Smoothness of sums of convex sets with real analytic boundaries, *Math. Scand.* **66** (1990), 225–230.

[7] J. M. Borwein and A. S. Lewis, Partially finite convex programming, Part I: Quasi relative interiors and duality theory, *Math. Programming* **57** (1992), 15–48.

[8] J. M. Borwein and A. S. Lewis, Partially finite convex programming, Part II: Explicit lattice models, *Math. Programming* **57** (1992), 49–83.

[9] M. Bougeard, J.-P. Penot and A. Pommellet, Towards minimal assumptions for the infimal convolution regularization, *J. Approx. Theory* **64** (1991), 245–270.

[10] M. M. Day, "Normed Linear Spaces," Third Edition, Springer-Verlag, 1973.

[11] I. Ekeland and R. Temam, "Convex Analysis and Variational Problems," North-Holland, Amsterdam, 1976.

[12] J.-B. Hiriart-Urruty, Extension of Lipschitz functions, *J. Math. Anal. Appl.* **77** (1980), 539–544.

[13] A. D. Ioffe and V. M. Tihomirov, "Theory of Extremal Problems," North-Holland, Amsterdam, 1979.

[14] C. O. Kiselman, How smooth is the shadow of a smooth convex body?, *J. London Math. Soc.* **33** (1986), 101–109.

[15] C. O. Kiselman, Smoothness of vector sums of plane convex sets, *Math. Scand.* **60** (1987), 239–252.

[16] P.-J. Laurent, "Approximation et optimisation," Hermann, Paris, 1972.

[17] D. McLaughlin, R. Poliquin, J. Vanderwerff and V. Zizler, Second-order Gâteaux differentiable bump functions and approximations in Banach spaces, *Can. J. Math.* **45** (1993), 612–625.

[18] G. B. Passty, The parallel sum of nonlinear monotone operators, *Nonl. Anal. T. M. A.* **10** (1986), 215–227.

[19] R. Poliquin, J. Vanderwerff and V. Zizler, Convex composite representation of lower semicontinuous functions and renorming, *C. R. Acad. Sci. Paris* **317** (1993), 545–549.

[20] R. T. Rockafellar, Extension of Fenchel's duality theorem, *Duke Math. J.* **33** (1966), 81–89.

[21] H. H. Schaefer, "Topological Vector Spaces," Springer-Verlag, New York, 1971.

[22] S. Simons, The occasional distributivity of $\circ$ over $+_e$ and the change of variable formula for conjugate functions, *Nonl. Anal. T. M. A.* **14** (1990), 1111–1120.

[23] T. STRÖMBERG, An operation connected to a Young-type inequality, *Math. Nachr.* **159** (1992), 227–243.

[24] T. STRÖMBERG, Representation formulae for infimal convolution with applications, *in* "Analysis, Algebra, and Computers in Mathematical Research—Proceedings of the Twenty-First Nordic Congress of Mathematicians" (M. Gyllenberg, L. E. Persson, Eds.), Marcel Dekker, to appear in April 1994.

[25] T. STRÖMBERG, On the epigraphical sum of functions uniformly continuous on bounded sets, *Sém. Anal. Convexe Montpellier* **22** (1992), 19.1–19.16.

[26] J. VANDERWERFF, Smooth approximations in Banach spaces, *Proc. Amer. Math. Soc.* **115** (1992), 113–120.

[27] C. ZĂLINESCU, On uniformly convex functions, *J. Math. Anal. Appl.* **95** (1983), 344–374.

HÖGSKOLAN I LULEÅ
LULEÅ UNIVERSITY, SWEDEN

ISBN 1994:139D
ISSN 0348-8373
ISRN HLU-TH-T--139-D--SE

Institution/Department: Matematik

Upplaga /Number of copies:

Avdelning /Division: Matematik

Datum /Date: 1994-01-20

Titel/Title

A Study of the Operation of Infimal Convolution

Författare/Author(s)

Thomas Strömberg

Uppdragsgivare/Commissioned by

Typ/Type

- [x] Doktorsavhandling/Ph.D. thesis
- [ ] Licentiatuppsats/Licentiate thesis
- [ ] Forskningsrapport/Resarch report
- [ ] Teknisk rapport/Technical report
- [ ] Examensarbete/Final project report
- [ ] Övrig rapport/Other report

Språk/Language

- [ ] Svenska/Swedish
- [x] Engelska/English
- [ ] ______

Sammanfattning, högst 150 ord/Abstract, max 150 words

Nyckelord, högst 8/Keywords, max 8

Infimal Convolution, epigraphical addition, convex function, regularization, Inequalities, Moreau-Yosida approximation, variational convergence

Underskrift av granskare/handledare / Signature of examiner/supervisor

Lars Erik Persson