UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA MATEMÁTICA

SUBDIFFERENTIAL CALCULUS IN THE FRAMEWORK OF EPI-POINTED VARIATIONAL ANALYSIS, INTEGRAL FUNCTIONS, AND APPLICATIONS

TESIS PARA OPTAR AL GRADO DE DOCTOR EN CIENCIAS DE LA INGENIERÍA,
MENCIÓN MODELACIÓN MATEMÁTICA

PEDRO ANTONIO PÉREZ AROS

PROFESOR GUÍA:
RAFAEL CORREA FONTECILLA
PROFESOR CO-GUÍA:
ABDERRAHIM HANTOUTE

MIEMBROS DE LA COMISIÓN:
RÉNE HENRION
MARCO LÓPEZ
BORIS MORDUKHOVICH
HÉCTOR RAMÍREZ CABRERA

Este trabajo ha sido parcialmente financiado por CONICYT-PCHA/doctorado Nacional/
2014-21140621 y CMM- Conicyt PIA AFB170001

SANTIAGO DE CHILE
ENERO 2018

RESUMEN DE LA MEMORIA PARA OPTAR
AL TÍTULO DE DOCTOR EN CIENCIAS DE LA INGENIERÍA, MENCIÓN MODELACIÓN MATEMÁTICA
POR: PEDRO ANTONIO PÉREZ AROS
FECHA: ENERO 2018
PROFESOR GUÍA: RAFAEL CORREA FONTECILLA
PROFESOR CO-GUÍA: ABDERRAHIM HANTOUTE

SUBDIFFERENTIAL CALCULUS IN THE FRAMEWORK OF EPI-POINTED VARIATIONAL ANALYSIS, INTEGRAL FUNCTIONS, AND APPLICATIONS

La investigación de esta tesis es presentada en seis capítulos, desde el Capítulo 2 al Capítulo 7.

El capítulo 2 proporciona una demostración directa de una caracterización reciente de convexidad dada en el marco de los espacios de Banach en [J. Saint Raymond, J. Convexo no lineal Anal., 14 (2013), pp. 253-262]. Estos resultados también extienden esta caracterización a espacios localmente convexos bajo condiciones más débiles y se basa en la definición de una función epi-puntada.

El Capítulo 3 proporciona una extensión del Teorema Brøndsted-Rockafellar, y algunas de sus importantes consecuencias, a las funciones convexas semicontinuas inferiores definidas en espacios localmente convexos. Este resulado es demostrado usando un nuevo enfoque basado en un principio variacional simple, que también permite recuperar los resultados clásicos de una manera natural.

El Capítulo 4 continúa el estudio de la epi-puntadas no convexas, bajo una definición general de subdiferencial. Este trabajo proporciona una generalización del teorema del valor medio de Zagrodny. Posteriormente este resultado es aplicado a los problemas relacionados con la integración de subdiferenciales y caracterización de la convexidad en términos de la monotonicidad del subdiferencial.

El Capítulo 5 proporciona una fórmula general para ε-subdiferencial de una función integral convexa en términos de ε-subdiferenciales de la funcion integrante. Bajo condiciones de calificación, esta fórmula recupera los resultados clásicos en la literatura. Además, este trabajo investiga caracterizaciones del subdiferencial en términos de selecciones medibles que convergen al punto de interés.

El Capítulo 6 proporciona fórmulas secuenciales para subdiferenciales bornológicos de un funcional integral no convexo. También son presentadas fórmulas exactas para el subiferencial Limiting/Mordukovich, el subdiferencial Geometrico de Ioffe y el subdiferencial de Clarke-Rockafellar.

El Capítulo 7 proporciona fórmulas para el subdiferencial de funciones de probabilidad bajo distribuciones Gaussianas. En este trabajo la variables de decisión esta tomada en un espacio infinito dimensional. Estas fórmulas se basan en la descomposición esférico-radial de vectores aleatorios Gaussianos.

ABSTRACT OF THE THESIS TO GET
PhD OF ENGINEERING SCIENCE, WITH EMPHASIS IN MATHEMATICAL MODELING
BY: PEDRO ANTONIO PÉREZ AROS
PH.D. ADVISOR: RAFAEL CORREA FONTECILLA
PH.D. CO-ADVISOR: ABDERRAHIM HANTOUTE

SUBDIFFERENTIAL CALCULUS IN THE FRAMEWORK OF EPI-POINTED VARIATIONAL ANALYSIS, INTEGRAL FUNCTIONS, AND APPLICATIONS

The research of this thesis is given in six chapters, from Chapter 2 to Chapter 7.

Chapter 2 provides a direct proof of a recent characterization of convexity given in the setting of Banach spaces in [J. Saint Raymond, J. Nonlinear Convex Anal., 14 (2013), pp. 253–262]. Our results also extend this characterization to locally convex spaces under weaker conditions and is based in the definition of an epi-pointed function.

Chapter 3 gives an extension of the Brøndsted-Rockafellar Theorem, and some of its important consequences, to proper convex lower-semicontinuous epi-pointed functions defined in locally convex spaces. We use a new approach based on a simple variational principle, which also allows recovering the classical results in a natural way.

Chapter 4 continues the study the subdifferential of nonconvex epi-pointed functions, under a general definition of the subdifferential. This work provides a generalization of Zagrodny's Mean Value Theorem, and gives several applications to problems related to integration of the subdifferential and the characterization of convexity in terms of the monotonicity of the subdifferential.

Chapter 5 provides a general formula for the ε-subdifferential of a convex integral functional in terms of the ε-subdifferential of the data functions. Under classical qualification conditions, we recover classical results in the literature. Also, this work investigates exact rules to characterize the subdifferential of the integral functional at a given point x in terms of measurable selections in the subdifferential of the data, which converge to the point x. These formulas generalize some results of [A. D. Ioffe, J. Convex Anal., 13(3-4) (2006, pp 759-772] and [O. Lopez and L. Thibault, J. Nonlinear Convex Anal., 9(2) (2008),pp 295-308].

Chapter 6 gives sequential formulae for the bornological subdifferential of a non-convex integral functional and also exact formulae for the Limiting/Mordukovich subdifferential, the G-subdifferential of Ioffe and the generalized Clarke-Rockafellar subdifferential.

Chapter 7 provides subdifferential formulae of probabilistic functionals in the case of Gaussian distributions for possibly infinite-dimensional decision variables and nonsmooth (locally Lipschitzian) input data. On the other hand these formulae are based on the spheric-radial decomposition of Gaussian random vectors and on the other hand on a cone of directions of moderate growth. By successively adding additional hypotheses, conditions are satisfied under which the probability function is locally Lipschitzian or even differentiable.

This thesis is dedicated to my family.

Acknowledgements

I would like to thank...

A mis padres, por todo su apoyo, cariño y amor entregados durante mis 27 años de vida.

A mi amada Carla, quien me ha acompañado y entregado su apoyo en cada momento, siempre motivándome a ser mejor, pese a que el tiempo libre de un matemático es siempre escaso.

A toda mi familia, especialmente a mis tías María, Rita y mi tío Victor quienes me acogieron en su hogar durante mis últimos diez años de estudio.

A toda mi familia mi abuela, tíos, tías, primos, primas sobrinos y demás integrantes en general quienes siempre mostraron gran apoyo y orgullo hacia mi desarrollo como estudiante y mi intensión de continuar toda mi vida estudiando matemáticas.

A David Salas y Emilio Vilches con quienes he compartido ideas y proyectos y cuyo formidable trabajo ha sido fuente de admiración y motivación para mi tesis.

A todos los funcionarios y académicos del departamento de ingeniería matemática y del centro de modelamiento matemático, quienes con su amabilidad y cordialidad hicieron de este proceso uno de los más agradables de mi vida, haciéndome sentir muchas veces no en un establecimiento educacional y de investigación, sino más bien dentro de una familia. De manera muy especial agradezco a María Antonieta Serei quien siempre me aconsejó con sus amables palabras y solucionó incontables problemas administrativos.

To my advisor, Professor Rafael Correa who always motivates and encourages me to continue my studies, even from the beginning of my engineering degree. I do not have enough words of thanks to show him my gratitude.

To my advisor, Professor Abderrahim Hantoute who shares many ideas and works during the last four years, always with his words of encouragement and confidence about my research.

To Professor Réne Henrion who introduced me to the fantastic theory of stochastic optimization and probabilistic constraint and for his help during my internship in the Weierstrass Institute and also for helping me to understand the importance of showing my research to the scientific world.

To Professor Boris Mordukhovich for his words and motivation about my word and also for his time and explanations about some topics of his research.

To Professor Marco Lopez for his help during my last state in Alicante and specially for his time, ideas and feedback about my research.

To my thesis committee members for all of their guidance through this process; your discussion, ideas, and feedback have been absolutely invaluable.

To CONICYT, University of Chile, Department of Mathematical Engineering (DIM), Center for Mathematical Modeling (CMM) and University of O'higgins for their grants and financial support to develop my research.

To all the people who have contributed somehow to finish my PhD thesis.

Contents

Introduction

The aim of this manuscript is to present the results of the research carried out during my PhD work, and relying on the *subdifferential theory and variational analysis*.

In this initial part of the manuscript I am going to summarize each of the six works that make up this thesis, given in six chapters, from chapters 2 through 7.

For easy of reading I have preferred to keep in this introduction the numbering of the theorems, corollaries, examples,..., as they appear in each chapter.

On the Klee-Saint Raymond's Characterization of Convexity

The Chapter 2, which corresponds to my first work, concerns a characterization of convexity using variational properties. It is common in learning the classic optimization theory to observe that if f is a nonconvex continuous function on the real line $\mathbb{R}$, which satisfies the coercivity property

$$\lim_{|x|\to\infty} \frac{f(x)}{|x|} = +\infty,$$

then one can find an affine function h such that $f(x) \geq h(x)$ for all $x \in \mathbb{R}$, and the set of points where the function $f - h$ vanishes is a nonconvex set (see Figure 1). Therefore, a natural question to ask is whether this fact could characterize convex functions on an infinite-dimensional Banach space X.

The first author to consider this observation was J. Saint Raymond in [108], in answer to the question proposed by B. Ricceri in private communications about characterizing convex functions defined on a reflexive Banach space. The question was the following: *If X is a reflexive Banach space and $f : X \to \mathbb{R}$ is a weakly lower semicontinuous function such that for all continuous linear functional $x^* \in X^*$, the function $f - x^*$ attains its global minimum at a single point, then, must f be convex?* He also inquired: *If $f : X \to \mathbb{R}$ is a C^1 function satisfying $\lim_{\|x\|\to\infty} \frac{f(x)}{\|x\|} = +\infty$, then must the function be convex?*

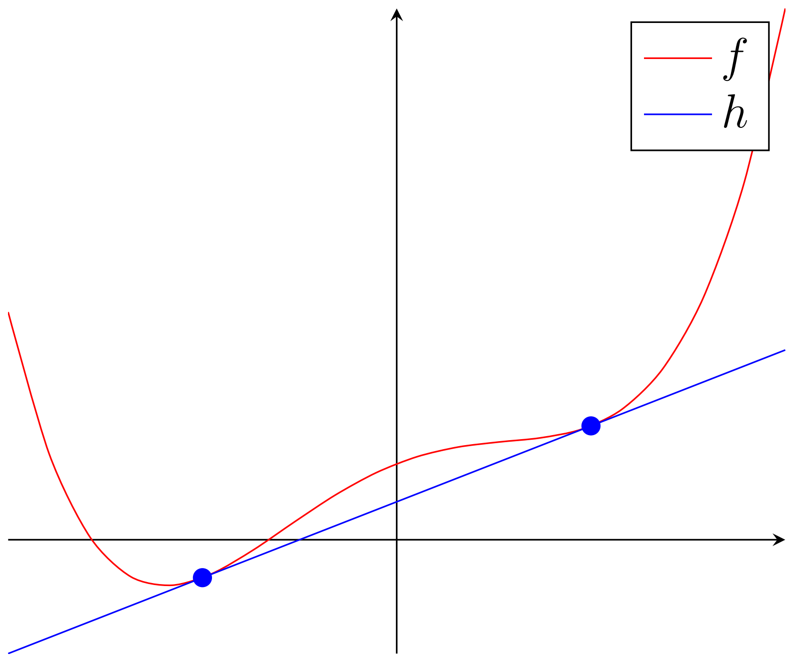

Figure 1: argmin of the function $f - h$.

The answers can be seen in [108, Theorem 3.3 and Corollary 3.4]. See also Ricceri's answer [97, Corollary 14]. Both hypotheses on the *growth of the function* presented in these works, the coercivity of f and the non-emptiness of the argmin of $f - x^*$ for all $x^* \in X^*$, appear to be different, but Saint Raymond also proved that both assumptions are equivalent in reflexive Banach spaces (see [108, Theorem 2.3.]). Later, the same author gave the following generalization of his result, which can be found in [107, Theorem 10].

Theorem A *Let X be a Banach space and let $f : X \to \overline{\mathbb{R}}$ be an weakly lsc proper function such that for every $x^* \in X^*$, $f - x^*$ attains its minimum. If the set where the function $f - x^*$ attains its minimum is convex for all x^* in a convex dense set, then necessarily f is convex.*

Although the mathematical statement emerges as an interesting result of convex analysis, the author used deep tools of functional analysis on Banach spaces, like James' Theorem and Brouwer's Fixed-Point Theorem for multifunctions, among many others. Also Ricceri's proof involves techniques of operator theory, all of which are far away, or not directly related to convex analysis.

Making a scrutiny of the hypotheses of the above result in terms of convex analysis one can notice that the assumption $f - x^*$ *attains its minimum for* $x^* \in X^*$, implies that the conjugate of f is finite on X^*. Moreover in [108, Theorem 2.4] the author present a result being the functional counterpart of James Theorem (see e.g. [46]), and this result prove that the above assumption also implies the inf-compactness of $f - x^*$ for all x^*, which in terms of convex analysis, corresponds the continuity of f^* on X^* with respect to the Mackey topology $\tau(X^*, X)$, which corresponds to the class of *Epi-pointed functions* (see Definition 1.7).

The aim of Chapter 2 is to use techniques of convex analysis to give a direct proof of a generalization of Theorem A for functions defined on locally convex spaces. This generalization, given in Corollary 2.8, is an immediate consequence of the main result of this work which provides an explicit expression of the closed convex hull of a function; see Theorem 2.5. Our hypotheses are weaker than those used in Theorem A, and rely on a property of compactness of functions, which have been called, by many authors, *epi-pointedness property.*

This chapter together with include the Saint Raymond's characterization of convexity in the theory of convex analysis, also introduce us the class of epi-pointed functions, which motives the study of chapters 3 and also 4.

The main result of this chapter is the following theorem.

Theorem 2.5 *If $f : X \to \overline{\mathbb{R}}$ is a weakly lsc epi-pointed function such that*

$$\operatorname{argmin}\{f - x^*\} \textit{ is convex for all } x^* \in D,$$

where D is a convex dense subset of $\operatorname{dom} f^*$. *Then we have that*

$$f^{**} = \sigma_{\operatorname{dom} f^*} \square f. \tag{1}$$

Here $\square$ denotes the inf-convolution between functions.

The above equation gives us the convexity of f under the additional assumption that $\operatorname{dom} f^*$ is dense, which in particular occurs under the hypotheses of Saint Raymond's Theorem.

To give a geometrical regard to this result in terms of epigraphs. Let us recall that when the inf-convolution is exact, then its epigraph is the sum of the epigraphs of the two functions. Since convolution in the equality (1) is exact, we see that Theorem 2.5 corresponds to the set equality

$$\overline{\mathrm{co}}(\operatorname{epi} f) = \operatorname{epi} f + \operatorname{epi} \sigma_{\operatorname{dom} f^*}. \tag{2}$$

Another interesting formulation can be given when we consider the asymptotic cone of $\operatorname{epi} f$ defined by (see for example [40])

$$(\operatorname{epi} f)_\infty := \bigcap_{\varepsilon > 0} \overline{]0, \varepsilon] \operatorname{epi} f},$$

which is the epigraph of the asymptotic function f^∞, and using the equality $\overline{\mathrm{co}}(f^\infty) = \sigma_{\operatorname{dom} f^*}$ when f is lsc and epi-pointed (see [28, Theorem 7]), we can rewrite (1) as

$$f^{**} = f \square \overline{\mathrm{co}}(f^\infty),$$

and (2) as

$$\overline{\mathrm{co}}(\operatorname{epi} f) = \overline{\mathrm{co}}((\operatorname{epi} f)_\infty) + \operatorname{epi} f.$$

It is worth noting that this characterization does not involve dual objects.

In the final part of this chapter and keeping in mind a possible application to finite-dimensional spaces, we consider the *relative interior* of the domain of the conjugate of f (see Definition 1.1), and we obtain:

Theorem 2.12 *Let $f : X \to \overline{\mathbb{R}}$ be a function with a proper conjugate, and denote $F := \overline{\mathrm{aff}}(\mathrm{dom}\, f^*)$. We suppose that the following conditions hold:*

(a) *The restriction of f^* to F, ${f^*}_{|F}$, is continuous on* $\mathrm{ri}(\mathrm{dom}\, f^*)$.

(b) *There exists $x_0^* \in \mathrm{ri}(\mathrm{dom}\, f^*)$ such that $f - x_0^*$ is weakly lsc and weakly inf-compact.*

(c) *There exists a convex set $D \subseteq \mathrm{ri}(\mathrm{dom}\, f^*)$, with $\mathrm{ri}(\mathrm{dom}\, f^*) \subset \overline{D}$, such that for all $x^* \in D$, $\mathrm{argmin}\{f - x^*\}$ is convex.*

Then we have

$$\sigma_{\mathrm{dom}\, f^*} \Box f = f^{**}.$$

Moreover, if $\overline{D} = F$, then

$$\sigma_{\mathrm{dom}\, f^*} \Box f = f^{**} = f_F,$$

where $f_F(x) := \inf\{f(w) : w \in x + (F - x_0^*)^{\perp}\}$.

Also in this chapter we give three interesting examples: The first one shows that equality (1) cannot be improved by the equality $f = f^{**}$ with the assumption of Theorem 2.5 (see Example 2.6). The second one shows the necessity of the convexity assumption of D in Corollary 2.8 (see Example 2.10). The final example shows that it is not possible to get the equality $f = f^{**}$ instead of $f_{F-x_0^*} = f^{**}$ in Theorem 2.12 (see Remark 2.13).

On Brøndsted-Rockafellar's Theorem for convex lower semi-continuous epi-pointed functions in locally convex spaces

The previous chapter opens our minds to the class of epi-ponted functions, which appear to be promising for developing variational analysis outside of Banach spaces.

It is known that Brøndsted-Rockafellar Theorem is not valid outside Banach spaces for all lsc proper convex functions (see [19]), more precisely Brøndsted and Rockafellar found an lsc proper convex function defined on a locally convex topological vector space with empty subdifferential everywhere, showing that many classical statements of convex analysis are not valid outside the framework of Banach spaces, for instance Brøonsted-Rockafellar Theorem, Maximal monotonicity of the subdifferential and Fuzzy calculus rules. This observation motivates the work to provide a suitable family of lsc proper convex functions defined on a locally convex space, which satisfies these theorems.

The main features of Chapter 3 are:

(1) To show that epi-pointed lsc convex functions, defined on any locally convex space X, satisfy the Brøndsted-Rockafellar theorem.

Theorem 3.8 *Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a convex, lsc and epi-pointed function. Consider $\varepsilon \geq 0$, $\beta \in [0, \infty)$, a continuous seminorm p, $\lambda > 0$ and $x_0 \in X$. If*

$x_0^* \in \partial_\varepsilon f(x_0) \cap \operatorname{int}(\operatorname{dom} f^*)$, *then there are* $x_\varepsilon \in X$, $y_\varepsilon^* \in \mathbb{B}_p(0,1)^\circ$ *and* $\lambda_\varepsilon \in [-1,1]$ *such that:*

(a) $p(x_0 - x_\varepsilon) + \beta|\langle x_0^*, x_0 - x_\varepsilon\rangle| \leq \lambda,$

(b) $x_\varepsilon^* := x_0^* + \frac{\varepsilon}{\lambda}(y_\varepsilon^* + \beta\lambda_\varepsilon x_0^*) \in \partial f(x_\varepsilon),$

(c) $|\langle x_\varepsilon^*, x_0 - x_\varepsilon\rangle| \leq \varepsilon + \frac{\lambda}{\beta},$

(d) $|f(x_0) - f(x_\varepsilon)| \leq \varepsilon + \frac{\lambda}{\beta}$

(e) $x_\varepsilon^* \in \partial_{2\varepsilon} f(x_0).$

Corollary 3.9 *Let* $f : X \to \mathbb{R} \cup \{+\infty\}$ *be a convex lsc epi-pointed function. Then for every* $x \in \operatorname{dom} f$ *there exist nets* $\{(x_\alpha),\ (x_\alpha^*)\}_{\alpha\in\mathbb{A}}$ *such that* $x_\alpha^* \in \partial f(x_\alpha)$, $x_\alpha \to x$ *and* $f(x_\alpha) \to f(x)$.

(2) To provide a different proof of Brøndsted-Rockafellar's Theorem for this class of epi-pointed functions, in the sense that it is based on a very simple variational principle (Lemma 3.7), which is valid in any locally convex space X, without requiring such tools as Ekeland's or Bishop-Phelps' variational principles.

Lemma 3.7 *Let* $x_0 \in X$ *and let* $f : X \to \mathbb{R} \cup \{+\infty\}$ *and* $\rho : X \to [0,\infty)$ *be two convex lsc functions such that* $\rho(0) = 0$ *and the function* $f(\cdot) + \rho(\cdot - x_0)$ *is epi-pointed. For any* $\varepsilon \geq 0$, $x_0 \in X$ *and* $x_0^* \in \partial_\varepsilon f(x_0)$ *with* $x_0^* \in \operatorname{int}(\operatorname{dom}(f + \rho(\cdot - x_0))^*)$, *there exists* $x_\varepsilon \in X$ *such that:*

(a) $\rho(x_0 - x_\varepsilon) \leq \varepsilon,$

(b) $x_0^* \in \partial(f + \rho(\cdot - x_0))(x_\varepsilon),$

(c) $|f(x_0) - f(x_\varepsilon)| \leq |\langle x_0^*, x_0 - x_\varepsilon\rangle| + \varepsilon.$

(3) Since every convex function in Banach spaces can be adequately perturbed to obtain an epi-pointed function, we can recover in the Banach setting the usual Brøndsted-Rockafellar theorem (see Section 3.3).

(4) We also obtain other important results in the same spirit, Theorem 3.10 for the Maximal monotonicity of the subdifferential, Theorems 3.12 and 3.14 for the subdifferential limiting calculus rules for functions defined in locally convex spaces.

Theorem *3.10 Let* $f : X \to \mathbb{R} \cup \{+\infty\}$ *be a convex lsc function. If either* f *or* f^* *is epi-pointed, then* ∂f *and* ∂f^* *are maximal monotone operators.*

Theorem *3.12 Let* X, Z *be two lcs,* $A \in \mathcal{L}(X,Z)$, $g \in \Gamma_0(Z)$ *be an epi-pointed function,* $f := g \circ A$ *and* $x \in \operatorname{dom} f$. *Then* $x^* \in \partial f(x)$ *if and only if there exists a net* $(z_{\mathrm{i}}, z_{\mathrm{i}}^*)_{\mathrm{i}\in I} \in Z \times Z^*$ *such that* $z_{\mathrm{i}}^* \in \partial g(z_{\mathrm{i}})$, $z_{\mathrm{i}} \to y = Ax$, $g(z_{\mathrm{i}}) \to g(y)$, $\langle z_{\mathrm{i}} - z, z_{\mathrm{i}}^*\rangle \to 0$ *and* $A^*(z_{\mathrm{i}}^*) \xrightarrow{w^*} x^*$.

Theorem *3.14* *Let $f_1, f_2 \in \Gamma_0(X)$ be two epi-ponted functions and $x \in \mathrm{dom}(f_1 + f_2)$. Then $x^* \in \partial (f_1 + f_2)(x)$ if and only if there exist two nets $(x_{k,\mathrm{i}}, x^*_{k,\mathrm{i}})_{\mathrm{i} \in I} \subset X \times X^*$ such that $x^*_{k,\mathrm{i}} \in \partial f_k(x_{k,\mathrm{i}})$ $k = 1, 2$, $x_{k,\mathrm{i}} \to x$, $f_k(x_{k,\mathrm{i}}) \to f_k(x)$, $\langle x_{k,\mathrm{i}} - x, x^*_{k,\mathrm{i}} \rangle \to 0$, for $k = 1, 2$, and $(x^*_{1,\mathrm{i}} + x^*_{2,\mathrm{i}}) \xrightarrow{w^*} x^*$.*

Extensions of Zagrodny's Mean Value Theorem, integration theorems and characterization the of convexity for a generalized subdifferential of functions defined in locally convex spaces

After finishing the previous work we face the nonconvex case and investigate the behavior of epi-pointed functions under some notions of generalized differentiation.

As a result of the discovery of variational principles in Banach spaces and examples which show the vacuity of the subdifferential of an lsc convex function, the theory of subdifferential has centred its attention on the setting of Banach spaces with appropriate smoothness properties for each subdifferential. The main purpose of our work in this chapter was to explore *subdifferential-techniques* outside the framework of Banach spaces. We are forewarned that this task is impossible for all lsc functions, therefore our efforts are concentrated in two frameworks: First in the class of epi-pointed functions in arbitrary locally convex spaces, and second for the class of lsc functions defined in a special class of locally convex spaces.

With this purpose we begin by recalling some definitions of subdifferentials available in the literature, that we classify into two groups. A subdifferential which is defined at a given point and it does not take into account variational properties of a function in its neighborhood. Usually, such subdifferentials come from some classical notions of differentiability; for example: *Fréchet, Gâteaux, Dini, Hadamard*, etc. The other group incorporates differential properties of a function near a given point. Usually, these subdifferentials can be represented as (some kinds of) limits of simple previous ones. Examples of these subdifferentials are the famous *generalized gradient of Clarke-Rockafellar, the Limiting/Mordukhovich subdifferential, the approximate subdifferential and the G-Subdifferential of Ioffe.*

Among the mentioned nonconvex subdifferentials, two ideas motivated the notion of subdifferential adopted in Chapter 4 (see Definition 4.1). The first corresponds to the idea of using the concept of the ε-enlargements of the Fréchet subdifferential (see also [62] for the ε-enlargement of the Dini-Hadamard subdifferential). This notion allowed Mordukhovich to define his subdifferential in arbitrary Banach spaces and establish many suitable calculus rules, see for example the book [84, Chapter I]. Following the mentioned book we recall the Limiting/Mordukhovich and the singular Limiting/Mordukhovich subdifferentials for a function f at a point $x \in \operatorname{dom} f$ as the sets:

$$\begin{aligned} \partial_M f(x) &:= \limsup_{\substack{x_n \xrightarrow{f} x \\ \varepsilon_n \downarrow 0}} \partial_F^{\varepsilon} f(x_n), \\ \partial_M^{\infty} f(x) &:= \limsup_{\substack{x_n \xrightarrow{f} x \\ \lambda_n, \varepsilon_n \downarrow 0}} \lambda_n \, \partial_F^{\varepsilon} f(x_n), \end{aligned} \tag{3}$$

respectively, where the limits correspond to the sequential *Painlevé-Kuratowski upper limits* taken with respect to the norm in X and the weak* topology in X^*. Here $\partial_F^{\varepsilon} f$ denotes the ε-enlargements of the Fréchet subdifferential, which is defined as:

$$\partial_F^{\varepsilon} f(x) = \{x^* \mid \liminf_{h \to 0} \frac{f(x+h) - f(x) - \langle x^*, h \rangle}{\|h\|} \geq -\varepsilon\}. \tag{4}$$

The second idea, which motivates our definition (see Definition 4.1), was the ideas given by Ioffe, with his construction of the approximate subdifferential and G-subdifferential (see for example [65]). In [61] the author considers an extension to every locally convex space, using the family $\mathcal{F}$ of all finite dimensional subspaces of X. Then Ioffe introduced the A-subdifferential and the singular A-subdifferential of f at x as the sets

$$\begin{aligned} \partial_A f(x) &:= \bigcap_{L \in \mathcal{F}} \limsup_{x' \xrightarrow{f} x} \partial_D^- f_{x'+L}(x'), \\ \partial_A^{\infty} f(x) &:= \bigcap_{L \in \mathcal{F}} \limsup_{\substack{x' \xrightarrow{f} x \\ \lambda \downarrow 0}} \lambda_n \, \partial_D^- f_{x'+L}(x'), \end{aligned} \tag{5}$$

respectively. Here $f_{x'+L}$ denotes the function $f + \delta_{x'+L}$ and ∂_D^- means the Dini-Hadamard subdifferential, given by

$$\partial^- f_D f(x) = \{x^* \mid \langle x^*, h \rangle \leq \mathrm{d}^- f(x; h), \text{ for all } h \in X\},$$

where $\mathrm{d}^- f(x;h)$ is the (lower) Dini-directional derivative of f at x

$$\mathrm{d}^- f(x;h) := \liminf_{\substack{t \to 0^+ \\ u \to h}} \frac{f(x+tu) - f(x)}{t}.$$

The Limiting/Mordukhovich and the Approximate subdifferentials motivate us to start considering a family of subdifferentials instead of only one. Then we introduce Definition 4.1 as our starting point. In this definition we consider a family of set-valued operators $\partial_{L,\mathrm{i}} : X \rightrightarrows X^*$ with $(L, \mathrm{i}) \in \mathcal{L} \times \mathbb{I}$, where $\mathcal{L}$ is a family of linear subspaces which covers X and $(\mathbb{I}, \preceq)$ is a directed set. The family $\partial_{F,\mathrm{i}}$ is allowed to satisfy a certain approximate Fermat rule.

The reader can notice that our motivation to consider the family $\mathcal{L}$ comes from (5), where the subdifferential is indexed by the family of finite dimensional subspaces. In [61, 62] the author proved that this family can be changed by a smaller one with certain nice properties for the subdifferential. Our motivation to consider the order set $(\mathbb{I}, \preceq)$ comes from the Limiting/Mordukhovich subdifferential and its ε-enlargements (see (3) and (4)).

With the previous definition in mind we propose to prove an extension of *Zagrodny's Mean Value Theorem*. Given that one of the most important (among many others of course) and more general tools in nonconvex subdifferential theory is Zagrodny's Mean Value Theorem (see [130] and see [114] for an extension to a more general class of subdifferentials) we extend this result using Definition 4.1. Our result establishes the following.

Theorem 4.9 *Consider a family of subdifferentials $\{\hat{\partial}_{\mathrm{i},L} : \mathrm{i} \in \mathbb{I},\ L \in \mathcal{L}\}$ for a given proper lsc function f. Assume that one of the following conditions holds:*

(a) The topology on X is generated by a family of seminorms $(\rho_L)_{L\in\mathcal{L}}$, where for every $L \in \mathcal{L}$, (L, ρ_L) is a Banach space, and $\rho_M \leq \rho_L$ for all $M \subseteq L$.

(b) f is a w-lsc and epi-pointed function.

Then, for every $a, b \in X$ with $a \in \operatorname{dom} f$ and $a \neq b$, $r \in \mathbb{R}$ with $r \leq f(b)$ and every continuous seminorm p such that $p(a-b) \neq 0$, there exist $c \in [a, b[$ and $(x_\alpha, x_\alpha^*)_{\alpha\in\mathbb{A}} \in X \times X^*$ such that $\{(\mathrm{i}_\alpha, L_\alpha)\}_{\alpha\ \in\mathbb{A}}$ is cofinal in $\mathbb{I} \times \mathcal{L}$, $(x_\alpha)_{\alpha\in\mathbb{A}} \xrightarrow{f} c$, $x_\alpha^* \in \hat{\partial}_{\mathrm{i}_\alpha,L_\alpha} f(x_\alpha)$ with $x_\alpha \in L_\alpha$, and

(i) $r - f(a) \leq \liminf_\alpha \langle b - a, x_\alpha^* \rangle$,

(ii) $0 \leq \liminf_\alpha \langle c - x_\alpha, x_\alpha^* \rangle$,

(iii) $\frac{p(b-c)}{p(b-a)} (r - f(a)) \leq \liminf_\alpha \langle b - x_\alpha, x_\alpha^* \rangle$,

(iv) $\frac{p(b-a)}{p(c-a)} (f(c) - f(a)) \leq (r - f(a))$.

Moreover, if $c \neq a$, then one has $\lim_\alpha \langle b - a, x_\alpha^* \rangle = r - f(a)$.

It is important to discuss the two cases explored in the above result. The epi-pointed case comes from the interest of the authors in exploring the variational properties of nonconvex epi-pointed functions. The second case is not only motivated by the ideas of Ioffe and the consideration of $\mathcal{L}$ as the family of all finite-dimensional subspaces; this case also comes from the observation that many spaces for applications, which are not Banach spaces, can be obtained by union of a countable family of Banach spaces, or they have a countable family of Banach spaces which their union is dense in the space; for example:

i) The space of k-times continuously differentiable functions from an open set Ω to $\mathbb{R}$ with compact support $C_0^k(\Omega, \mathbb{R})$ with the topology generated by the uniform convergence of all its derivatives over compact sets. This space is generated by the spaces $C_0^k(K, \mathbb{R})$ with K being a compact set included in Ω.

ii) The space of all locally p-integrable (also called p-locally integrable) functions. This space has a dense subspace generated by all the integrable functions which vanish outside of a compact set.

We hope that our result could bring new ideas from variational analysis in Banach spaces to applications in locally convex spaces.

Using Theorem 4.9 we were able to extend other important results, such as the density of the domain of subdifferentials, known also as Like-Brøndsted-Rockafellar Theorems:

Corollary *4.11* *In the setting of Theorem 4.9, if* $x \in \operatorname{dom} f$*, then there exists* $(x_\beta, x_\beta^*)_{\beta\in\mathbb{U}} \subseteq \operatorname{gph} \hat{\partial}_{L_\beta,\mathrm{i}_\beta} f$ *such that* $x_\beta \xrightarrow{f} x$ *and* $\liminf\langle x - x_\beta, x_\beta^*\rangle \geq 0$*. In particular,* $\operatorname{dom} f \subseteq \limsup\limits_{L,\mathrm{i}} \operatorname{dom} \hat{\partial}_{L,\mathrm{i}} f$*.*

Also we obtain results concerning the integration of subdifferentials. The most general result is

Theorem *4.13* *In the setting of Theorem 4.9, consider a function* $g \in \Gamma(X)$ *and a continuous seminorm* ρ*. Suppose there exist* $\varepsilon \geq 0$ *and a net* $(\varepsilon_\mathrm{i})_{\mathrm{i}\in I} \downarrow \varepsilon$ *and an open convex set* U *with* $U \cap \operatorname{dom} f \neq \emptyset$ *such that*

$$\hat{\partial}_{L,\mathrm{i}} f(x) \subseteq \partial g(x) + \varepsilon_\mathrm{i}\mathbb{B}_\rho(0,1)^\circ + L^\perp, \;\; \forall x \in C, \; \forall \mathrm{i} \in \mathbb{I}, \forall L \in \mathcal{L}.$$

Then $C \cap \operatorname{dom} f = C \cap \operatorname{dom} g$ *and for all* $x \in C$*,* $y \in C \cap \operatorname{dom} f$ *one has*

$$g(x) - g(y) - \varepsilon\rho(x-y) \leq f(x) - f(y) \leq g(x) - g(y) + \varepsilon\rho(x-y). \tag{6}$$

The reader can notice that (6) implies the equality $f(x) - f(y) = g(x) - g(y)$ when $\varepsilon = 0$. More precisely, the following corollary holds.

Corollary *4.16* *In the setting of Theorem 4.13, assume that* $\varepsilon = 0$ *and* $C = X$*. Then there exists* $c \in \mathbb{R}$ *such that* $f = g + c$*.*

Theorem 4.13 also allows to prove two recent results about integration of subdifferentials. The first one is due to Thibault [81, Theorem 1.2] (see Corollary 4.17 for our result) and the second one is a particular case of a more general result due to Correa-Hantoute-Salas [32, Theorem 4.8] (see Corollary 4.19 for our result).

The final goal of this Chapter was to establish a characterization of the convexity by means of the monotonicity of the subdifferential. This result is due to Correa-Jofre-Thibault [35, Theorem 2.2 and 2.4] for an lsc function defined in a Banach space X. We extend this result for weakly lsc epi-pointed functions defined in an arbitrary locally convex space.

Theorem *4.20* *In the setting of Theorem 4.9, assume that* $\hat{\partial}$ *is a subdifferential and consider the following statements:*

(i) f *is convex.*

(ii) $\hat{\partial} f$ *is monotone.*

(iii) $\hat{\partial} f(x) \subseteq \partial f(x)$*, for all* $x \in X$*.*

Then (ii) *and* (iii) *are equivalent and each one implies* (i)*. In addition, if* $\hat{\partial} f$ *coincides with the Moreau-Rockafellar subdifferential for every convex function* f*, then the three statements are equivalent.*

A complete characterization of the subdifferential of convex integral functions

Several problems in applied mathematics such calculus of variation, control theory and stochastic programming among others concern the study of integral functionals. By this is meant an expression of the form

$$\hat{I}_f(x) := \int_T \max\{f(t,x(t)),0\}\mathrm{d}\mu(t) + \int_T \min\{f(t,x(t)),0\}\mathrm{d}\mu(t), \quad x(\cdot) \in \mathfrak{X}, \tag{7}$$

where $(T,\mathcal{A},\mu)$ is a complete σ-finite measure space and $\mathfrak{X}$ is some linear space of $\mathcal{A}$-measurable functions with values on X. The function $f : T \times \mathbb{R}^n \to \overline{\mathbb{R}}$ is the associated *integrand* (see Definition 1.5 and 1.20 for more details).

The aim of this Chapter is to bring formulas for the subdifferential of the convex integral functional $\hat{I}_f$ when the space of measurable functions is the constant functions, that is to say, $x \in X \to I_f(x) = \int_T f(t,x)\mathrm{d}\mu(t)$; for this reason we give to this functional the special notation I_f instead of $\hat{I}_f$; this particular case is also known as a continuous sum.

A well-known formula, given by Ioffe-Levin [66], shows that under certain assumptions of continuity of the integral the following formula holds for the subdifferential of I_f

$$\partial I_f(x) = \int_T \partial f(t,x)\mathrm{d}\mu(t) + N_{\mathrm{dom}_f}(x), \text{ for all } x \in \mathbb{R}^n, \tag{8}$$

where the set $\int_T \partial f(t,x)\mathrm{d}\mu(t)$ is understood in the sense of the Aumman integral, that is to say, as the set of points of the form $\int_T x(t)\mathrm{d}\mu(t)$ where x is an integrable function such that $x(t) \in \partial f(t,\cdot)(x)$ for almost all $t \in T$ (see Definition 1.23). One can compare (8) with its discrete counterpart, which states that for every two convex lsc functions f_1, f_2 such that f_1 is continuous at some point of the domain of f_2 one gets $\partial(f_1+f_2)(x) = \partial f_1(x) + \partial f_2(x)$ for all $x \in \mathbb{R}^n$. Thus, a reasonable idea is give similar formulas as the given by Hiriart-Urruty and Phelps without qualification conditions (see Proposition 1.2). Consequently it feels natural to think in a generalization of (8) as

$$\partial I_f(x) = \bigcap_{\eta>0} \mathrm{cl}\left\{ \int_T \partial_\eta f(t,x)\mathrm{d}\mu(t) + N_{\mathrm{dom}_f}(x) \right\} \text{ for all } x \in \mathbb{R}^n. \tag{9}$$

Unfortunately, such a mathematical expression does not hold without any qualification conditions; indeed, one can even show counterexamples where the set $\int_T \partial_\varepsilon f(t,x)\mathrm{d}\mu(t)$ is empty and the integrand $f_t(\cdot)$ is smooth at the point of interest.

Example ***5.12*** *Consider the function* $f(x) := \frac{b}{a}x + b + \delta_{[-\eta,\eta]}$ $a,b,\eta > 0$, *then we have* $\partial_\varepsilon f(0) = \left[-\frac{\varepsilon}{\eta} + \frac{b}{a}, \frac{\varepsilon}{\eta} + \frac{b}{a}\right]$. *Indeed,*

$$\begin{aligned} \alpha \in \partial_\varepsilon f(0) &\Leftrightarrow \alpha \cdot x \le f(x) - f(0) + \varepsilon \;\; \forall x \in [-\eta,\eta] \\ &\Leftrightarrow \alpha \cdot x \le \frac{b}{a}x + b - b + \varepsilon \;\; \forall x \in [-\eta,\eta] \\ &\Leftrightarrow (\alpha - \frac{b}{a}) \cdot x \le \varepsilon \;\; \forall x \in [-\eta,\eta]. \end{aligned}$$

Then, it is clear that the last inequality holds if and only if $\alpha \in \left[-\frac{\varepsilon}{\eta} + \frac{b}{a}, \frac{\varepsilon}{\eta} + \frac{b}{a}\right]$. *Using the previous function one constructs a normal integrand* $f :]0,1] \times \mathbb{R} \to [0, +\infty]$ *by* $f(t,x) := \frac{b(t)}{a(t)} x + b(t) + I_{[-\eta(t), \eta(t)]}$, *with* $b(t) = \frac{1}{\sqrt{t}} + 1$, $a(t) = \delta(t) = t$. *Hence we compute*

$$I_f(x) = \begin{cases} \int\limits_0^1 (1 + \frac{1}{\sqrt{t}}) \mathrm{d}t & \text{if } x = 0, \\ +\infty & \text{if } x \neq 0. \end{cases}$$

That implies $\partial I_f(0) = \mathbb{R}$, *but* $\int_T \partial_\varepsilon f(t,0) = \emptyset$ *for* $\varepsilon < 1$.

The above example motivates us to use larger sets than $\int_T \partial_\varepsilon f(t,x) \mathrm{d}\mu(t)$ to generalize (8). The main result of Chapter 5 is Theorem 5.9, which solves the emptiness of the integral of the set-valued map adding the indicator of the domain of the function I_f.

Theorem ***5.9*** *Let* $f : T \times X \to \overline{\mathbb{R}}$ *be such that for every finite dimensional space* F *of* X, $f_{|F} : T \times F \to \overline{\mathbb{R}}$ *a is convex normal integrand. Then for every* $\varepsilon \geq 0$ *and* $x \in X$ *we have the formulas*

$$\partial_\varepsilon I_f(x) = \bigcap_{L \in \mathcal{F}(x)} \bigcup_{\substack{\varepsilon = \varepsilon_1 + \varepsilon_2 \\ \varepsilon_1, \varepsilon_2 \geq 0 \\ \ell \in \mathcal{I}(\varepsilon_1)}} \left\{ \int\limits_T \partial_{\ell(t)} (f_t + \delta_{\operatorname{aff}\{L \cap \operatorname{dom} I_f\}})(x) \mathrm{d}\mu(t) + N^{\varepsilon_2}_{\operatorname{dom} I_f \cap L}(x) \right\} \tag{10}$$

$$= \bigcap_{L \in \mathcal{F}(x)} \bigcup_{\ell \in \mathcal{I}(\varepsilon)} \left\{ \int\limits_T \partial_{\ell(t)} (f_t + \delta_{L \cap \operatorname{dom} I_f})(x) \mathrm{d}\mu(t) \right\}, \tag{11}$$

where $\mathcal{F}(x) := \{V \subseteq X : V$ *is finite dimensional linear space and* $x \in V\}$ *and* $\mathcal{I}(\eta) := \{\ell \in L^1(T, \mathbb{R}_+) : \int_T \ell(t) \mathrm{d}\mu(t) \leq \eta\}$.

Using Theorem above we established simplifications under some additional assumptions, in particular a generalization of Formula (8) is provided for the case of the ε-subdifferential.

Corollary ***5.23*** *In the setting of Theorem 5.9 assume that one of the following conditions holds.*

(a) $(T, \mathcal{A}) = (\mathbb{N}, \mathcal{P}(\mathbb{N}))$.

(b) X, X^* *are Suslin spaces and* f *convex normal integrand.*

In addition, assume that I_f *and* $(f(t,\cdot))_{t \in T}$ *are continuous at some* x_0. *Then for every* $x \in X$

$$\partial_\varepsilon I_f(x) = \bigcup_{\substack{\varepsilon = \varepsilon_1 + \varepsilon_2 \\ \varepsilon_1, \varepsilon_2 \geq 0 \\ \ell \in \mathcal{I}(\varepsilon_1)}} \left\{ (w) - \int\limits_T \partial_{\ell(t)} f_t(x) \mathrm{d}\mu(t) + N^{\varepsilon_2}_{\operatorname{dom} I_f}(x) \right\}.$$

In Example 5.25 it is shown the necessity of considering the weak integral of a multifunction instead of the strong integral.

From the fact that the measure space $(T, \mathcal{A}, \mu)$ could be any complete σ-finite measure, Theorem 5.9 also provides us important corollaries about the subdifferential of finite sums of convex functions and also series of convex functions. In the case of finite sums of convex functions Corollary 5.14 shows that when some qualification condition holds for a particular function, for instance the continuity at some point of its domain, then the formulas provided by Hiriart-Urruty and Phelps (see Proposition 1.2) can be simplified using the exact subdifferential for the function which satisfies the qualification condition. The exact statement is the following.

Corollary ***5.17*** *Let* $\{g_{\mathrm{i}}\}_{\mathrm{i}=1}^{k}, \{g_{\mathrm{i}}\}_{\mathrm{i}=k+1}^{p} \subseteq \Gamma_0(X)$ *such that*

$$\bigcap_{\mathrm{i}\leq k} \mathrm{ri}_{\mathrm{aff}(\mathrm{dom}\, g_{\mathrm{i}})}(\mathrm{dom}\, g_{\mathrm{i}}) \cap \bigcap_{j\geq k+1} \mathrm{dom}\, g_j \neq \emptyset$$

and for every $j \leq k$, g_j *is continuous on* $\mathrm{ri}_{\mathrm{aff}(\mathrm{dom}\, g_{\mathrm{i}})}(\mathrm{dom}\, g_{\mathrm{i}})$*. Then for all* $x \in X$

$$\partial\left(\sum_{\mathrm{i}=1}^{p} g_{\mathrm{i}}\right)(x) = \bigcap_{\substack{\varepsilon_{\mathrm{i}}>0, \mathrm{i}>k \\ \sum \varepsilon_{\mathrm{i}}=\varepsilon}} \mathrm{cl}\left\{\sum_{\mathrm{i}\leq k} \partial\, g_{\mathrm{i}}(x) + \sum_{\mathrm{i}>k} \partial_{\varepsilon_{\mathrm{i}}}\, g_{\mathrm{i}}(x)\right\}.$$

The reader can understand the above formulation as an intermediate result between Hiriart-Urruty and Phelps and Rockafellar formulas for the sum of convex functions.

In the case of series of convex functions the results found in this chapter correspond to a generalization of the formulas presented in [121, 131]. Also the following example gives an answer to the question proposed in [121, Question 2.12].

Example ***5.27*** *Consider* $(T, \mathcal{A}) = (\mathbb{N}, \mathcal{P}(\mathbb{N}))$ *and* $X = \ell^1$*, and let* $(\mathrm{e}_n)_n$ *be the canonical basis of* ℓ^1*, and* μ *be the finite measure given by* $\mu(\{n\}) = 1$*. Define the integrand* $f : \mathbb{N} \times \ell^1 \to \mathbb{R}$ *as* $f(n, x) := |\langle \mathrm{e}_n, x\rangle|^{1+1/n}$*, so that* $I_f(x) = \sum |\langle \mathrm{e}_n, x\rangle|^{1+1/n} < +\infty$*. Since each* f_n *is a Fréchet-differentiable convex function such that* $\nabla f_n(x) = (1 + \frac{1}{n})|\langle x, \mathrm{e}_n\rangle|^{1/n}\mathrm{e}_n$*, according to Corollary 5.26* I_f *is Gâteaux-differentiable on* ℓ^1*, with Gâteaux-differential equal to* $\sum \nabla f_n(x) := \int_{\mathbb{N}} \nabla f_n(x)\mathrm{d}\mu(n) = \sum(1 + \frac{1}{n})|\langle x, \mathrm{e}_n\rangle|^{1/n}\mathrm{e}_n$ *(by Corollary 5.14). Thus, if* I_f *would be Fréchet-differentiable at* $x = 0$*, then we would have*

$$\frac{I_f(n^{-1}\mathrm{e}_n) - I_f(0) - n^{-1}\langle \nabla I_f(0), \mathrm{e}_n\rangle}{n^{-1}} = nn^{-1-\frac{1}{n}} = n^{-\frac{1}{n}} \to 1,$$

which is a contradiction.

From the fact that our formulas involve the use of the Normal cones, we also provide characterizations of this object in terms of the data f_t (see for example Corollary 5.34, Lemma 5.36 and Theorem 5.37).

In [64] Ioffe proved that the subdifferential of the convex integral functional I_f can be characterized as the limits of measurable functions $x^*(t) \in \partial\, f(t, x(t))$ with certain properties. More precisely, he stated the following result, which can be found in [64, Theorem 1 and 2].

Theorem B *Let X be a separable Banach space and let $f : T \times \mathbb{R}^n \to \overline{\mathbb{R}}$ be a convex normal integrand satisfying the following lower bound condition: There exist $a^* \in \mathrm{L}^1_{w^*}(T, X^*)$ and $\alpha \in \mathrm{L}^1(T, \mathbb{R})$ such that*

$$f(t,x) \geq \langle a^*(t), x\rangle + \alpha(t), \quad \textit{for all } t \in T,\ x \in X. \tag{12}$$

Then one has $x^* \in \partial I_f(x)$ if and only if there exist nets of integrable functions $x_\nu \in \mathrm{L}^1(T, X)$ and $x^*_\nu \in \mathrm{L}^1_{w^*}(T, X^*)$ such that:

(a) $x^*_\nu(t) \in \partial f(t, x_\nu(t))$ *ae,*

(b) $w^*\text{-}\lim \int\limits_T x^*_\nu \mathrm{d}\mu(t) = x^*$,

(c) $\lim \int\limits_T f(t, x_\nu(t)\mathrm{d}\mu(t) = I_f(x)$,

(d) $\lim \int\limits_T \langle x^*_\nu(t), x - x_\nu(t)\rangle \mathrm{d}\mu(t) = 0.$

In addition, if X is reflexive, the preceding nets can be replaced by sequences and the w^*-convergence of $\int\limits_T x^*_\nu \mathrm{d}\mu(t)$ can be taken in the norm topology.

Also in the same paper the author made the statement that the measurable selection x_ν can be taken in $\mathrm{L}^p_{w^*}(T, X^*)$ for all $p \in [1, +\infty)$, but he could not prove the case $p = +\infty$. Recently, Lopez and Thibault [80] generalized the last result for any $p \in [1, +\infty]$. They understood the integral functional I_f as the composition of the operator $\hat{I}_f$ defined on $\mathrm{L}^p(T, X)$ and the linear functional $X \ni x \to x\mathbb{1}_T \in \mathrm{L}^p(T, X)$. Then, they applied the well-known *fuzzy composition rule* (see e.g. [116, Theorem 1]) to provide a generalization of the above result. Section 5.7 also responds partially to the question made by Ioffe in a different way, in the sense that the approximate measurable selections can also converge uniformly almost everywhere in various forms, see Theorems 5.38 and 5.41 for more details.

It is important to mention that this Chapter gives an upper-estimate for the Clarke's subdifferential of an integral functional in an arbitrary locally convex space, the reader can see [23, Theorem 2.7.2] for comparing the classical result with Proposition 5.29. Also our upper-estimate can be compared with a recent result due to Murdokhovich-Sagara in an arbitrary Asplund space [86]. Finally, Proposition 5.29 gives us Corollary 5.30, which guarantees the closedness of the integral of measurable multifunction with convex closed values.

Sequential and exact formulae for the subdifferential of nonconvex integral functionals

The aim of Chapter 6 is the study of Hadammard, Fréchet, Proximal, Mordukhovich, G- and Clarke subdifferentials of the integral functional

$$I_f(x) = \int_T f(t,x)\mathrm{d}\mu(t),$$

when the normal integrand f is not necessarily convex with respect to x.

The first step was to find some sequential approximation as in the result of Ioffe [64] and Lopez-Thibault [80]. For this reason in section 6.1.1, inspired by the work of Lopez-Thibault in the convex case, we treated the integral functional I_f as the composition between the operator $\hat{I}_f$ defined on $\mathrm{L}^p(T,X)$ and the linear functional $x \to x\mathbb{1}_T \in \mathrm{L}^p(T,X)$. In this section we apply a recent results on the calculus of the Fréchet subdifferential of the functional $\hat{I}_f$ on $\mathrm{L}^p(T,X)$ with $p \in (1,+\infty)$ under the assumption that the measure space $(T,\mathcal{A},\mu)$ is non-atomic (see [90, Theorem 12 and Theorem 22]), then using theses theorems we apply *the fuzzy chain rule* for the Fréchet subdifferential (see e.g. [14, Theorem 3.5.2]) to establish a sequential formula for the Fréchet subdifferential of I_f in the spirit of the result of Ioffe and Lopez-Thibault mentioned previously (see Theorem 6.1 for our result).

Unfortunately, in this part of the chapter we also showed that we cannot study the functional I_f as a simple matter of a composition between $\hat{I}_f$ and the linear functional $x \to x\mathbb{1}_T \in \mathrm{L}^1(T,X)$ (see Example 6.2). Also the lack of formulas for the case of non-atomic measures and for other subdifferentials impedes us to continue the study of the functional I_f using these ideas. Facing this inconvenience in Section 6.1.2 we adapt a notion of *robusted local minimum* for this class of functionals (see Definition 6.3). This definition opens a gate to apply *Borwein-Preiss's variational principle* (see e.g. [14, Theorem 2.4.1 and Theorem 2.5.2]) as in the case of discrete sums. Posteriorly, in Section 6.1.3 we established approximate formulae for *bornological subdifferentials* and for the *proximal subdifferential* of the integral function I_f by means of measurable selections in the subdifferential of the functions $f(t,\cdot)$. Our results establish the following (see Theorems 6.10, 6.11 and 6.12 for another versions of this statement).

Theorem *6.12* *Let $f : T \times X \to [0,+\infty]$ be a normal integrand and assume that X is separable, or $(T,\mathcal{A}) = (\mathbb{N},\mathcal{P}(\mathbb{N}))$. Consider β a bornology on X and $p,q \in [1,+\infty]$ with $1/p+1/q=1$. Consider $x^* \in \hat{\partial} I_f(x)$. If $\hat{\partial}$ and X satisfy one of the following conditions:*

(i) $\hat{\partial} = \partial_\beta^-$, X is a β-smooth space, or

(ii) $\hat{\partial} = \partial_P$, X is a Hilbert space.

Then for every w^*-continuous seminorm ρ in X^*, there exist sequences $x_n \in \mathrm{L}^p(T,X)$, $x_n^* \in \mathrm{L}^q_{w^*}(T,X^*)$ such that

(a) $x_n^*(t) \in \hat{\partial} f(t,x_n(t))$ *ae.*

(b) $\|x-y_n\| \to 0, \|x-x_n(\cdot)\|_p \to 0.$

(c) $\displaystyle\int_T \|x_n^*(t)\|\|x_n(t)-y_n\|\mathrm{d}\mu(t) \to 0.$

(d) $\displaystyle\int_T \langle x_n^*(t), x_n(t)-x\rangle \mathrm{d}\mu(t) \to 0.$

(e) $\displaystyle\rho\Big(\int_T x_n^*(t)\mathrm{d}\mu(t) - x^*\Big) \to 0.$

(f) $\displaystyle\int_T |f(t,x_n(t)) - f(t,x)|\mathrm{d}\mu(t) \to 0.$

A straightforward application of the techniques of separable reduction gives us the following corollary.

Corollary 6.15 *The statement of Theorem 6.12 holds with $\hat{\partial} = \partial_F$ if we assume that X is a non-separable Asplund space and $(T, \mathcal{A}) = (\mathbb{N}, \mathcal{P}(\mathbb{N}))$.*

Theorem 6.12 gives us Corollary 6.16, which characterizes the subdifferential of a convex integral functional. Also an appropriate comment was written about the non-necessity of the smoothness assumption in this result when the measure is $(T, \mathcal{A}) = (\mathbb{N}, \mathcal{P}(\mathbb{N}))$ (see Remark 6.17).

Corollary ***6.16*** *In the setting of Theorem 6.12 assume that f is a convex normal integrand (i.e. f_t is convex for all $t \in T$). Then one has $x^* \in \partial I_f(x)$ if and only if there are nets $x_\nu \in \mathrm{L}^\infty(T, X)$ and $x^*_\nu \in \mathrm{L}^1(T, X^*)$ such that $x^*_\nu(t) \in \partial f(t, x_\nu(t))$ ae, $\|x - x_\nu(\cdot)\|_\infty \to 0$, $\int_T x^*_\nu(t) \mathrm{d}\mu(t) \overset{w^*}{\to} x^*$, $\int_T \langle x^*_\nu(t), x_\nu(t) - x\rangle \mathrm{d}\mu(t) \to 0$ and $\int_T |f(t, x_\nu(t)) - f(t, x)| \mathrm{d}\mu(t) \to 0$. If the space X is reflexive we can take sequences instead of nets and the w^*-convergence of $\int_T x^*_\nu(t) \mathrm{d}\mu(t)$ is in norm topology.*

It is worth comparing our approach to the work due to Ioffe and Lopez-Thibault. First the lower bound condition (12) can always be carried to our setting considering the function $h(t, x) := f(t, x) - \langle a^*(t), x\rangle - \alpha(t)$ (see also Theorem 6.11 and 6.12 for other lower bound assumptions). Second our theorem improves the approximation found by Ioffe. Third, our sequential approximation does not involve the use of *singular elements* in the dual of $\mathrm{L}^\infty(T, X)$ to characterize the subgradients as in the Theorem due to Lopez-Thibault. Finally, we bypass the convexity assumption in the results of Ioffe and Lopez-Thibault.

Finally, in Section 6.2 we encountered with calculus of the Limiting/Mordukhovich subdifferential, the G-subdifferential and the Clarke's subdifferential. Due to the multiple subdifferentials involved in the following results we adopt the following notation: If X is an F-smooth space $\hat{\partial} = \partial^-_F$, $\partial_L = \partial_M$, $\partial^\infty_L = \partial^\infty_M$, otherwise $\hat{\partial} = \partial^-_H$, $\partial_L = \tilde{\partial}^\infty_G$, $\partial^\infty_L = \tilde{\partial}^\infty_G$.

Since the interchange between the limits and the sign of integral is not for free, we had to impose an additional assumption. We use a condition, which generalizes the classical condition about the Lipschitz continuity of the integral functional (see e.g. [23, Theorem 2.7.2]). In the classical Lipschitz condition it is required that for a point of interest $\bar{x}$ there exist a neighborhood of the point U and an integrable function K such that

$$|f(t, y) - f(t, z)| \leq K(t)\|y - z\|, \text{ for all } y, z \in U \text{ and all } t \in T \tag{13}$$

Instead of the above condition we impose a subdifferential inclusion which is

$$\hat{\partial} f(t, y) \subseteq K(t)\mathbb{B}(0, 1) + C(t), \text{ for all } y, z \in U \text{ and all } t \in T,$$

where K is an integrable function and the multifunction $C(t)$ is a set-valued mapping satisfying the *Integrable compact sole* property (see Definition 6.20), which in the particular case when C is a constant convex closed cone, is equivalent to assuming that the polar set

$$C^- := \{\ x \in X : \langle x^*, x\ \rangle \leq 0, \text{ for all } x^* \in C\},$$

has a non-empty interior (see Lemma 6.22 for a more detailed equivalence of this definition). Also the reader can notice that when C is reduced to the zero cone, then our definition reduces to (13).

With this assumption the we obtain an upper-estimate for the Mordukhovich suddifferential, the G-suddifferential and the Clarke-Rockafellar subdifferential of the integral functional I_f (see Theorem 6.21 and Corollary 6.27). One of the most important characteristics of this condition is that it gives us an upper-estimate for the subdifferential of integral functionals even when the function I_f is not Lipschitz continuous. As far as we know this is the first formula for nonconvex and non Lipschitz integral functionals. Now we present the main result of Section 6.2.

Theorem ***6.21*** *Let X be a separable Banach space and let $x \in \operatorname{dom} I_f$. Suppose there exist $\delta > 0$, a measurable multifunction $C : T \rightrightarrows X^*$ satisfying the integrable compact sole property, and an integrable function $K(\cdot) > 0$ such that*

$$\hat{\partial} f(t, x') \subseteq K(t)\mathbb{B}(0,1) + C(t), \forall x' \in \mathbb{B}(x,\delta),\ \forall t \in T,$$

Then

$$\partial_L I_f(x) \subseteq \bigcap \left\{ \int\limits_T \partial_L f(t,x) \mathrm{d}\mu(t) + UI(C)^- + W^\perp \right\},$$

$$\partial_L^\infty I_f(x) \subseteq \bigcap \left\{ \int_T \partial_L^\infty f(t,x) \mathrm{d}\mu(t) + UI(C)^- + W^\perp \right\},$$

where the intersection is over all finite dimensional subspaces W of X and

$$UI(C) := \{\xi \in X : \sigma_{C(\cdot)}(\xi)^+ \in \mathrm{L}^1(T, \mathbb{R})\}.$$

Consequently,

$$\partial_C I_f(x) \subseteq \overline{\mathrm{co}}^{w^*} \left\{ \int\limits_T \partial_L f(t,x) \mathrm{d}\mu + \int_T \partial_L^\infty f(t,x) \mathrm{d}\mu(t) + UI(C)^- \right\}.$$

Now let we illustrate the above result with a simple example.

Example ***6.29*** *Consider the integrand $f :]0,1] \times \mathbb{R} \to [0, +\infty)$ given by*

$$f(t,x) = \begin{cases} x^{3/2} t^{-1+x} & \text{if} \quad x > 0, \\ 0 & \text{if} \quad \textit{not}. \end{cases}$$

It is easy to check that f is continuously differentiable with respect to x and

$$I_f(x) = \begin{cases} \sqrt{x} & \text{if} \quad x > 0, \\ 0 & \text{if} \quad x \le 0. \end{cases}$$

Then one gets easily $\partial_M I_f(0) = \partial_M^\infty I_f(0) = [0, +\infty)$,

$$\partial_F f(t,x) = \begin{cases} \frac{3}{2} x^{1/2} t^{-1+x} + x^{3/2} \ln(t) t^{-1+x} & \text{if} \quad x > 0, \\ 0 & \text{if} \quad x \le 0, \end{cases}$$

and $\partial_M f(t,x) = \{0\}$. Then we can consider $C(t) = [0, +\infty)$, applying our result we get

$$\partial I_f(0) = \int_{]0,1]} \partial_M f_t(0) \mathrm{d}\mu(t) + UI(C)^- = \{0\} + [0, +\infty)$$

and $\partial^\infty_M I_f(0) = [0, +\infty)$. *The same example can be modified as*

$$f(t,x) = \begin{cases} x^2 t^{-1+x} & \text{i}f \quad x > 0, \\ 0 & \text{i}f \quad x \leq 0. \end{cases}$$

Then one has

$$I_f(x) = \begin{cases} x & \text{i}f \quad x > 0, \\ 0 & \text{i}f \quad x \leq 0. \end{cases}$$

Therefore, the integral functional I_f *is Lipschitz continuous, but it is not true that* $\partial_M I_f(0) = [0,1]$ *is included in* $\int_{]0,1]} \partial_M f(t,0) \mathrm{d}\mu(t) = \{0\}$ *as in classical results (see [85, Lemma 6.18] and also [86] for an extension of this result). However, our results guarantee the non-trivial inclusion* $\partial_M I_f(0) \subseteq \int_{]0,1]} \partial_M f(t,0) \mathrm{d}\mu(t) + [0, +\infty)$.

As a final comment we recall that in finite dimensional setting two lsc functions f_1, f_2 satisfy the sum rule inclusion $\partial_M (f_1 + f_2)(x) \subseteq \partial_M f_1(x) + \partial_M f(x)$ at a point x provided that the asymptotic qualification condition $\partial^\infty_M f_1(x) \cap \partial^\infty_M f_2(x) = \{0\}$ holds. However, the reader can notice that in the above example the integrand is continuously differentiable and then the singular subdifferential $\partial^\infty_M f_t(0) = \{0\}$ for all $t \in T$. The last shows that it is not possible to recover similar criteria, as in the finite sums, in terms of the singular subdifferentials, to get an inclusion of the form $\partial I_f(x) \subseteq \int_T \partial_M f_t(x) \mathrm{d}\mu(t)$.

Subdifferential characterization of probability functions under Gaussian distribution

For a *probability function* it is understood an assignment of the form

$$x \to \varphi(x) := \mathbb{P}(g(x,\xi) \leq 0), \tag{14}$$

where $g : X \times \mathbb{R}^m \to \mathbb{R}^s$ is a mapping defining a system of random inequalities and ξ is an m-dimensional random vector defined in some probability space $(\Omega, \mathcal{A}, \mathbb{P})$. The reader also can notice that from the theoretical viewpoint of the problem one can assume $s = 1$ passing to the maximum over the component $\max g_j$. Typical applications for this class of functions can be found in water management, telecommunications, electricity network expansion, mineral blending, chemical engineering, etc, where the problem can be model as:

$$\begin{gathered} \min c(x) \\ \text{s.t } \mathbb{P}(g(x,\xi) \leq 0) \geq p, \\ x \in \mathcal{C}. \end{gathered} \tag{15}$$

Here c represents a cost function, $\mathcal{C}$ a non-empty set, typically a set of the form $\{x \mid h_{\mathrm{i}}(x) \leq 0, h_j(x) = 0, \text{ for all } \mathrm{i} = 1, ..., N;\ j = 1, ..., M\}$ with smooth functions h_{i}, h_j, and the inequality $\mathbb{P}(g(x,\xi) \leq 0) \geq p$ expresses that a decision vector x is feasible if and only if the random inequality $g(x,\xi) \leq 0$ is satisfied with probability at least p.

When one solves probabilistic constrained problems via numerical nonlinear optimization methods, almost any algorithms needs not only to calculate the values of the functional φ; one also needs to have an access to gradients of φ, or subgradients when the optimization problems is nonsmooth. It has been proved that particular cases, for example when the probability function has the form $\varphi(x) = \mathbb{P}(D(x)\xi \leq c(x))$ under *Gaussian distribution* and with a possible singular matrix $D(x)$, the computation of the values of φ and also the gradient of the probability function can be analytically reduced to the computation of Gaussian distribution functions (see e.g. [55, 122, 124–126]). However, when the models on the variable involving ξ are nonlinear, a reduction to distribution cannot be applied. In this setting, another approach, the so-called *spherical radial decomposition of Gaussian random vectors* (see for example [49]) appears to be promising for calculating both functions values and gradients of φ. Basically for a m-dimensional Gaussian vector distributed according to $\xi \sim \mathcal{N}(0, A)$ for some correlation matrix A; the spherical radial decomposition decomposes the vector ξ into the product of two random vectors η and ζ in the form $\xi = \eta L\zeta$, where $A = L \times L^T$ is the Cholesky decomposition of the matrix A, η has a χ-distribution with m-degrees of freedom, and ζ has a uniform over the Euclidean unit sphere of $\mathbb{R}^m$, $\mathbb{S}^{m-1} := \{z \in \mathbb{R}^m \mid \sum z_{\mathrm{i}}^2 = 1\}$. This decomposition allows us to write the probability distribution as an nonconvex integrand functional, as studied in Chapter 6. More precisely,

$$\varphi(x) := \int_{S^{m-1}} \mathrm{e}(v, x)\mathrm{d}\mu_\zeta(v),$$

where

$$\mathrm{e}(v, x) = \mu_\eta(\{r \geq 0 : g(x, rLv) \leq 0\}). \tag{16}$$

Here μ_η is the one-dimensional Chi-distribution with m degrees of freedom, and μ_ζ refers to the uniform distribution on $\mathbb{S}^{m-1}$.

The last formulation resulted crucial and appropriate for solving this problem using the results obtained in Chapter 6. Nevertheless, the variational behavior of the function $\mathrm{e}(v, \cdot)$ involves nonsmoothness, even when function g is continuously differentiable in both variables. The following example given in [123] shows that it is not enough considering the function g to be smooth to obtain the differentiability of the function φ

Example *Let ξ have a one-dimensional standard Gaussian distribution, and define*

$$g(x_1, x_2, x_3, \xi) := (\xi - x_1, \xi - x_2, -\xi - x3).$$

Then, with $\Phi(t) := \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{t} \mathrm{e}^{-\tau^2/2}\mathrm{d}\tau$ being the one-dimensional standard Gaussian distribution function, one has that

$$\varphi(x_1, x_2, x_3) = \max\{\ \min\{\Phi(x_1), \Phi(x_2)\} - \Phi(x_3), 0\}.$$

Clearly, φ fails to be differentiable at $x := (0, 0, -1)$.

Even when it is considered only one single continuously differentiable inequality the nonsmoothness of the function φ could be appear.

Example 7.15 *Define the function $g : \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}$ by*

$$g(x, z_1, z_2) := \alpha(x)\mathrm{e}^{h(z_1)} + z_2 - 1,$$

where

$$\alpha(x) := \begin{cases} x^2 & x \geq 0 \\ 0 & x < 0, \end{cases}$$

$$h(t) := -1 - 4\log(1 - \Phi(t)); \quad \Phi(t) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} \mathrm{e}^{-\tau^2/2} \mathrm{d}\tau.$$

Moreover, let ξ have a bivariate standard normal distribution, i.e.,

$$\xi = (\xi_1, \xi_2) \sim \mathcal{N}\left((0,0), \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right).$$

Then φ fails to be differentiable at $x = 0$, see Figure 2 for a graphical inspection and see Section 7.6 for details.

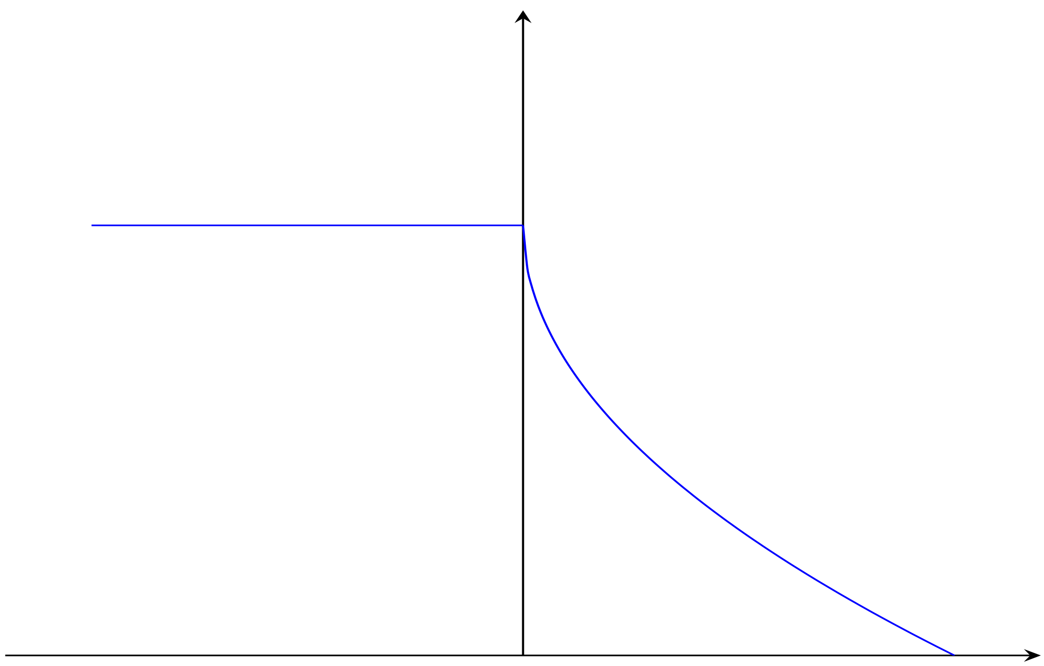

Figure 2: Probability function φ of Example 7.15.

Above examples show that it is not sufficient to study the differentiability of the function φ. To the best knowledge of the author, Henrion and Van Ackooij in [123] where the first authors in to study the Clarke's subdifferential of the probability function. They provided criteria for the differentiability and Lipschitz continuity of the function φ together with formulae for the gradient and subgradients of the function φ at a point of interest $\bar{x}$ under the assumption that each component g_{i} is continuously differentiable and convex with respect to the second argument. To avoid pathological situations the authors impose a growth condition over the gradient $\nabla_x g(x,z)$ in a neighborhood of the point $\bar{x}$. They assumed that the components g_{i} satisfy the *exponential growth condition*, which is

$$\exists l > 0 : \|\nabla_x g_{\mathrm{i}}(x,z)\| \leq l \mathrm{e}^{\|z\|} \quad \forall x \in \mathbb{B}(\bar{x}, 1/l), \ \|z\| \geq l, \ \mathrm{i} = 1, \ldots, p. \tag{17}$$

The aim of this work is to extend the previous results to infinite-dimensional setting on X, using a weaker growth condition and assuming only local Lipschitz continuity of g. Our main assumptions of this chapter are:

(H)

1. X is a reflexive and separable Banach space.
2. Function g is locally Lipschitzian as a function of both arguments simultaneously, and convex as a function of the second argument.
3. The random vector ξ is Gaussian of type $\xi \sim \mathcal{N}(0, A)$, where A is a correlation matrix.

Before presenting our results let me introduce some notations. To understand the function $\mathrm{e}(v,x)$ we associate the *finite* and *infinite directions* defined respectively as

$$\begin{aligned} F(x) &:= \{v \in \mathbb{S}^{m-1} \mid \exists r \geq 0 : g(x, rLv) = 0\}, \\ I(x) &:= \{v \in \mathbb{S}^{m-1} \mid \forall r \geq 0 : g(x, rLv) < 0\}. \end{aligned}$$

It is easily observed that $F(x) \cap I(x) = \emptyset$ and that $F(x) \cup I(x) = \mathbb{S}^{m-1}$ by continuity of g. Moreover, the number $r \geq 0$ satisfying $g(x, rLv) = 0$ in the case of $v \in F(x)$ is uniquely defined, due to the convexity of g in the second argument. We use the auxiliary function $\rho : \mathbb{S}^{m-1} \times X \to [0, +\infty]$ called *radius function* defined as:

$$\rho(v, x) := \begin{cases} r \text{ such that } g(x, rLv) = 0 & \text{if } v \in F(x), \\ +\infty & \text{if } v \in I(x). \end{cases} \tag{18}$$

This definition allows us to rewrite the radial probability function e from (16) in the form

$$\mathrm{e}(v, x) = \mu_\eta([0, \rho(v, x)]) = F_\eta(\rho(v, x)) \tag{19}$$

whenever $g(x, 0) < 0$. Here, F_η refers to the Chi-distribution with m degrees of freedom, so that $F'_\eta(t) = \chi(t)$, where χ is the corresponding density:

$$\chi(t) := \frac{2^{1-\frac{m}{2}}}{\Gamma(\frac{m}{2})} t^{m-1} \mathrm{e}^{-t^2/2} \quad \forall t \geq 0.$$

The second equation in (19) follows from $F_\eta(0) = 0$. We formally put $F_\eta(\infty) := 1$ which translates the limiting property $F_\eta(t) \overset{t \to +\infty}{\longrightarrow} 1$ of cumulative distribution functions.

Roughly speaking, our weaker growth condition was motivated by the results given in Chapter 6. These results allow us to give an upper estimate of the subdifferential of the function φ, even when it could be non-Lipschitz, considering the *l-cone of nice directions* at x defined as:

$$C_l(x) := \{h \in \mathbb{R}^n \mid g^\circ(\cdot, z)(y; h) \leq l \|z\|^{-m} \exp(\frac{\|z\|^2}{2\|L\|^2}) \|h\| \;\; \forall y \in \mathbb{B}_{1/l}(x), \, \|z\| \geq l\}.$$

With this cone one can obtain the following inclusion (see Theorem 7.10)

$$\partial_F \mathrm{e}(v, y) \subseteq \mathbb{B}(0, r)(0) - C_l^-(x) \quad \forall y \in U, \, v \in \mathbb{S}^{m-1},$$

where U is a neighborhood of x and r is some constant. Then the above inclusion fit perfectly with the formulas found in Chapter 6 and then applying our result to this problem we obtain the following upper estimate.

Theorem ***7.12*** *Let $x_0 \in X$ be such that $g(x_0, 0) < 0$. Assume that the cone $C_l(x_0)$ has a non-empty interior for some $l > 0$. Then,*

(i) $\partial_M \varphi(x_0) \subseteq \mathrm{cl}^* \left\{ \int\limits_{\mathbb{S}^{m-1}} \partial_M \mathrm{e}(v, x_0) \mathrm{d}\mu_\zeta(v) - C_l^-(x_0) \right\}.$

(ii) *Provided that X is finite-dimensional,*

$$\partial_M \varphi(x_0) \subseteq \int\limits_{\mathbb{S}^{m-1}} \partial_M \mathrm{e}(v, x_0) \mathrm{d}\mu_\zeta(v) - C_l^-(x_0).$$

(iii) $\partial^\infty \varphi(x_0) \subseteq -C_l^-(x_0)$.

(vi) $\partial_C \varphi(x_0) \subseteq \overline{\mathrm{co}} \left\{ \int\limits_{\mathbb{S}^{m-1}} \partial_M \mathrm{e}(v, x_0) \mathrm{d}\mu_\zeta(v) - C_l^-(x_0) \right\}.$

The above result obviously yields simplifications under additional assumptions, see Proposition 7.14 and obviously one can conclude the continuously differentiability of the probability function φ, see Proposition 7.18.

The final goal of this chapter was an application to a finite system of smooth inequalities. The probability function under study is given by

$$\varphi(x) := \mathbb{P}\left(g_{\mathrm{i}}(x, \xi) \le 0,\ \mathrm{i} = 1, \ldots, p\right),\ x \in X, \tag{20}$$

where each g_{i} is a continuously differentiable function convex in the second argument. Clearly, as we said previously, this can be recast in the form of (14) upon defining

$$g := \max_{\mathrm{i}=1,\ldots,p} g_{\mathrm{i}}.$$

Then the function g satisfies our assumptions. To continue we may associate with each component its radius function ρ_{i} as in (18) and the corresponding radius function ρ associated to g. Then applying Theorem 7.12 we obtain:

Theorem ***7.20*** *Fix $x_0 \in X$ with $g(x_0, 0) < 0$, and assume that for some $l > 0$ it holds, for $\mathrm{i} = 1, \ldots, p$,*

$$\|\nabla_x g_{\mathrm{i}}(x, z)\| \le l \|z\|^{-m} \mathrm{e}^{\frac{\|z\|^2}{2\|L\|^2}} \quad \forall x \in \mathbb{B}(x_0, 1/l),\ \|z\| \ge l. \tag{21}$$

Then the probability function (20) is locally Lipschitz near x_0 and there exists a nonnegative number $R \le \sup\{\|x^*\| \mid x^* \in \partial_M \mathrm{e}(x_0, v)$ and $v \in I(x_0)\}$ such that

$$\partial_C \varphi(x_0) \subseteq -\int\limits_{v \in F(x_0)} \mathrm{co} \left\{ \bigcup_{\mathrm{i} \in T(v)} \frac{\chi(\rho(x_0, v)) \nabla_x g_{\mathrm{i}}(x_0, \rho(x_0, v) Lv)}{\langle \nabla_z g_{\mathrm{i}}(x_0, \rho(x_0, v) Lv), Lv \rangle} \right\} \mathrm{d}\mu_\zeta(v)$$
$$+ \mu_\zeta(I(x_0)) \mathbb{B}(0, R).$$

Here, $T(v) := \{\mathrm{i} \in \{1, \ldots, p\} \mid \rho_{\mathrm{i}}(x, v) = \rho(v, x)\}$.

This result intermediately entails the following corollary about the differentiability of the probability function (20) .

Corollary ***7.21*** *If in the setting of Theorem* 7.20 *one has that* $\mu_\zeta(I(x_0)) = 0$, *or the constant* R *in Theorem 7.20 is zero, then*

$$\partial^C \varphi(x_0) \subseteq -\int_{v\in\mathbb{S}^{m-1}} \operatorname{co}\left\{\bigcup_{\mathrm{i}\in T(v)} \frac{\chi\left(\rho\left(x_0,v\right)\right)\nabla_x g_{\mathrm{i}}\left(x_0,\rho\left(x_0,v\right)Lv\right)}{\left\langle\nabla_z g_{\mathrm{i}}\left(x_0,\rho\left(x_0,v\right)Lv\right),Lv\right\rangle}\right\} \mathrm{d}\mu_\zeta(v).$$

If, in addition, for μ_ζ*-a.e.* $v\in\mathbb{S}^{m-1}$ *we have that* $\#T(v)=1$ *(say:* $T(v)=\{\mathrm{i}^*(v)\}$*), then the probability function* (7.28) *is strictly differentiable with gradient*

$$\nabla\varphi(x_0) = -\int_{v\in\mathbb{S}^{m-1}} \frac{\chi\left(\rho\left(x_0,v\right)\right)\nabla_x g_{\mathrm{i}^*(v)}\left(x_0,\rho\left(x_0,v\right)Lv\right)}{\left\langle\nabla_z g_{\mathrm{i}^*(v)}\left(x_0,\rho\left(x_0,v\right)Lv\right),Lv\right\rangle}\mathrm{d}\mu_\zeta(v).$$

Consequently, if X *is finite-dimensional and* $\#T(v)=1$ *holds true in some neighborhood of* x, *then* φ *is even continuously differentiable at* x.

It is worth mentioning that under the strengthened exponential growth condition (compared (17) with (21)) the constant R in Theorem 7.20 and Corollary above is zero (see also [123, Theorem 3.6 and Theorem 4.1]).

Finally, in this Chapter we provide two examples: The first one is Example 7.15, which on the one hand, serves as an illustration of our main result Theorem 7.12 and, on the other hand, shows that even for a continuously differentiable inequality $g(x,\xi)\le 0$, satisfying a basic constraint qualification, the associated probability function φ may fail to be differentiable, actually even to be locally Lipschitzian (though it is continuous due to the constraint qualification; see Theorem 7.5). The second one shows a probability function which does not satisfy the exponential growth condition (17). Nevertheless, using our results one can prove that the probability function is continuously differentiable. Indeed, the following example together with Example 7.15 will be discussed with more details in the final part of Chapter 7.

Example ***7.16*** *Define the function* $g:\mathbb{R}\times\mathbb{R}^2\to\mathbb{R}$ *by*

$$g\left(x,z_1,z_2\right) := \alpha(x)\frac{\exp(z_1^2/2)}{z_1^2+4}+z_2-1,$$

where

$$\alpha(x) := \begin{cases} x^2 & x\ge 0 \\ 0 & x<0. \end{cases}$$

Moreover, let ξ *have a bivariate standard normal distribution, i.e.,*

$$\xi=(\xi_1,\xi_2)\sim\mathcal{N}\left((0,0),\begin{pmatrix}1&0\\0&1\end{pmatrix}\right).$$

Then the following properties are shown in Appendix 7.6 (see Figure 3):

1. *g is continuously differentiable.*

2. *g is convex in (z_1, z_2).*

3. *$g(0,0,0) < 0$.*

4. *$C_1(0) = \mathbb{R}$.*

5. *g does not satisfy the exponential growth condition at $x_0 = 0$.*

6. *φ is continuously differentiable at 0.*

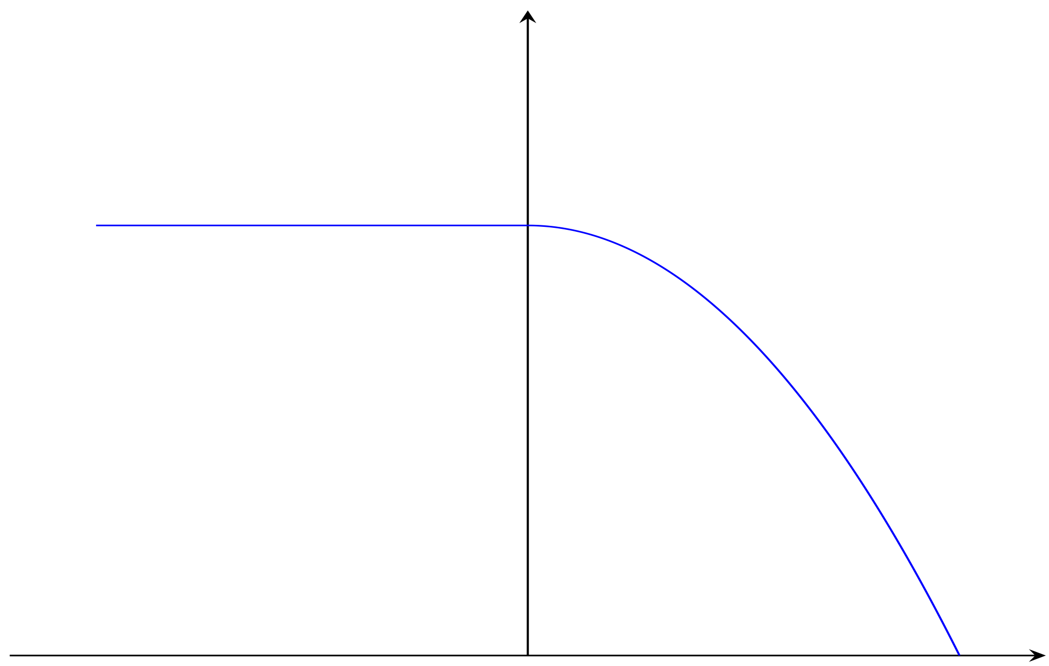

Figure 3: Probability function φ of Example 7.16.

Chapter 1

Notation and preliminary results

1.1 Topological and convex tools

In the following, X and X^* will be two (separated) locally convex spaces (lcs) in duality by the bilinear symmetric form $\langle \cdot, \cdot \rangle : X^* \times X \to \mathbb{R}$, $\langle x^*, x \rangle = \langle x, x^* \rangle = x^*(x)$. When the spaces X and X^* will be abstract lcs they will endowed with compatible topologies τ_X and τ_{X^*}. For a point $x \in X$ (resp. $x^* \in X^*$) $\mathcal{N}_x(\tau_X)$ (resp. $\mathcal{N}_{x^*}(\tau_{X^*})$) represent the neighborhood system of x (resp. x^*) with respect to the topology τ_X (resp. τ_{X^*}) and we omit the symbol τ_X (resp. τ_{X^*}) when there is no confusion. Examples of τ_{X^*} are the weak* topology denoted by $w(X^*, X)$ (w^*, for short) which is the smallest topology compatible with the dual pair (X, τ_x), (X^*, τ_{X^*}), the Mackey topology denoted by $\tau(X^*, X)$, which is the largest topology compatible with the dual pair (X, τ_x), (X^*, τ_{X^*}), and the strong topology denoted by $\beta(X^*, X)$, which is the topology generated by the uniform convergence over bounded sets on X. We will write $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$ and adopt the conventions that $0 \cdot \infty = 0 = 0 \cdot (-\infty)$ and $\infty + (-\infty) = (-\infty) + \infty = \infty$. The notation $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{Q}$ will be reserved for the set of *natural numbers*, the *integer numbers* and the *rational numbers*. For any set A, the symbol $\mathcal{P}(A)$ denotes the set of all subsets of A and $\#A$ denotes the cardinal of a set A. Consider a *preordered set* $(\Re, \preceq)$ (i.e., $\preceq$ is a *reflexive, antisymmetric, and transitive* binary relation on $\Re$) a subset $\Im$ of $\Re$ is said to be *cofinal* if for every $r \in \Re$ there exists $s \in \Im$ such that $r \preceq s$.

A function $\rho : X \to [0, +\infty)$ is called a seminorm if for every $x, y \in X$ and $\lambda \in \mathbb{R}$, $\rho(\lambda x) = |\lambda|\rho(x)$ and $\rho(x + y) \leq \rho(x) + \rho(y)$. For a seminorm ρ, $x \in X$, and $r > 0$ we denote $\mathbb{B}_\rho(x, r) := \{y \in X : \rho(x - z) \leq r\}$ the closed ball with radius r around x with respect to the seminorm ρ. When X will be a Banach space the norm on X and X^* are simply denoted by $\| \cdot \|$ and the closed balls in these spaces are denoted $\mathbb{B}(x, r)$ and $\mathbb{B}(x^*, r)$ for points $x \in X$ and $x^* \in X^*$, for our convenience in this cases is useful to introduce the bidual space of X defined and denoted by $X^{**} := (X^*, \| \cdot \|)^*$. When there is a risk of confusion about the space, the balls will be denoted with a sub-index referring to the space, that is to say, $\mathbb{B}_X(x, r)$ according to the respective space X. For a seminorm ρ, a set C and $x \in X$, we define the distance from x to C with respect to ρ as $\mathrm{d}_C^\rho(x) := \inf_{z \in C} \rho(x - z)$. For two functions $h, g : X \to \overline{\mathbb{R}}$ the notation $g \leq h$ means $g(x) \leq h(x)$ for all $x \in X$.

For a given function $f : X \to \overline{\mathbb{R}}$, the (effective) *domain* of f is $\operatorname{dom} f := \{x \in X \mid f(x) < +\infty\}$. We say that f is *proper* if $\operatorname{dom} f \neq \emptyset$ and $f > -\infty$. The function f is said to be *lower semicontinuous* (lsc) if for every $\lambda \in \mathbb{R}$ the sublevel set $\operatorname{lev}_\alpha f := \{x \in X \mid f(x) \leq \lambda\}$ is closed, it is said to be *inf-compact* if for every $\lambda \in \mathbb{R}$ the sublevel set $\operatorname{lev}_\alpha f$ is compact; the function f is said to be *sequentially τ-inf-compact* if for every $\lambda \in \mathbb{R}$ and every sequence $x_n \in [f \leq \lambda] := \{x \in X \mid f(x) \leq \lambda\}$ there exists a subsequence $x_{n_k} \xrightarrow{\tau} x \in [f \leq \lambda]$. We denote by $\Gamma_0(X)$ the class of proper lsc convex functions on X. The *conjugate* of f is the function $f^* : X^* \to \overline{\mathbb{R}}$ defined by

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

and the *biconjugate* of f is $f^{**} := (f^*)^* : X \to \overline{\mathbb{R}}$.

The directional derivative of a convex lsc function f at $x \in \operatorname{dom} f$ in the direction u is given by

$$f'(x; u) := \inf_{s>0} \frac{f(x + su) - f(x)}{s}.$$

It is well-known that the directional derivative $f'(x; \cdot)$ is a positively homogeneous convex function and $\partial f(x) = \partial f'(x; \cdot)(0)$.

For a subspace F of X, the restriction of the function f to F is denoted by $f_{|F}$. The notation $\overline{\text{co}}_F f : F \to \overline{\mathbb{R}}$ is the function such that $\operatorname{epi}(\overline{\text{co}}_F f) = \overline{\text{co}}(\operatorname{epi} f \cap (F \times \mathbb{R}))$. Then, the *closed convex hull of* f is denoted and defined by $\overline{\text{co}}\, f := \overline{\text{co}}_X f$.

The *indicator* and the *support* functions of a set A $(\subseteq X, X^*)$ are, respectively,

$$\delta_A(x) := \begin{cases} 0 & x \in A \\ +\infty & x \notin A, \end{cases} \qquad\qquad \sigma_A := \delta_A^*.$$

The *inf-convolution* of $f, g : X \to \overline{\mathbb{R}}$ is the function $f \square g := \inf_{z \in X}\{f(z) + g(\cdot - z)\}$; it is said to be exact at x if there exists z such that $f \square g(x) = f(z) + g(x - z)$.

For a pair of points we denote the interval between a, b as $[a, b] := \{(1-\lambda)a + \lambda b : \lambda \in [0, 1]\}$ and $[a, b[= [a, b]\backslash\{b\}$, $]a, b] = [a, b]\backslash\{a\}$ and $]a, b[= [a, b]\backslash\{a, b\}$. For a set $A \subseteq X$ (or $\subseteq X^*$), $\operatorname{int}(A)$, $\overline{A}$ (or $\operatorname{cl} A$), $\operatorname{co}(A)$, $\overline{\text{co}}(A)$, $\operatorname{lin}(A)$ and $\operatorname{aff}(A)$, the interior, the closure, the *convex hull*, the *closed convex hull*, the *linear subspace* and the *affine subspace* generated by A.

Definition 1.1 *The* relative interior *of A with respect to an affine subspace F, denoted by* $\operatorname{ri}_F(A)$, *is the interior of A with respect to F. By the symbol* $\operatorname{ri}(A)$, *we denote* $\operatorname{ri}_{\operatorname{aff}(A)}(A)$ *if* $\operatorname{aff}(A)$ *is closed, and the empty set otherwise.*

The *polar* and the *negative polar* cone of A are the set

$$A^o := \{x^* \in X^* \mid \langle x^*, x\rangle \leq 1, \forall x \in A\},$$
$$A^- := \{x^* \in X^* \mid \langle x^*, x\rangle \leq 0, \forall x \in A\},$$

respectively, and the *recession cone* of A (when A is convex and close) is the set

$$A_\infty := \{x \in X \mid \lambda x + y \in A \text{ for some } y \text{ in } A \text{ and all } \lambda \geq 0\}.$$

For two lcs spaces (Y, τ_Y) and (Z, τ_z) the set $\mathcal{L}(Y, \tau_Y, Z, \tau_z)$ ($\mathcal{L}(Y, Z)$ for short) is defined as the set of all linear continuous functions from Y to Z.

For a net x_k, a point x and a function f (or a set C) the notation $x_k \xrightarrow{f} x$ (respectively $x_k \xrightarrow{C} x$) means $(x_k, f(x_k)) \to (x, f(x))$ (respectively $x_k \in C$ and $x_k \to x$).

Given a set A and a topological space Z, a multifunction (or a set-valued mapping) M from A to Z is a function form A to the power set of Z (i.e the set of all subsets of Z) and it will be denoted as $M : A \rightrightarrows Z$. The *domain*, the *range (or image)* and *the graph* of the multifunction M are defined and denoted as $\operatorname{dom} M := \{a \in A : M(a) \neq \emptyset\}$, $\operatorname{rge} M := \bigcup_{a \in A} M(a)$ and $\operatorname{gph} M := \{(a, b) \in A \times Z : b \in M(a)\}$, respectively. Moreover, M will be called closed (open, compact, convex, etc, respectively) valued if for every $a \in A$ the set $M(a)$ is closed (open, compact, convex, etc, respectively).

1.2 Subdifferential theory

For $\varepsilon \geq 0$ and a function $f : X \to \overline{\mathbb{R}}$ the ε-*subdifferential* of f at a point $x \in X$ where it is finite is the set

$$\partial_\varepsilon f(x) := \{x^* \in X^* \mid \langle x^*, y - x\rangle \leq f(y) - f(x) + \varepsilon, \ \forall y \in X\};$$

if $f(x)$ is not finite, we set $\partial_\varepsilon f(x) := \emptyset$. The special case $\varepsilon = 0$ is the *convex Subdifferential*[1] and it is denoted by $\partial f(x)$.

The ε-normal set of A at x is $N_A^\varepsilon(x) := \partial_\varepsilon \delta_A(x)$. The case $\varepsilon = 0$ is the well-known normal cone and it is simply denoted by $N_A(x)$.

We proceed in this section by presenting formulas for the calculus of the composition and the sum of convex functions with and without qualification conditions. The proofs can be found in [56, Theorems 3.1 and 3.2] and [132].

Proposition 1.2 *[132, Corollary 2.6.7] Let $f, g \in \Gamma_0(X)$. Then for every $x \in X$ we have*

$$\partial_\varepsilon (f + g)(x) = \overline{\bigcup_{\substack{\varepsilon_1, \varepsilon_2 \geq 0 \\ \varepsilon_1 + \varepsilon_2 = \varepsilon}} (\partial_{\varepsilon_1} f(x) + \partial_{\varepsilon_2} g(x))}^{w^*} \text{ for all } \varepsilon > 0,$$

and

$$\partial (f + g)(x) = \bigcap_{\eta > 0} \overline{(\partial_\eta f(x) + \partial_\eta g(x))}^{w^*}.$$

[1]Also called for some authors the Fenchel subdifferential or the Moreau-Rockafellar subdifferential.

Proposition 1.3 *[132, Corollary 2.6.5] Let Y, Z be two lcs, $f \in \Gamma_0(Z)$ and $A \in \mathcal{L}(Y,Z)$. Then for every $y \in Y$ we have*

$$\partial_\varepsilon (f \circ A)(y) = \overline{A^* (\partial_{\eta+\varepsilon} f(Ay))}^{w^*} \text{ for all } \varepsilon \geq 0.$$

Proposition 1.4 *[132, Theorem 2.8.3] Let Y, Z be two lcs, $f \in \Gamma_0(Z)$ and $A \in \mathcal{L}(Y,Z)$. Assume that f is finite and continuous at some point. Then, for every $y \in Y$ we have that*

$$\partial_\varepsilon (f \circ A)(y) = A^* (\partial_\varepsilon f(Ay)) \text{ for all } \varepsilon \geq 0.$$

Now consider a Banach space X. A *bornology* β on X is a collection of closed bounded and symmetric subsets of X which covers X and satisfies the following properties: For any two elements of $A, B \in \beta$ there exists an element $C \in \beta$ such that $A \cup B \subseteq C$ and $aU \in \beta$ whenever $U \in \beta$ and $a > 0$. The β^* topology on X^* is the topology generated by the basis of neighborhoods of zero, $\{U^o : U \in \beta\}$, where $U^o := \{x^* \in X^* \mid \langle x^*, x\rangle \leq 1, \forall x \in U\}$ denotes the polar cone of the set U. Two of the most useful bornologies considered in the literature are the *Fréchet bornology* F consisting of all closed bounded sets and the *Hadamard bornology* H of all norm-compact sets. A function $\phi : X \to \mathbb{R}$ is said to be β-differentiable at $x \in X$ if there exists $x^* \in X$ such that for every $U \in \beta$

$$\lim_{s \to 0^+} \sup_{h \in U} \left| \frac{\phi(x+sh) - \phi(x)}{s} - \langle x^*, h\rangle \right| = 0.$$

We denote $\nabla_\beta \phi(x) = x^*$. A function $\phi : U \to \mathbb{R}$ from an open set U is said to be *β-smooth* if it is β-differentiable at every $x \in U$ and the derivative $\nabla_\beta g$ is a continuous function from $(U, \|\cdot\|)$ to (X^*, β^*). Furthermore, ϕ is a *bump function* if $\operatorname{supp} \phi := \{\, x \in X : \phi(x) \neq 0\}$ is non-empty and bounded. A Banach space X is said to be β-smooth if there is a Lipschitz continuous, bump β-smooth function from X to $\mathbb{R}$. Each separable Banach space admits an H-smooth renorm, if the dual is also separable, then the space admits an F-smooth renorm (see for example [41, 94]). Therefore, these spaces are H-smooth and F-smooth respectively. A function ϕ is said to be C^2 on U if it is F-smooth on U and if the mapping $\nabla_F \phi : U \to X^*$ is continuous differentiable, that is to say, there exists a continuous function $\nabla^2 \phi : U \to \mathcal{L}(X, X^*)$, which satisfies $\lim\limits_{h \to 0} \| \nabla_F \phi(x+h) - \phi(x) - \nabla^2 \phi(x)(h)\| / \|h\| = 0$ for all $x \in U$.

Now let f be a function on X finite at x. Then

$$\begin{aligned}
\partial_F f(x) :=& \{x^* \in X^* \mid \liminf_{h \to 0} \frac{f(x+h) - f(x) - \langle x^*, h\rangle}{\|h\|} \geq 0\}, \\
\partial_P f(x) :=& \{x^* \in X^* \mid \liminf_{h \to 0} \frac{f(x+h) - f(x) - \langle x^*, h\rangle}{\|h\|^2} > -\infty\}
\end{aligned}$$

are called the *Fréchet subdifferential* and the *Proximal subdifferential* of f at x, respectively. For $k > 0$ we set

$$\partial^-_{\beta,k} f(x) := \left\{ x^* \in X^* : \begin{array}{c} \text{there is a neighborhood } U \text{ of } x \text{ and a } \beta\text{-smooth function} \\ \phi : U \to \mathbb{R} \text{ with Lipschitz constant } k \text{ such that} \\ \nabla_\beta \phi(x) = x^* \text{ and } f - \phi \text{ attains a local minimum at } x \end{array} \right\},$$

and the *(viscosity) β-subdifferential* is defined by $\partial_\beta^- f(x) := \bigcup_{k>0} \partial_{\beta,k}^- f(x)$ (see for example [11] and the reference therein). It is important to mention that when the space X has a Fréchet smooth renorm, then ∂_β^- and ∂_F coincide (see [13, Remark 1.4]).

Let $S \subseteq X$ and $x \in S$, then $x^* \in X^*$ is *G-normal* to S at x if there exists $\lambda > 0$ such that for any $\varepsilon > 0$ and any finite dimensional subspace $F \subseteq X$, there are $y \in \mathbb{B}(x;\varepsilon)$ and $y^* \in X^*$ such that

$$\langle y^* - x^*, h\rangle \leq \varepsilon\|h\| \text{ and } \langle y^*, h\rangle \leq \mathrm{d}_S^D(x,h),\ \forall h \in F,$$

where $\mathrm{d}_S(y) := \inf_{s\in S} \|s - y\|$ and $\mathrm{d}_S^D(x,h) := \liminf_{r\to 0^+} r^{-1}(\mathrm{d}_S(x+rh) - \mathrm{d}_S(x))$. The set of all G-normals is denoted by $N_G(S,x)$ (see e.g. [62, 63] and the reference therein). The *Clarke-normal set* to S at x is denoted by $N_C(S,x) := \overline{\mathrm{co}}^{w^*} N_G(S,x)$ (see [63]). Then for a function f and $x \in \operatorname{dom} f$ the *G-subdifferential*, the singular G-subdifferential , and the *Clarke subdifferential*[2] , are defined by

$$\begin{aligned}
\partial_G f(x) :=& \{x^* \in X^* : (x^*. -1) \in N_G(\operatorname{epi} f, (x, f(x)))\},\\
\partial_G^\infty f(x) :=& \{x^* \in X^* : (x^*, 0) \in N_G(\operatorname{epi} f, (x, f(x)))\},\\
\partial_C f(x) :=& \{x^* \in X^* : (x^*. -1) \in N_C(\operatorname{epi} f, (x, f(x)))\},
\end{aligned}$$

respectively.

For locally Lipschitzian functions f at $\bar{x}$, the following classical definition of Clarke's subdifferential applies:

$$\partial_C f(\bar{x}) = \{x^* \in X^* \mid \langle x^*, h\rangle \leq f^\circ(\bar{x}; h),\ \forall h \in X\}, \tag{1.1}$$

where

$$f^\circ(\bar{x}; h) := \limsup_{\substack{x\to\bar{x}\\ t\downarrow 0}} \frac{f(x+th) - f(x)}{t}$$

denotes *Clarke's directional derivative* of f at $\bar{x}$ in the direction h.

When the space X is an Asplund space (recall that X is Asplund if and only if every separable subspace of X has separable dual), the *limiting/Mordukhovich subdifferential* and the *singular limiting/Mordukhovich subdifferential* can be defined as (see e.g. [84, 85])

$$\begin{aligned}
\partial_M f(x) :=& \{w^*\text{-}\lim x_n^* : x_n^* \in \partial_F f(x_n),\ \text{and } x_n \xrightarrow{f} x\},\\
\partial_M^\infty f(x) :=& \{w^*\text{-}\lim \lambda_n x_n^* : x_n^* \in \partial_F f(x_n),\ x_n \xrightarrow{f} x \text{ and } \lambda_n \to 0^+\},
\end{aligned}$$

respectively. If $|f(x)| = +\infty$, we set $\partial f(x) := \emptyset$ for any of the previous subdifferentials.

A similar sequential representation for G-subdifferential and the Clarke-Rockafellar subdifferential is given by the following result.

[2]Some authors call Clarke subdifferential, the original definition of Clarke given only for locally Lipschitz functions and they give the name *Clarke-Rockafellar subdifferential* to the extension presented in this Thesis.

Proposition 1.5 *[69, Theorem 1.6 and Theorem 1.8] Assume that either (i) X is β-smooth, or (ii) X is an Asplund space and $\beta = F$. Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be lsc and $x \in \operatorname{dom} f$. Then*

$$\partial_G f(x) = \bigcup_{k>0} \operatorname{cl}^{w^*} \left\{ w^*\text{-}\lim x_n^* : x_n^* \in \partial^-_{\beta,k} f(x_n), \text{ and } x_n \xrightarrow{f} x \right\},$$

$$\partial_G^\infty f(x) = \bigcup_{k>0} \operatorname{cl}^{w^*} \left\{ w^*\text{-}\lim \lambda_n \cdot x_n^* : x_n^* \in \partial^-_{\beta,k} f(x_n), \text{ and } x_n \xrightarrow{f} x,\ \lambda_n \to 0^+ \right\},$$

$$\partial_C f(x) = \overline{\operatorname{co}}^{w^*} \left\{ \tilde{\partial}_G f(x) + \tilde{\partial}_G^\infty f(x) \right\},$$

where

$$\tilde{\partial}_G f(x) := \bigcup_{k>0} \left\{ w^* - \lim x_n^* : x_n^* \in \partial^-_{\beta,k} f(x_n), \text{ and } x_n \xrightarrow{f} x \right\},$$

$$\tilde{\partial}_G^\infty (x) := \bigcup_{k>0} \left\{ w^* - \lim \lambda_n \cdot x_n^* : x_n^* \in \partial^-_{\beta,k} f(x_n), \text{ and } x_n \xrightarrow{f} x,\ \lambda_n \to 0^+ \right\}.$$

It is important to recall when f is convex, proper and lsc all of these subdifferentials coincide with the convex subdifferential.

1.3 Epi-pointed functions

Before presenting the definition of an *epi-pointed function*, which plays a key role in our analysis, let us introduce the history about this family and some properties in finite-dimension which motivate our definition.

The class of epi-pointed functions was introduced in finite dimension by Hiriart-Urruty and Benoist [6] in the nineties, but the original definition goes back to the Nobel laureate mathematical economist Gérard Debreu in the fifties [39].

To introduce this class of functions, the authors used the notion of asymptotic function (also in [106, Chapter 3, Definition 3.17] called the horizon function) f^∞, for a proper function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, defined by

$$f^\infty(x) := \liminf_{\substack{t \to 0^+ \\ y \to x}} t f(t^{-1} y).$$

So, f is said to be epi-pointed if the epigraph of its asymptotic function (which is clearly a cone) is pointed, that is,

$$\text{for all } \xi_1, ..., \xi_p \in \operatorname{epi} f^\infty \text{ and } \xi_1 + ... + \xi_p = 0 \implies \xi_{\mathrm{i}} = 0 \text{ for all } \mathrm{i} = 1, ..., p.$$

However, the original definition given in [39] was only for sets: roughly speaking, if $S \subseteq \mathbb{R}^n$ is a closed set, the asymptotic cone of S is defined as

$$S^\infty := \{x \in \mathbb{R}^n : \text{there exist } t_n \to 0^+, x_n \in S \text{ such that } t_n x_n \to x\}. \tag{1.2}$$

Then, S is termed pointed if the asymptotic cone S^∞ contains no straight line.

The following proposition allows us to appreciate better the definition of an epi-pointed function.

Proposition 1.6 *Consider $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ lsc and proper. Then the following statements are equivalent:*

(a) f is epi-pointed.

(b) There exist $\bar{u} \in \mathbb{R}^n$, $\alpha > 0$ and $r \in \mathbb{R}$ such that:

$$f(x) \geq \langle \bar{u}, x \rangle + \alpha \|x\| + r \quad \forall x \in \mathbb{R}^n$$

(c) There exists $\bar{u} \in \mathbb{R}^n$ such that f^* is bounded from above on a neighborhood of $\bar{u}$.

Proof. First we will prove that

$$\liminf_{\|x\| \to \infty} \frac{f(x)}{\|x\|} = \inf_{\|x\|=1} f^\infty(x). \tag{1.3}$$

Indeed, consider sequences (x_k), t_k, y_k and points w_0, w_1 such that:

- $\|x_k\| \to +\infty$, $\liminf\limits_{\|x\| \to \infty} \frac{f(x)}{\|x\|} = \lim \frac{f(x_k)}{\|x_k\|}$ and $\frac{x_k}{\|x_k\|} \to w_0$.
- $t_k \to 0^+$, $y_k \to w_1$ and $\inf\limits_{\|x\|=1} f^\infty(x) = f^\infty(w_1) = \lim t_k f(\frac{y_k}{t_k})$ (f^∞ is lsc).

Then

$$\begin{aligned}\liminf_{\|x\| \to \infty} \frac{f(x)}{\|x\|} &= \lim \frac{f(x_k)}{\|x_k\|} = \lim \frac{f(\frac{x_k}{\|x_k\|}\|x_k\|)}{\|x_k\|} \geq f^\infty(w_0) \geq f^\infty(w_1) \\ &= \lim t_k f(\frac{y_k}{t_k}) = \lim \|y_k\| \frac{t_k}{\|y_k\|} f(\frac{y_k}{t_k}) \geq \liminf_{\|x\| \to \infty} \frac{f(x)}{\|x\|}.\end{aligned}$$

Now we will prove the proposition:

- $(a) \Rightarrow (b)$ Suppose that f is epi-pointed. Then $\inf\limits_{\|x\|=1} f^\infty(x) =: \gamma > -\infty$ (if this does not happen, $\{0\} \times \mathbb{R} \subset \operatorname{epi} f^\infty$). Now we will show that there exists $\bar{u} \in \mathbb{R}^n$ such that for every $\|x\| = 1$, $f^\infty(x) > \langle \bar{u}, x \rangle$. In fact, consider the set $K := \overline{\text{co}}\{(x, \alpha) \in \operatorname{epi} f^\infty : \|x\| = 1\}$. We have $(0,0) \notin K$, because if this is not true, then there would exist $(x_n, \alpha_n) = \sum_{\mathrm{i}=0}^n \lambda_{\mathrm{i}}^n (x_{\mathrm{i}}^n, \beta_{\mathrm{i}}^n)$ (n is fixed, by Carathéodory's Theorem) with $(x_{\mathrm{i}}^n, \beta_{\mathrm{i}}^n) \in K$, $\lambda_{\mathrm{i}}^n \geq 0$ and $\sum_{\mathrm{i}=0}^n \lambda_{\mathrm{i}}^n = 1$ with $(x_n, \alpha_n) \to (0,0)$. By taking a subsequence we can suppose that for every i, $x_{\mathrm{i}}^n \to x_{\mathrm{i}}$ and $\lambda_{\mathrm{i}}^n \to \lambda_{\mathrm{i}}$; moreover, because $\beta_{\mathrm{i}}^n \geq \gamma$, we conclude that $\lambda_{\mathrm{i}}^n \beta_{\mathrm{i}}^n \to \beta_{\mathrm{i}}$ (taking another subsequence). From the fact that f^∞ is lsc and positively homogeneous, we get $(\lambda_{\mathrm{i}} x_{\mathrm{i}}, \beta_{\mathrm{i}}) \in \operatorname{epi} f^\infty$, that is to say, there are $(w_{\mathrm{i}}, \alpha_{\mathrm{i}}) \in \operatorname{epi} f^\infty$ (not all identically zero) such that $(w_0, \alpha_0) + \ldots + (w_n, \alpha_n) = 0$, which contradicts the epi-pointed assumption.

Now we apply the Hahn-Banach Theorem (if $K = \emptyset$ the result is trivial) to conclude the existence of $w \in \mathbb{R}^n$, $\mathrm{d} \in \mathbb{R}$ and η such that $\mathrm{d}\alpha + \langle w, x\rangle > \eta > 0$ for every $(x, \alpha) \in K$, then necessarily $\mathrm{d} \geq 0$. So, taking $\bar{u} = -\mathrm{d}^{-1}w$ if $\mathrm{d} > 0$ or $\bar{u} = -2\frac{|\gamma|}{\eta}w$ if $\mathrm{d} = 0$, we get $\inf\limits_{\|x\|=1}(f - \bar{u})^\infty(x) = \inf\limits_{\|x\|=1} f^\infty(x) - \langle \bar{u}, x\rangle > 0$. Now using Equation 1.3 we conclude

$$\liminf_{\|x\|\to\infty} \frac{f(x) - \langle \bar{u}, x\rangle}{\|x\|} > 0.$$

Then there are $\alpha > 0$ and $M > 0$ such that $f(x) \geq \langle \bar{u}, x\rangle + \alpha\|x\|$ for every $\|x\| \geq M$. So, taking $r := \min\limits_{w\in\mathbb{B}(0,M)} \{f(w) - \langle \bar{u}, w\rangle - \alpha\|w\|, 0\}$ (because f is lsc and proper) we conclude (b).

- $(b) \Rightarrow (a)$ Suppose there exists $\{(x_{\mathrm{i}}, \alpha_{\mathrm{i}})\}_{\mathrm{i}=0}^p \subset \operatorname{epi} f^\infty$ (not all identically zero) such that $(x_0, \alpha_0) + ... + (x_p, \alpha_p) = 0$. Then, taking $h(x) = f(x) - \langle \bar{u}, x\rangle$ we have

$$\liminf_{\|x\|\to\infty} \frac{h(x)}{\|x\|} > 0,$$

and so we get

$$\inf_{\|x\|=1} h^\infty(x) = \inf_{\|x\|=1} f^\infty(x) - \langle \bar{u}, x\rangle > 0.$$

From the fact $\sum\limits_{\mathrm{i}=0}^p \{f^\infty(x_{\mathrm{i}}) - \langle \bar{u}, x_{\mathrm{i}}\rangle\} \leq \sum\limits_{\mathrm{i}=0}^p \alpha_{\mathrm{i}} - \sum\limits_{\mathrm{i}=0}^p \langle \bar{u}, x_{\mathrm{i}}\rangle = 0$, there exists some $x_{\mathrm{i}} \neq 0$ such that $f^\infty(x_{\mathrm{i}}) - \langle \bar{u}, x_{\mathrm{i}}\rangle \leq 0$, and then $h^\infty(\frac{x_{\mathrm{i}}}{\|x_{\mathrm{i}}\|}) \leq 0$, which is a contradiction.

- $(b) \Rightarrow (c)$ Take an arbitrary $w \in \mathbb{B}(0, \alpha)$. Then for every $x \in \mathbb{R}^n$, $\langle w, x\rangle - \alpha\|x\| - r \geq \langle \bar{u} + w, x\rangle - f(x)$. Therefore, $f^*(\bar{u} + w) \leq -r$ for every $w \in \mathbb{B}(0, \alpha)$.

- $(c) \Rightarrow (b)$ Let $M \geq 0$ and $\alpha > 0$ such that $f^*(\bar{u} + w) \leq M$ for all $w \in \mathbb{B}(0, \alpha)$. Then $\langle \bar{u} + w, x\rangle - f(x) \leq M$ for every $w \in \mathbb{B}(0, \alpha)$ and every $x \in \mathbb{R}^n$. Then taking $w := \alpha\frac{x}{\|x\|}$ we conclude (b).

$\square$

This equivalence shows that the property of epi-pointedness is characterized by the continuity of f^* at some point $\bar{u}$ and it justifies the definition for functions defined on a locally convex space X, that we will adopt in this thesis.

Definition 1.7 *Let X, X^* be two lcs in duality. A function $f : X \to \overline{\mathbb{R}}$ is said to be epi-pointed if f^* is proper and $\tau(X^*, X)$-continuous at some point of its domain.*

Next, we give typical examples that illustrate the amplitude of this class of functions.

Example 1.8 Let $f : \mathbb{R} \to \mathbb{R}$ be a function such that $f(0) = 0$ and the closed convex envelope of f is positive; that is to say, $\overline{\mathrm{co}}\, f(x) > 0$ for all $x \in \mathbb{R}\backslash\{0\}$. Then f is an epi-pointed function.

Example 1.9 Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a 1-coercive function; that is, $\|x\|^{-1} f(x) \to +\infty$ as $\|x\| \to +\infty$. Then, f is epi-pointed.

Example 1.10 Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a proper lsc convex function. Then, for every $\alpha > 0$, the function $x \to f(x) + \alpha\|x\|^2$ is epi-pointed.

Example 1.11 Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ proper and prox-bounded function; that is, $f \geq -\mu\|x\|^2$ for some $\mu \geq 0$. Then for every $\varepsilon > 0$ the function $f + (\mu + \varepsilon)\| \cdot \|^2$ is epi-pointed.

Example 1.12 For every function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ and every bounded set $C \subset X$ such that $C \cap \operatorname{dom} f \neq \emptyset$ and f is minorized on C by a continuous affine form; the function $f + \delta_C$ is epi-pointed.

Example 1.13 If $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is minorized by a continuous affine form and is bounded from above in a neighborhood of some point, then $f^* : X^* \to \overline{\mathbb{R}}$ is epi-pointed.

Epi-pointed functions are expected to share many useful properties with convex functions. This class of functions has been successfully utilized recently (among other articles we refer to [5], [25], [26], [27], [28], [31], [33], [92]) with the purpose of extending results, which were known exclusively for Banach spaces or convex functions, to arbitrary locally convex spaces or for nonconvex functions respectively. In this scenario we can underline two important topics. The first concerns the extension of *Fenchel's duality*. In this respect, we recall the following result, which relies on the relationship between the convex subdifferential of f and its conjugate f^*, more precisely, one has

$$x^* \in \partial f(x) \text{ if and only if } x \in \partial f^*(x^*).$$

The result below extends in a fashionable way the Fenchel's well-known Theorem. The following result was established in a series of paper by the authors ([25–28]) under different stages of generality.

Proposition 1.14 [26, Corollary 6] *Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a weakly lsc epi-pointed function. Then for every $x^* \in X^*$ we have that*

$$\partial f^*(x^*) = \overline{\operatorname{co}}^{w^*} \left[(\partial f)^{-1}(x^*)\right] + N_{\operatorname{dom} f^*}(x^*).$$

The second one corresponds to problems related to recovering a function from its subdifferential; in the literature these topics are called problems of *integration of subdifferentials*. Determining a function from its first-order variations is a fundamental principle in nonlinear analysis. It is a very elemental fact in the first course of calculus that given two continuously differentiable functions $h, g : \mathbb{R}^n \to \mathbb{R}$ such that $\nabla h(x) = \nabla g(x)$ for all $x \in \mathbb{R}^n$, then the functions are equal up to a constant. This question becomes more involved when the nominal function fails to be differentiable. The problem involving convex functions was started and solved by Moreau in Hilbert spaces [87]. Posteriorly, Rockafellar [98, 100] in the sixties took the problem and asserted that every two proper lsc convex functions f, g such that $\partial f(x) \subseteq \partial g(x), \forall x \in \mathbb{R}^n$ must be equal up to a constant. With the passing of the years many authors have studied these problems with varying degrees of abstraction passing even to the nonconvex case, among other works we refer to [5, 25, 33, 34, 118, 119] where variants of the following result have been proved.

Proposition 1.15 *Let $f : X \to \overline{\mathbb{R}}$ be a weakly lsc epi-pointed function and let $g : X \to \overline{\mathbb{R}}$ be any function such that*

$$\partial f(x) \subseteq \partial g(x),\ \forall x \in X.$$

Then there exists $c \in \mathbb{R}$ such that $\overline{\operatorname{co}} f = \overline{\operatorname{co}}\, g \Box \sigma_{\operatorname{dom} f^*} + c$.

The following lemma is a compilation of classical results in convex analysis that we will use in Chapter 2, 3 and 4. They can be found in the pioneer reference of convex analysis [88] and also in [77, chapter 6].

Lemma 1.16 (a) *Given two lsc proper convex functions $g, h : X \to \overline{\mathbb{R}}$ such that g^* is continuous at some point of* $\operatorname{dom} h^*$*, then for all $x \in X$ there exist $x_1, x_2 \in X$ such that $x_1 + x_2 = x$ and*

$$\partial g(x_1) \cap \partial h(x_2) = \partial(g \Box h)(x_1 + x_2)$$

(b) *Given two functions $g, h : X \to \overline{\mathbb{R}}$, we have for all $x_1, x_2 \in X$*

$$\partial g(x_1) \cap \partial h(x_2) \subseteq \partial(g \Box h)(x_1 + x_2).$$

(c) *Given two functions $g, h : X \to \overline{\mathbb{R}}$, we have $(g \Box h)^* = g^* + h^*$.*

(d) *Let $g, h : X^* \to \overline{\mathbb{R}}$ be proper convex functions such that h is continuous at some point in* $\operatorname{dom} g$*. Then*

$$(g + h)^*(x) = (g^* \Box h^*)(x)\ \ \forall x \in X,$$

and the inf-convolution is exact.

(e) *Given a function $g : X \to \overline{\mathbb{R}}$, if $\partial g(x) \neq \emptyset$, then $g(x) = g^{**}(x)$.*

(f) *A lsc proper convex function $g : X^* \to \overline{\mathbb{R}}$ is $\tau(X^*, X)$-continuous at $x^* \in \operatorname{dom} g$ if and only if $g^* - x^*$ is $w(X, X^*)$-inf-compact. Consequently a w-lsc epi-pointed function f satisfies that $f - x^*$ is w-infcompact for every $x^* \in \operatorname{int}(\operatorname{dom} f^*)$.*

1.4 Measure theory

Through this thesis $(T, \mathcal{A}, \mu)$ will be a complete (non-negative) σ-finite measure space. One special case considered in this thesis is $(T, \mathcal{A}) = (\mathbb{N}, \mathcal{P}(\mathbb{N}))$, which will be crucial to give formulae for the case of finite or infinite sums. The notation $\mathcal{B}(X, \tau)$ ($\mathcal{B}(X)$ if there is no confusion about the topology τ) will be reserved for the σ-algebra generated by the topology τ. For a subset $M \subseteq \mathbb{R}$ we denote by $\mathrm{L}^1(T, M)$ the set of all integrable function from T to M. Given a function $f : T \to \overline{\mathbb{R}}$, we denote

$$\mathcal{D}_f := \{g \in \mathrm{L}^1(T, \mathbb{R}) : f(t) \leq g(t)\ \mu\text{-almost everywhere}\},$$

and define the *upper integral* of f by

$$\int_T f(t)\mathrm{d}\mu(t) := \inf_{g\in\mathcal{D}_f} \int_T g(t)\mu(t) \tag{1.4}$$

whenever $\mathcal{D}_f \neq \emptyset$. If $\mathcal{D}_f = \emptyset$, we set $\int_T f(t)\mathrm{d}\mu(t) := +\infty$.

Definition 1.17 (Poilsh space) *A topological space P is said to be a Polish space if it is complete, metrizable and separable.*

Definition 1.18 (Suslin space [15, 22, 109]) *A Hausdorff topological space S is said to be a Suslin space if there exist a Polish space P and a continuous surjection from P to S.*

A function $f : T \to U$, with U being a topological space, is called *simple* if there are $k \in \mathbb{N}$, a partition $T_\mathrm{i} \in \mathcal{A}$ and elements $x_\mathrm{i} \in X$, $\mathrm{i} = 0, ..., k$, such that $f = \sum_{\mathrm{i}=0}^{k} x_\mathrm{i} \mathbb{1}_{T_\mathrm{i}}$ here, $\mathbb{1}_{T_\mathrm{i}}$ denotes the characteristic function (or the indicator function in the sense of measure theory) of T_i, equaling to 1 in T_i and 0 outside. A function f is called strongly measurable (measurable, for short) if there exists a countable family f_n of simple functions such that $f(t) = \lim_{n\to\infty} f_n(t)$ for almost every (ae, for short) $t \in T$. A strongly measurable function $f : T \to X$ is said to be strongly integrable, and we write $f \in \mathcal{L}^1(T, X)$, if $\int_T \sigma_B(f(t))\mathrm{d}\mu(t) < \infty$ for every bounded balanced subset $B \subset X^*$. Observe that in the Banach spaces setting $\mathcal{L}^1(T, X)$ is the set of Bochner integrable functions (see, e.g., [42, Chapter II]).

A function $f : T \to X$ is called (weakly or scalarly integrable) weakly or scalarly measurable if for every $x^* \in X^*$, $t \to \langle x^*, f(t)\rangle$ is (integrable, resp.,) measurable. We denote $\mathcal{L}^1_w(T, X)$ the space of all weakly integrable functions f such that $\int_T \sigma_B(f(t))\mathrm{d}\mu(t) < \infty$ for every bounded balanced subset $B \subseteq X^*$. Similarly, for functions taking values in X^*, $f : T \to X^*$, if for every $x \in X$, the mapping $t \to \langle x, f(t)\rangle$ is (integrable, resp.,) measurable, then we say that f is (w*-integrable, resp.) w*-measurable. Also, we denote $\mathcal{L}^1_{w^*}(T, X^*)$ the space of all w*-integrable functions f such that $\int_T \sigma_B(f(t))\mathrm{d}\mu(t) < \infty$ for every bounded balanced subset $B \subseteq X$.

It is clear that every strongly integrable function is weakly integrable. However, the weak measurability of a function f does not necessarily imply the measurability of the function $\sigma_B(f(\cdot))$, and so the corresponding integral of this last function must be understood in the sense of (1.4) (see [68, Chapter VI] where the integral of the norm of a function f is understood in this sense). Also, observe that if in addition X is a Suslin space, then every $(\mathcal{A}, \mathcal{B}(X))$-measurable function $f : T \to X$ (that is, $f^{-1}(B) \in \mathcal{A}$ for all $B \in \mathcal{B}(X)$) is weakly measurable, where $\mathcal{B}(X)$ is the Borel σ-Algebra of the open (equivalently, weak open) set of X (see, e.g., [22, Theorem III.36]).

The quotient spaces $\mathrm{L}^1(T, X)$ and $\mathrm{L}^1_w(T, X)$ of $\mathcal{L}^1(T, X)$ and $\mathcal{L}^1_w(T, X)$, respectively, are those given with respect to the equivalence relations $f = g$ ae, and $\langle f, x^*\rangle = \langle g, x^*\rangle$ ae for all $x^* \in X^*$, respectively (see, for example, [68]). In a similar way one defines the spaces $\mathrm{L}^1(T, X^*)$ and $\mathrm{L}^1_{w^*}(T, X^*)$.

It is worth observing that when X is a separable Banach space, both notions of (strong and weak) measurability coincide; hence, if, in addition, $(X^*, \|\cdot\|)$ is separable, then $\mathcal{L}^1(T, X^*) = \mathcal{L}^1_{w^*}(T, X^*)$ (see [42, Chapter II, Theorem 2]). It will be more clarifying recalling that when the space X is separable, but the its dual X^* is not $\|\cdot\|$-separable $\mathcal{L}^1(T, X^*)$ and $\mathcal{L}^1_{w^*}(T, X^*)$ may not coincide (see [42, Chapter II Example 6]). For every w*-integrable function $f : T \to X^*$ and every $E \in \mathcal{A}$, the function $x^\sharp_E$ defined on X as $x^\sharp_E(x) := \int_E \langle f, x\rangle \mathrm{d}\mu$ is a linear mapping (not necessary continuous), which we call the weak integral of f over E, and we write $\int_E f\mathrm{d}\mu := x^\sharp_E$ (see [16, 17]). Moreover, if f is strongly integrable, this element $\int_E f\mathrm{d}\mu$ also refers to the strong integral of f over E. Observe that, in general, $\int_E f\mathrm{d}\mu$ may not be in X^*. However, when the space X is Banach, and function $f : T \to X^*$ is w*-integrable, $\int_E f\mathrm{d}\mu \in X^*$ and is called the Gelfand integral of f over E (see [42, Chapter II, Lemma 3.1] and details therein).

The space $\mathrm{L}^\infty(T, X)$ ($\mathrm{L}^\infty_{w^*}(T, X)$ respectively) is the set of all (equivalence classes) measurable functions (w^*-measurable functions, respectively) $f : T \times X$ such that $f(T)$ is essentially bounded . When X is Banach, $\mathrm{L}^\infty(T, X)$ is the classical normed space and its norm is given by $\|x\|_\infty := \operatorname{ess\,sup}\{\|x(t)\| : t \in T\} < \infty$. A functional $\lambda^* \in \mathrm{L}^\infty(T, X)^*$ is called singular if there exists a sequence of measurable sets T_n such that $T_{n+1} \subseteq T_n$, $\mu(T_n) \to 0$ as $n \to \infty$ and $\lambda^*(g\mathbb{1}_{T_n^c}) = 0$ for every $g \in \mathrm{L}^\infty(T, X)$. We will denote $\mathrm{L}^{\mathrm{sing}}(T, X)$ the set of all singular functionals. It is well-known that each functional $\lambda^* \in \mathrm{L}^\infty(T, X)^*$ can be uniquely written as the sum $\lambda^*(\cdot) = \int_T \langle \lambda_1^*(t), \cdot\rangle \mathrm{d}\mu(t) + \lambda_2^*(\cdot)$, where $\lambda_1^* \in \mathrm{L}^1_{w^*}(T, X^*)$ and $\lambda_2^* \in \mathrm{L}^{\mathrm{sing}}(T, X)$ (see, for example, [22, 79]). In a similar way the space $\mathrm{L}^\infty_{w^*}(T, X^*)$ is defined as the set of all weakly* measurable essentially bounded function from T to X^*.

When (X, τ) is a lcs, we say that a net of weakly measurable functions $(g_{\mathrm{i}})_{\mathrm{i}\in I} \subset X^T$ converges uniformly τ-ae (or τ-essentially uniformly, also essentially uniformly) if, for every τ-continuous seminorm ρ on X it holds $\|\ \rho(g_{\mathrm{i}} - g)\|_\infty \to 0$.

A family of functions $(\varphi_{\mathrm{i}})_{\mathrm{i}\in\mathbb{I}} \subseteq \mathrm{L}^1(T, \mathbb{R})$ is said to be uniformly integrable if

$$\lim_{a\to\infty} \sup_{\mathrm{i}\in\mathbb{I}} \int\limits_{\{|\varphi_{\mathrm{i}}(t)|\geq a\}} |\varphi_{\mathrm{i}}(t)|\mathrm{d}\mu(t) = 0.$$

When the space X will be a Banach space we will also need the following L^p spaces. For $p \in [1, \infty)$ we denote by $\mathrm{L}^p(T, X)$ and $\mathrm{L}^p_{w^*}(T, X^*)$ the sets of all (equivalence classes by the relation $f = g$ ae) measurable and w^*-measurable functions f such that $\|f(\cdot)\|^p$ is integrable, as usual the norm in these spaces is $\|f\|_p := (\int_T \|f(t)\|^p \mathrm{d}\mu(t))^{1/p}$.

The next definition corresponds to the notion of an *integral functional.*

Definition 1.19 *For a vector space L of function $x : T \to X$, where X is endowed with a lcs topology τ, by an integral functional on L we mean an extended-real-valued functional $\hat{I}_f$ of the form*

$$x(\cdot) \in L \to \hat{I}^{\mu,L}_f(x(\cdot)) := \int\limits_T f(t, x(t))\mathrm{d}\mu(t), \tag{1.5}$$

where $f : T \times X \to \overline{\mathbb{R}}$ is any function and the integral is considered in the sense of (1.4). The notation $\hat{I}^{\mu,p}_f$ will be reserved for the integral functional, when it is defined over $\mathrm{L}^p(T, X)$.

When there is not an ambiguity we simply write $\hat{I}_f$.

The next definition corresponds to the classical notion of *normal integrands* and *convex normal integrands.*

Definition 1.20 *A function $f : T \times X \to \overline{\mathbb{R}}$ is called a τ-normal integrand (or, simply, normal integral when no confusion occurs), if f is $\mathcal{A} \otimes \mathcal{B}(X, \tau)$-measurable and the functions $f(t, \cdot)$ are lsc for ae $t \in T$. In addition, if $f(t, \cdot) \in \Gamma_0(X)$ for ae $t \in T$, then f is called convex normal integrand. For simplicity, we denote $f_t := f(t, \cdot)$.*

When L is the linear space of constant functions, we also consider the integral function I_f defined on X as

$$x \in X \to I_f(x) := \int_T f(t, x) \mathrm{d}\mu(t).$$

A multifunction $G : T \rightrightarrows X$ is called $\mathcal{A}$-$\mathcal{B}(X)$-measurable (measurable, for simplicity) if its graph, $\operatorname{gph} G := \{(t, x) \in T \times X : x \in G(t)\}$, is an element of $\mathcal{A} \otimes \mathcal{B}(X)$. We say that G is weakly measurable if for every $x^* \in X^*$, $t \to \sigma_{G(t)}(x^*)$ is a measurable function.

When the space X is a Suslin space the literature provides many powerful results. Among these one finds the theorems concerning measurable selections (see for example [22]). However, the theory outside of separable spaces has not prospered as much as the theory in Suslin spaces. A new result concerning *weakly**-measurable selections in nonseparable Asplund spaces was proved in [20]. The next two theorems deal with measurable selections in both Suslin spaces and non-separable Banach spaces.

Proposition 1.21 *[22, Theorem III.22] Let S be a Suslin space and $G : T \rightrightarrows S$ be a measurable multifunction with non-empty values. Then there exists a sequence (g_n) of $(\mathcal{A}, \mathcal{B}(S))$-measurable selections of $G(t)$ such that $\{g_n(t)\}_{n \geq 1}$ is dense in $G(t)$ for every $t \in T$.*

Proposition 1.22 *[20, Corrolary 3.11] Assume that $(T, \mathcal{A}, \mu)$ is finite (complete measure), and assume that X is Asplund. Then every w^*-measurable multifunction $C : T \rightrightarrows X^*$ with nonempty and weak*-compact values admits a w^*-measurable selection.*

Definition 1.23 *For a (non-necessarily measurable) multifunction $G : T \rightrightarrows X^*$ the strong and the weak integrals of M are defined by*

$$\int_T G(t) \mathrm{d}\mu(t) := \left\{ \int_T m(t) \mathrm{d}\mu(t) \in X^* : m \text{ is strong integrable and } m(t) \in G(t) \text{ ae} \right\},$$

$$(w)\text{-} \int_T G(t) \mathrm{d}\mu(t) := \left\{ \int_T m(t) \mathrm{d}\mu(t) \in X^* : m \text{ is } w^*\text{-integrable and } m(t) \in G(t) \text{ ae} \right\}.$$

In Chapter 6 and Chapter 7 the integral of mutifunctions will be understood as

$$\int_T G(t) \mathrm{d}\mu(t) := \left\{ \int_T m(t) \mathrm{d}\mu(t) \in X^* : m \in \mathrm{L}^1_{w^*}(T, X^*) \text{ and } m(t) \in G(t) \text{ ae} \right\}. \tag{1.6}$$

It is important recalling that the original definition of integral of set-valued mappings is due to R. J. Aumann and it was given for multifunction defined on a closed interval $[0, T]$ in $\mathbb{R}$; see for example [3]. For this reason many authors give the name of *Aumann Integral* to (1.6).

Chapter 2

On the Klee-Saint Raymond's Characterization of Convexity

2.1 Introduction

J. Saint Raymond observes [107] that for a given nonconvex continuous function $f : \mathbb{R} \to \mathbb{R}$, which satisfies $\lim\limits_{|x| \to +\infty} \frac{f(x)}{|x|} = +\infty$, there exists an affine function h that minorizes f, such that $f - h$ vanishes on a nonconvex set. This fact characterizes the convexity of a function. More generally, in [107, Theorem 10] the author stipulated:

Theorem 2.1 *Let X be a Banach space and let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a weakly lower semicontinuous proper function such that $f - x^*$ is weakly inf-compact for all $x^* \in X^*$. If the argmin set of the function $f - x^*$ is convex for all x^* in a convex dense subset of X^*, then f is convex.*

Observe that in the original statement of [107, Theorem 10] the hypothesis of weak inf-compactness used above is replaced by the equivalent fact that the function $f - x^*$ attains its minimum, for every $x^* \in X^*$. This equivalence, being a functional counterpart of James Theorem ([46, Theorem 3.130]), has been established in [108, Theorem 2.4]. We call this work Klee-Saint Raymond characterization of convexity because in the framework of Hilbert spaces, Theorem 2.1 is equivalent to the famous characterization given by Klee [75] for the convexity of weakly closed sets. See [9] for a recent review of this problem related to Chebychev sets.

To prove Theorem 2.1 the author uses classical deep tools of Banach space theory, like James' theorem and Brouwer's fixed-point theorem for multifunctions, among others. More recently, another proof has been given in [97], under the assumption that X is a reflexive Banach space, by using techniques of operator theory.

In this work we use techniques of convex analysis to give a direct proof for a generalization of Theorem 2.1 for functions defined on locally convex spaces. This generalization, given in Corollary 2.8, is an immediate consequence of the main result of this work that provides an explicit expression of the closed convex hull of a function; see Theorem 2.5. Our hypotheses are weaker than those used in Theorem 2.1, and rely on the epi-pointedness property that has been successfully utilized recently for many authors with the purpose of extending results, which were known exclusively for Banach spaces or convex functions, to locally convex spaces and nonconvex functions.

2.2 The characterization of convexity

First we start proving the following lemma, which is a slight extension of [2, Theorem 2.40] to the case of nets of functions. For completeness we give a proof.

Lemma 2.2 *Let X be a topological space and let $(f_\alpha)_{\alpha\in D}$ be a net of lsc proper functions defined on X such that*

$$\alpha_1, \alpha_2 \in D,\ \alpha_1 \le \alpha_2 \Rightarrow f_{\alpha_1} \le f_{\alpha_2}. \tag{2.1}$$

For $\varepsilon_\alpha \underset{\alpha\in D}{\to} 0^+$, let $(x_\alpha)_{\alpha\in D}$ be a relatively compact net such that $x_\alpha \in \varepsilon_\alpha\text{-}\operatorname{argmin} f_\alpha$ for each α. Then

$$\inf_{x\in X} \sup_{\alpha\in D} f_\alpha(x) = \sup_{\alpha\in D} \inf_{x\in X} f_\alpha(x),$$

and every accumulation point of (x_α) is a minimizer of the function $\sup_{\alpha\in D} f_\alpha$.

Proof. It is easy to see that every subnet of (f_α) has a subnet that preserves property (2.1); so, without loss of generality, we may assume that $x_\alpha \to \bar{x} \in X$. We start by showing that for every $V \in \mathcal{N}_{\bar{x}}$, the neighborhood system of $\bar{x}$, we have

$$\sup_{\alpha\in D} \inf_{v\in V} f_\alpha(v) \le \sup_{\alpha\in D} \inf_{x\in X} f_\alpha(x). \tag{2.2}$$

Given $\delta > 0$ and $V \in \mathcal{N}_{\bar{x}}$, we choose $\alpha_0 \in D$ such that $x_\alpha \in V$ and $\varepsilon_\alpha < \delta$, for all $\alpha \ge \alpha_0$. Then for any $\eta \in D$ we get

$$\inf_{v\in V} f_\eta(v) \le \sup_{\alpha\ge\alpha_0} f_\alpha(x_\alpha) \le \sup_{\alpha\ge\alpha_0} \left\{ \inf_{x\in X} f_\alpha(x) + \varepsilon_\alpha \right\} \le \sup_{\alpha\in D} \inf_{x\in X} f_\alpha(x) + \delta.$$

This yields (2.2) by taking the supremum on $\eta \in D$ and the limit as $\delta \to 0^+$.

Now, (2.2) leads us to

$$\inf_{x\in X} \sup_{\alpha\in D} f_\alpha(x) \le \sup_{\alpha\in D} f_\alpha(\bar{x}) = \sup_{\alpha\in D} \sup_{V\in\mathcal{N}_{\bar{x}}} \inf_{v\in V} f_\alpha(v) = \sup_{V\in\mathcal{N}_{\bar{x}}} \sup_{\alpha\in D} \inf_{v\in V} f_\alpha(v) \le \sup_{\alpha\in D} \inf_{x\in X} f_\alpha(x),$$

which yields

$$\inf_{x\in X} \sup_{\alpha\in D} f_\alpha(x) \le \sup_{\alpha\in D} f_\alpha(\bar{x}) \le \sup_{\alpha\in D} \inf_{x\in X} f_\alpha(x) \le \inf_{x\in X} \sup_{\alpha\in D} f_\alpha(x),$$

and, so, the proof is completed. □

We continue with a comparison between the subdifferentials of an epi-pointed function and its biconjungate.

Proposition 2.3 *Let $f : X \to \overline{\mathbb{R}}$ be a weakly lsc epi-pointed function and denote*

$$M_f := \{x^* \in X^* : \operatorname{argmin}\{f - x^*\} \text{ is convex}\}.$$

Then for every $x \in X$ we have that

$$\operatorname{int}(\operatorname{dom} f^*) \cap M_f \cap \partial f^{**}(x) \subseteq \partial f(x).$$

Proof. We choose $x^* \in \operatorname{int}(\operatorname{dom} f^*) \cap M_f \cap \partial f^{**}(x)$ and $x \in X$. Since $(\partial f)^{-1}(x^*) = \operatorname{argmin}\{f - x^*\}$ is convex and weakly closed (f is weakly lsc), according to Proposition 1.14 we have that

$$\partial f^*(x^*) = (\partial f)^{-1}(x^*).$$

Hence, since $x^* \in \partial f^{**}(x)$ we have that $x \in \partial f^*(x^*) = (\partial f)^{-1}(x^*)$, which is equivalent to $x^* \in \partial f(x)$. □

We give now a first relation between an epi-pointed function and its biconjugate.

Proposition 2.4 *Let $f : X \to \overline{\mathbb{R}}$ be a weakly lsc epi-pointed function and M_f as in proposition above. Then, for all nonempty, convex and compact set $C \subset \operatorname{int}(\operatorname{dom} f^*) \cap M_f$, we have that*

$$\sigma_C \Box f^{**} = \sigma_C \Box f.$$

Proof. We fix $x \in X$. By Lemma 1.16(a), applied with $g := \sigma_C$ and $h := f^{**}$, since $h^* = f^*$ is continuous at any point of $\operatorname{dom} g^* = C$, there are $x_1, x_2 \in X$ such that $x = x_1 + x_2$ and

$$\partial \sigma_C(x_1) \cap \partial f^{**}(x_2) = \partial(\sigma_C \Box f^{**})(x). \tag{2.3}$$

Since $\partial \sigma_C(x_1) \subseteq C \subseteq \operatorname{int}(\operatorname{dom} f^*) \cap M_f$, Proposition 2.3 gives us the relation

$$\partial \sigma_C(x_1) \cap \partial f^{**}(x_2) \subseteq \partial \sigma_C(x_1) \cap \partial f(x_2).$$

So, invoking Lemma 1.16(b), from (2.3) we infer that

$$\partial(\sigma_C \Box f^{**})(x) \subseteq \partial(\sigma_C \Box f)(x). \tag{2.4}$$

On the one hand, by Lemma 1.16(c) we have $(\sigma_C \Box f)^* = f^* + \delta_C$ and so, invoking Lemma 1.16(d), we get

$$(\sigma_C \Box f)^{**} = \sigma_C \Box f^{**}. \tag{2.5}$$

On the other hand, since

$$(\sigma_C \Box f)^{**}(x) = (\delta_C + f^*)^*(x) = \sup_{x^* \in C} \{\langle x^*, x\rangle - f^*(x^*)\},$$

the continuity of f^* and the compactness of C yield the existence of some $\overline{x}^* \in C$ such that

$$(\sigma_C \Box f)^{**}(x) = \langle \overline{x}^*, x\rangle - f^*(\overline{x}^*) = \langle \overline{x}^*, x\rangle - (f^* + \delta_C)(\overline{x}^*) = \langle \overline{x}^*, x\rangle - (\sigma_C \Box f)^*(\overline{x}^*).$$

In other words, in view of (2.5) and (2.4), respectively,

$$\overline{x}^* \in \partial \left(\sigma_C \Box f\right)^{**}(x) = \partial \left(\sigma_C \Box f^{**}\right)(x) \subseteq \partial(\sigma_C \Box f)(x)$$

and, so, $\partial(\sigma_C \Box f)(x) \neq \emptyset$. Due to Lemma 1.16(e), and using (2.5) again, this implies that

$$\sigma_C \Box f(x) = \left(\sigma_C \Box f\right)^{**}(x) = \sigma_C \Box f^{**}(x).$$

This finishes the proof, since x was arbitrarily chosen. $\Box$

We are now able to prove the main result of this work, which has as a consequence the required characterization of convexity.

Theorem 2.5 *Let $f : X \to \overline{\mathbb{R}}$ be a weakly lsc epi-pointed function such that*

$$\operatorname{argmin}\{f - x^*\} \text{ is convex for all } x^* \in D,$$

where D is a convex dense subset of $\operatorname{dom} f^*$. *Then we have that*

$$f^{**} = \sigma_{\operatorname{dom} f^*} \Box f.$$

Proof. Since $\operatorname{int}(\operatorname{dom} f^*) \neq \emptyset$, without loss of generality, we assume that $D \subseteq \operatorname{int}(\operatorname{dom} f^*)$. Define $\mathfrak{C} := \{\operatorname{co}(F) : F \text{ is a finite subset of } D\}$. Clearly $(\mathfrak{C}, \supseteq)$ is a directed set.

It is easy to check, using Proposition 2.4, that

$$\sup_{C \in \mathfrak{C}} \sigma_C \Box f = \sup_{C \in \mathfrak{C}} \sigma_C \Box f^{**} \leq \sigma_{\operatorname{dom} f^*} \Box f^{**} \leq \sigma_{\operatorname{dom} f^*} \Box f.$$

Now we will prove that

$$\sup_{C \in \mathfrak{C}} \sigma_C \Box f = \sigma_{\operatorname{dom} f^*} \Box f, \tag{2.6}$$

and the conclusion will follow from Lemma 1.16(d), that shows that

$$\sigma_{\operatorname{dom} f^*} \Box f^{**} = \left(I_{\operatorname{dom} f^*} + f^*\right)^* = f^{**}.$$

We fix $x \in X$ and for any $C \in \mathfrak{C}$ define $g_C(y) := f(y) + \sigma_C(x - y)$ and $g(y) := f(y) + \sigma_{\operatorname{dom} f^*}(x - y)$. Clearly, g_C and g are weakly lsc functions and $g_C \nearrow g$ pointwise.

If there exists $C \in \mathfrak{C}$ such that $g_C \equiv +\infty$, (2.6) is trivially verified.

Assume that for every $C \in \mathfrak{C}$, $g_C \not\equiv +\infty$. From lemma 1.16(f) applied to f^* at $x^* \in C (\subset \operatorname{int}(\operatorname{dom} f^*))$ we conclude that $f^{**} - x^*$ is weakly inf-compact and from the fact that

$$g_C(\cdot) = f(\cdot) + \sigma_C(x - \cdot) \geq f^{**}(\cdot) - \langle x^*, \cdot \rangle + \langle x^*, x \rangle,$$

we see that g_C is weakly inf-compact.

Now, for every $C \in \mathfrak{C}$ we take $x_C \in \operatorname{argmin}\{g_C\}$ and fix some $C_0 \in \mathfrak{C}$. Then for every $K \in \mathfrak{C}$ such that $C_0 \subseteq K$ we have $x_K \in \Gamma := \{y \in X : g_{C_0}(y) \leq \sup_{C \in \mathfrak{C}} \{\sigma_C \Box f(x)\}\}$.

Finally, on the one hand, if $\sup_{C\in\mathfrak{C}}\{\sigma_C\Box f(x)\} = +\infty$, equality (2.6) is trivial. On the other hand, if $\sup_{C\in\mathfrak{C}}\{\sigma_C\Box f(x)\} < +\infty$, then Γ is compact and we apply Lemma 2.2 to the family g_C, x_C indexed by the directed set $\{C \in \mathfrak{C} : C_0 \subseteq C\}$, to obtain equality (2.6). $\square$

The following example illustrates the necessity of considering the support function of $\operatorname{dom} f^*$ in the formula for the biconjugate given in Theorem 2.5.

Example 2.6 *Let $f : \mathbb{R} \to \mathbb{R}$ be the lsc nonconvex function defined by*

$$f(x) = \begin{cases} |x| & \textit{if} \quad x \in [-1,1] \\ |x| + \mathrm{e}^{-|x|} & \textit{if} \quad x \in \mathbb{R}\backslash[-1,1]. \end{cases}$$

Then it is easy to prove that $f^* = I_{[-1,1]}$, which shows that f is epi-pointed, and

$$\operatorname{argmin}\{f - \alpha\} = \begin{cases} \{0\} & \text{if} \quad \alpha \in (-1,1) \\ [0,1] & \text{if} \quad \alpha = 1 \\ [-1,0] & \text{if} \quad \alpha = -1. \end{cases}$$

Hence, the hypothesis of Theorem 2.5 holds, but not the equality $f^{**} = f$. However, we easily check that

$$f^{**} = |\cdot| = \sigma_{\operatorname{dom} f^*}\Box f.$$

The following remark gives a geometrical interpretation of the conclusion of Theorem 2.5 in terms of the epigraph of involved functions.

Remark 2.7 It is well known that when the inf-convolution is exact, then its epigraph is the sum of the epigraphs of the two functions. Since it can be shown that convolution in the equality of Theorem 2.5 is exact, we see that Theorem 2.5 corresponds to the set equality

$$\overline{\mathrm{co}}(\operatorname{epi} f) = \operatorname{epi} f + \operatorname{epi} \sigma_{\operatorname{dom} f^*}. \tag{2.7}$$

On the other hand, if we consider the asymptotic cone of $\operatorname{epi} f$ given by (see, e.g., [40])

$$(\operatorname{epi} f)_\infty := \bigcap_{\varepsilon>0} \overline{]0,\varepsilon] \operatorname{epi} f},$$

which is the epigraph of the asymptotic function f_∞, since $\overline{\mathrm{co}}(f_\infty) = \sigma_{\operatorname{dom} f^*}$ when f is weakly lsc and epi-pointed (see [28, Theorem 7]), we can rewrite Theorem 2.5 as

$$f^{**} = f\Box\, \overline{\mathrm{co}}(f_\infty),$$

and the equation (2.7) as

$$\overline{\mathrm{co}}(\operatorname{epi} f) = \overline{\mathrm{co}}((\operatorname{epi} f)_\infty) + \operatorname{epi} f.$$

It is worth noting that this characterization does not involve dual objects.

The next corollary corresponds to the announced result of this work: The extension of Theorem 2.1 to the setting of locally convex spaces.

Corollary 2.8 *Let $f : X \to \overline{\mathbb{R}}$ be a weakly lsc epi-pointed function such that* $\operatorname{dom} f^* = X^*$*. If there exists a convex dense set $D \subset X^*$ such that* $\operatorname{argmin}\{f - x^*\}$ *is convex for all $x^* \in D$, then f is convex.*

The next remark compares the hypotheses of Theorem [107, Theorem 10] and Corollary 2.8

Remark 2.9 It is worth observing that in [107, Theorem 10] the density assumption on D is with respect to the norm topology in X^*. This is clearly stronger than the condition used in Corollary 2.8 asking that D is dense only with respect to the Mackey topology. On the other hand, according to Lemma 1.16(f) we see that the hypothesis of weakly infcompactness of $f - x^*$ is equivalent to the continuity of f^* over all X^* with respect to the Mackey topology.

The following example shows the necessity of the convexity assumption of D in Corollary 2.8.

Example 2.10 *Let h be a proper lsc convex function defined on a reflexive Banach space Z such that* $\operatorname{dom} h^* = Z^*$*; hence, h is epi-pointed (with respect to the norm-topology in Z^*). Choose any nonconvex positive and weakly lsc function g such that the function $f := h + g$ is not convex. Then f defines a weakly lsc epi-pointed function, due to the relation $h^* \geq f^*$ which implies that* $\operatorname{dom} f^* = Z^*$ *and f^* is norm-continuous on Z^*. Since f^* is Fréchet-differentiable in a (G_δ)-dense subset of Z^* (see, e.g., [94]), $D \subset Z^*$, by [25, Proposition 6] we deduce that, for all $x^* \in D$,*

$$\operatorname{argmin}\{f - x^*\} = (\partial f)^{-1}(x^*) = \{\nabla f^*(x^*)\}.$$

The next corollary shows that the epi-pointedness assumption in Corollary 2.8 can be replaced by the assumption of convexity of ε- $\operatorname{argmin}\{f - x^*\}$ when ε is sufficiently small and the density of D in the norm topology.

Corollary 2.11 *Suppose that X is a normed space, and let $f : X \to \overline{\mathbb{R}}$ be a weakly lsc function such that* $\operatorname{dom} f^* = X^*$*. Assume that there exist a convex (norm-)dense set $D \subset X^*$ such that for all $x^* \in D$ there is some $\delta > 0$ such that*

$$\varepsilon\text{- }\operatorname{argmin}\{f - x^*\} \textit{ is convex, for all } \varepsilon \in (0, \delta).$$

Then f is convex.

Proof. We consider the duality pair $((X^{**}, w^*), (X^*, \|\cdot\|_*))$ together with the function $\overline{f}^{w^*} : X^{**} \to \overline{\mathbb{R}}$ given by

$$\overline{f}^{w^*}(x^{**}) := \liminf_{\substack{x_\alpha \xrightarrow{w^*} x^{**} \\ x_\alpha \in X}} f(x_\alpha).$$

It is clear that $\overline{f}^{w^*}$ is weakly lsc and epi-pointed. Moreover, since for every $x^* \in D$ we have

$$\operatorname{argmin}\{\overline{f}^{w^*} - x^*\} = \bigcap_{\varepsilon > 0} \overline{\varepsilon\text{- }\operatorname{argmin}\{f - x^*\}}^{w^*},$$

we deduce that $\operatorname{argmin}\{\overline{f}^{w^*} - x^*\}$ is convex. Hence, by applying Corollary 2.8 we conclude that $\overline{f}^{w^*}$ is convex, and so is the function f. $\square$

In the final part, keeping in mind a possible application to the finite-dimensional case, we consider the relative interior within the definition of epi-pointed functions. More generally, we have:

Theorem 2.12 *Let $f : X \to \overline{\mathbb{R}}$ be a function with a proper conjugate, and denote $F := \overline{\operatorname{aff}}(\operatorname{dom} f^*)$. We suppose that the following conditions hold:*

(a) *The restriction of f^* to F, ${f^*}_{|F}$, is continuous on* $\operatorname{ri}(\operatorname{dom} f^*)$.

(b) *There exists $x_0^* \in \operatorname{ri}(\operatorname{dom} f^*)$ such that $f - x_0^*$ is weakly lsc and weakly inf-compact.*

(c) *There exists a convex set $D \subseteq \operatorname{ri}(\operatorname{dom} f^*)$, with $\operatorname{ri}(\operatorname{dom} f^*) \subset \overline{D}$, such that* $\operatorname{argmin}\{f - x^*\}$ *is convex for all $x^* \in D$.*

Then we have

$$\sigma_{\operatorname{dom} f^*} \square f = f^{**}.$$

Moreover, if $\overline{D} = F$, then

$$\sigma_{\operatorname{dom} f^*} \square f = f^{**} = f_F,$$

where $f_F(x) := \inf\{f(w) : w \in x + (F - x_0^*)^\perp\}$.

Proof. We may suppose without loss of generality that $x_0^* = 0$ so that the function f_F defined above is written as

$$f_F(x) := \inf\{f(w) : w \in x + F^\perp\}.$$

We also denote by $Z := X/F^\perp$ the quotient space of X by the orthogonal space of F, and introduce the function $h : Z \to \overline{\mathbb{R}}$ defined as

$$h([x]) := f_F(x).$$

Let us consider the dual pair $(Z, \sigma(Z, F), F, \tau(X^*, X))$ endowed with the bilinear form $\langle x^*, [x]\rangle = \langle x^*, x\rangle$. Then, from the relation

$$\{z \in Z : h(z) \leq \lambda\} = \Pi(\{x \in X : f(x) \leq \alpha\}), \ \ \forall \lambda \in \mathbb{R},$$

where $\Pi : X \to X/F^\perp$ is the canonical projection, i.e., $\Pi(x) = [x]$, it follows that h is weakly lsc. Also, since f is weakly inf-compact, the relation above together with the fact that $h^* = {f^*}_{|F}$ also imply that

$$\operatorname{int}(\operatorname{dom} h^*) = \operatorname{ri}(\operatorname{dom} f^*),$$

and h is epi-pointed.

Next, because f is weakly inf-compact, we get

$$\operatorname{argmin}\{h - x^*\} = \Pi(\operatorname{argmin}\{f - x^*\}),$$

which shows that $\operatorname{argmin}\{h - x^*\}$ is convex.

Now, we are able to apply Theorem 2.5 to get, for every $x \in X$,

$$\sigma_{\operatorname{dom} h^*} \Box h = \sup_{x^* \in F} \{\langle \cdot, x^* \rangle - h^*(x^*)\}.$$

Moreover, using the fact that $h^* = f^*{}_{|F}$ and $\sigma_{\operatorname{dom} f^*}(x - y - z) = \sigma_{\operatorname{dom} f^*}(x - y)$, for every $z \in F^\perp$, we see that for all $x \in X$

$$\sigma_{\operatorname{dom} h^*} \Box h([x]) = \sigma_{\operatorname{dom} f^*} \Box f(x),$$

and

$$\sup_{x^* \in F} \{\langle x, x^* \rangle - h^*(x^*)\} = f^{**}(x).$$

Finally, if $F \subseteq \overline{\operatorname{dom} f^*}$, then $x - y \notin F^\perp$, and so

$$\sigma_{\operatorname{dom} f^*}(x - y) = \sigma_F(x - y) = +\infty.$$

Therefore, $\sigma_{\operatorname{dom} f^*} \Box f = f_F$. $\square$

Remark 2.13 With the hypothesis of Theorem 2.12, it is not possible to get the equality $f = f^{**}$ instead of $f_{F - x_0^*} = f^{**}$. Indeed, consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ given by

$$f(x, y) = \frac{1}{2}x^2 + \ln(|y| + 1).$$

Then the conjugate is given by

$$f^*(\alpha, \eta) = \begin{cases} \frac{\alpha^2}{2} & \text{if} \quad \eta = 0 \\ +\infty & \text{if} \quad \eta \neq, 0 \end{cases}$$

so that $\operatorname{dom} f^* = \mathbb{R} \times \{0\}$ and for every $(\alpha, 0)$,

$$\operatorname{argmin}\{f - (\alpha, 0)\} = \{(\alpha, 0)\};$$

that is, $\operatorname{argmin}\{f - (\alpha, 0)\}$ is convex, while f is not convex.

Chapter 3

On Brøndsted-Rockafellar's Theorem for convex lower semicontinuous epi-pointed functions in locally convex spaces

3.1 Introduction

The first chapter opens our minds to the class of epi-ponted functions, which appear to be promising for develop variation analysis outside of Banach spaces. Before presenting the definition of an epi-pointed function that we adopted in infinite dimension locally convex spaces, let us introduce the history about this family and some properties in finite-dimension which motivate our definition.

To better understand the research done in this thesis we must explain the terminology *variational analysis* and recall some results in the theory of variational and convex analysis. One of the most simple and intuitive introductions to the concept of variational analysis is given in the acclaimed book by Jonathan M. Borwein and Qiji J. Zhu [14]. *Variational techniques refer to proofs by way of establishing that an appropriate auxiliary function attains a minimum. This can be viewed as a mathematical form of the principle of least action in physics. Since so many important results in mathematics, in particular, in analysis have their origins in the physical sciences, it is entirely natural that they can be related in one way or another to variational techniques. The use of variational arguments in mathematical proofs has a long history. This can be traced back to Johann Bernoulli's problem of the Brachistochrone and its solutions leading to the development of the calculus of variations. Since then the method has found numerous applications in various branches of mathematics. A simple illustration of the variational argument is the following example.*

Example 3.1 (Surjectivity of Derivatives) Suppose that $f : \mathbb{R} \to \mathbb{R}$ is differentiable everywhere and suppose that $\lim\limits_{|x|\to\infty} f(x)/|x| = +\infty$. Then $\{f'(x) : x \in \mathbb{R}\} = \mathbb{R}$.

Proof. Let r be an arbitrary real number. Define $g(x) = f(x) - rx$. We easily check that

$g(x) \to +\infty$ as $|x| \to +\infty$ and therefore attains a (global) minimum at some $\bar{x}$. Then $0 = g'(\bar{x}) = f'(\bar{x}) - r$. $\square$

Two conditions are essential in this variational argument. ***The first is compactness*** *(to ensure the existence of the minimum) and* ***the second is differentiability*** *of the auxiliary function (so that the differential characterization of the results is possible). Two important discoveries in the 1970's led to significant useful relaxation on both conditions. First, the discovery of general variational principles led to the relaxation of the compactness assumptions. Such principles typically assert that any lower semicontinuous (lsc) function, bounded from below, may be perturbed slightly to ensure the existence of the minimum. Second, the development of the nonsmooth analysis made possible the use of nonsmooth auxiliary functions.*

To the best of my knowledge this paragraph represents a brief, but concrete introduction to the techniques necessary to develop variational analysis. Summarizing the ideas one needs the existence of minima, or at least the existence of minima to perturbed problems and also some form of generalized differentiation is required.

In mathematical analysis, Ekeland's variational principle, discovered by *Ivar Ekeland* in 1979 is a theorem that asserts that there exists nearly optimal solutions to some perturbed optimization problems.

Theorem 3.2 *[45, I. Ekeland, Theorem 1] Let (X, d) be a complete metric space and let $f : X \to \overline{\mathbb{R}}$ be an lsc proper function, let $\varepsilon, \lambda > 0$ and $u \in X$ such that:*

$$f(u) \leq \inf f + \varepsilon.$$

Then there exists $u_\varepsilon \in X$ such that:

(a) $\mathrm{d}(u, u_\varepsilon) \leq \lambda$,

(b) $f(u_\varepsilon) + \varepsilon/\lambda \mathrm{d}(u, u_\varepsilon) \leq f(u)$,

(c) $f(u_\varepsilon) < f(y) + \varepsilon/\lambda \mathrm{d}(y, u_\varepsilon)$, $\forall y \in X \backslash \{u_\varepsilon\}$.

The above result has been utilized to prove, or to give alternative proofs to some of the most important theorems in nonsmooth analysis, in particular convex analysis. One can find many examples of the tremendous and influential power of this theorem by reading [14], for instance: Equivalence of the Ekeland's variational principle with the completeness of the space, a proof of the *Banach fixed point theorem* based in this variational principle, Borwein-Preiss variational principle, the theory of convex analysis from variational principles, nonconvex generalized differentiability, among many other topics.

Convex analysis has been one of the most important topics in optimization theory, in particular the subdifferential theory has granted diverse tools to resolve theoretical and applied problems.

One of the results which can be proved using Ekeland's variational principle is the well-known *Brøndsted-Rockafellar Theorem.* This fundamental proposition basically gives us the non-vacuity of the convex subdifferential in many points, since it is a powerful result in convex analyis, it has been redacted in many different forms. One of the form is:

Theorem 3.3 (Brøndsted, A. and Rockafellar, R. T. [19]) *Let X be a Banach space and let $f : X \to \overline{\mathbb{R}}$ be a convex proper lsc function. Consider $\varepsilon \geq 0$ and $x_0^* \in \partial_\varepsilon f(x_0)$. Then there exists $x_\varepsilon^* \in \partial f(x_\varepsilon)$ such that $\|x_\varepsilon - x_0\| \leq \sqrt{\varepsilon}$ and $\|x_\varepsilon^* - x_0^*\| \leq \sqrt{\varepsilon}$. In particular,* $\operatorname{dom} f \subset \overline{\operatorname{dom} \partial f}$ *and* $\operatorname{dom} f^* \subset \overline{\operatorname{rge} \partial f}$*.*

Another important result is the *maximal monotonicity of the subdifferential*. It is clear that for every function f the convex subdifferential ∂f is a monotone multifunction in the sense that

$$\langle x^* - y^*, x - y \rangle \geq 0, \text{ for all } (x, x^*), (y, y^*) \in \operatorname{gph} \partial f.$$

Moreover, one can also get easily that ∂f is cyclically monotone; that is for all $n \in \mathbb{N}$ and all $(x_\mathrm{i}, x_\mathrm{i}^*) \in \operatorname{gph} \partial f$, with $\mathrm{i} = 1, ..., n$ one has

$$0 \geq \langle x_0 - x_n, x_n^* \rangle + ... + \langle x_2 - x_1, x_1^* \rangle + \langle x_1 - x_0, x_0^* \rangle. \tag{3.1}$$

Rockafellar said in [98]: *The cyclic monotonicity condition can be viewed heuristically as a discrete substitute for two classical conditions: that a smooth convex function has a positive semi-definite second differential, and that all circuit integrals of an integrable vector field must vanish.* Also in the same paper he proved that every multifunction $M : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ satisfying the cyclically monotone relation (3.1) must be contained in the subdifferential of some proper convex function f, that is to say, $M(x) \subseteq \partial f(x)$ for all $x \in \mathbb{R}^n$. Besides, Rockafellar proved that for every lsc convex function f the subdifferential mapping $\partial f : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ cannot be extended to strictly larger mapping (in the sense of inclusion of theirs graphs), that is to say, it is a maximal monotone operator. The precise statement is the following:

Theorem 3.4 (Rockafellar, R. T. [100]) *Let X be a Banach space and let f be an lsc proper convex function on X. Then ∂f is a maximal monotone operator.*

This kind of assertion goes back to [83], where Minty proves it for a continuous convex function defined on a Hilbert space. Later, Moreau in [87] also gave a proof in Hilbert spaces without the continuity assumption, using duality and the Moreau-Yosida envelope. It is important to recall that the theorem above have gotten the attention of many authors who have made different proofs of this result, even when this result was proved by Rockafellar in any arbitrary Banach space, among others we refer to Borwein [8], Ivanov and Zlateva [70], Marques Alves and Svaiter [82], Simons [111], Taylor [112], Thibault [116], Zalinescu [133].

Now consider a linear operator $A : \mathbb{R}^m \to \mathbb{R}^n$ and convex proper lsc functions $g, h : \mathbb{R}^n \to \overline{\mathbb{R}}$. One important question from the nineties in the spirit of Brøndsted-Rockafellar Theorem was: Could it be possible for any point $x^* \in \partial g \circ A(x)$ to find sequences $y_n^* \in \partial g(y_n)$ such that $y_n \to A(x)$ and $A^*(y_n^*) \to x^*$? And in the same way for the sum of functions: Could it be possible for any point $x^* \in \partial(g + h)(x)$ to find sequences $y_n^* \in \partial g(y_n)$ and $z_n^* \in \partial g(z_n)$ such that $y_n, z_n \to x$ and $y_n^* + z_n^* \to x^*$? Both questions have been positively answered for many authors (see e.g. [48], [71], [89], [113], [115], [116]). This kind of formulas are commonly called limiting calculus rules for the convex subdifferential or Fuzzy calculus rules for the convex subdifferential. One of the classical statement, is given by:

Theorem 3.5 *[89, Theorem 2.2] Let X and Y be two Banach spaces and let $A : X \to Y$ be a linear operator and $g : Y \to \overline{\mathbb{R}}$ be a convex proper lsc function. Then for any $x \in X$ with $A(x) \in \operatorname{dom} g$ one has $x^* \in \partial\, g \circ A(x)$ if and only if there exist nets $y^*_\nu \in Y^*$, $y_\nu \in Y$ such that $y^*_\nu \in \partial\, g(y_\nu)$, $y_\nu \to A(x)$, $g(y_\nu) \to g(A(x))$ and $A^*(y_n)^* \xrightarrow{w^*} x^*$.*

Theorem 3.6 *[113, Theorem 3] Let X be a Banach space and let $g, h : X \to \overline{\mathbb{R}}$ be two proper lsc functions. Then for every $x \in \operatorname{dom} g \cap \operatorname{dom} h$ one has $x^* \in \partial(g+h)(x)$ if and only if there exist nets $y^*_\nu, z^*_\nu \in X^*$, $y_\nu, z_\nu \in X$ such that $y^*_\nu \in \partial\, g(y_\nu)$, $z^*_\nu \in \partial\, h(z_\nu)$, $y_\nu, z_\nu \to x$, $g(y_\nu) \to g(x)$, $h(z_\nu) \to h(x)$ and $y^*_\nu + z^*_\nu \xrightarrow{w^*} x^*$.*

It is known that the Brøndsted-Rockafellar's Theorem is not valid outside Banach spaces for all lsc proper convex functions (see, e.g., [19]), more precisely Brøndsted and Rockafellar found an lsc proper convex function defined on a locally convex topological vector space with empty subdifferential everywhere, and consequently all of the above statements are also not valid. This observation motivates the work to provide a suitable family of lsc proper convex functions defined on locally convex spaces, which satisfies the Brøonsted-Rockafellar Theorem.

The main features of this work are:

(1) Show that epi-pointed lsc convex functions, defined on any locally convex space, satisfy the Brøndsted-Rockafellar's Theorem (see Theorem 3.8 and Corollary 3.9).

(2) Provide a different proof of Brøndsted-Rockafellar's Theorem for this class of epi-pointed functions, in the sense that it is based on a very simple variational principle, which is valid in locally convex spaces, without requiring such tools as Ekeland's or Bishop-Phelps' variational principles (see Lemma 3.7).

(3) Since every convex function in Banach spaces can be adequately perturbed to obtain an epi-pointed function, we recover in the Banach setting the usual Brøndsted-Rockafellar's theorem (see Section 3.3).

(4) We also obtain other important results in the same spirit, as all the above Theorems, Theorems 3.10 for the Maximal monotonicity of the subdifferential and Theorems 3.12 and 3.14 for the subdifferential limiting calculus rules for convex functions.

3.2 Brøndsted-Rockafellar's Theorem and consequences

First we give in Lemma 3.7 a simple variational principle, for convex functions defined on lcs, that is the key tool in the proof of our main results.

Lemma 3.7 *Let $x_0 \in X$ and let $f : X \to \mathbb{R} \cup \{+\infty\}$ and $\rho : X \to [0, \infty)$ be two convex lsc functions such that $\rho(0) = 0$ and the function $f(\cdot) + \rho(\cdot - x_0)$ is epi-pointed. For any $\varepsilon \geq 0$, $x_0 \in X$ and $x^*_0 \in \partial_\varepsilon f(x_0)$ with $x^*_0 \in \operatorname{int}(\operatorname{dom}(f + \rho(\cdot - x_0))^*)$, there exists $x_\varepsilon \in X$ such that:*

(a) $\rho(x_0 - x_\varepsilon) \leq \varepsilon$,

(b) $x_0^* \in \partial(f + \rho(\cdot - x_0))(x_\varepsilon)$,

(c) $|f(x_0) - f(x_\varepsilon)| \leq |\langle x_0^*, x_0 - x_\varepsilon\rangle| + \varepsilon$.

Proof. Define $g := f + \rho(\cdot - x_0) - x_0^*$. By Proposition 1.16(f) we see that the (proper lsc convex) function g is inf-compact, and there exists $x_\varepsilon \in \operatorname{argmin} g$ (the minima of g), such that

$$f(x_\varepsilon) + \rho(x_\varepsilon - x_0) - x_0^*(x_\varepsilon) \leq f(x_0) - x_0^*(x_0) \leq f(x_\varepsilon) - x_0^*(x_\varepsilon) + \varepsilon.$$

Hence, $\rho(x_\varepsilon - x_0) \leq \varepsilon$. Now, since $0 \in \partial g(x_\varepsilon)$, by Proposition 1.16(a) we have

$$x_0^* \in \partial(f + \rho(\cdot - x_0))(x_\varepsilon).$$

Finally, $\langle x_0^*, x_0 - x_\varepsilon\rangle \leq f(x_0) - (f(x_\varepsilon) + \rho(x_\varepsilon - x_0)) \leq f(x_0) - f(x_\varepsilon) \leq \langle x_0^*, x_0 - x_\varepsilon\rangle + \varepsilon$, gives us the last statement. □

The following result gives the counterpart of Brøndsted-Rockafellar-Borwein Theorem (e.g., [8]) for convex lsc epi-pointed functions defined in locally convex spaces.

Theorem 3.8 *Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a convex, lsc and epi-pointed function. Consider $\varepsilon \geq 0$, $\beta \in [0, \infty)$, a continuous seminorm p, $\lambda > 0$ and $x_0 \in X$. If $x_0^* \in \partial_\varepsilon f(x_0) \cap \operatorname{int}(\operatorname{dom} f^*)$, then there are $x_\varepsilon \in X$, $y_\varepsilon^* \in \mathbb{B}_p(0,1)^\circ$ and $\lambda_\varepsilon \in [-1, 1]$ such that:*

(a) $p(x_0 - x_\varepsilon) + \beta|\langle x_0^*, x_0 - x_\varepsilon\rangle| \leq \lambda$,

(b) $x_\varepsilon^* := x_0^* + \frac{\varepsilon}{\lambda}(y_\varepsilon^* + \beta\lambda_\varepsilon x_0^*) \in \partial f(x_\varepsilon)$,

(c) $|\langle x_\varepsilon^*, x_0 - x_\varepsilon\rangle| \leq \varepsilon + \frac{\lambda}{\beta}$,

(d) $|f(x_0) - f(x_\varepsilon)| \leq \varepsilon + \frac{\lambda}{\beta}$

(e) $x_\varepsilon^* \in \partial_{2\varepsilon} f(x_0)$.

With the convention $\frac{1}{0} = +\infty$.

Proof. Define $\rho(x) = \frac{\varepsilon}{\lambda}(p(x) + \beta|\langle x_0^*, x\rangle|)$. We apply Lemma 3.7 to f and ρ to conclude the existence of $x_\varepsilon \in X$ such that $\rho(x_0 - x_\varepsilon) \leq \varepsilon$, $x_0^* \in \partial(f + \rho(\cdot - x_0))(x_\varepsilon)$ and $|f(x_0) - f(x_\varepsilon)| \leq |\langle x_0^*, x_0 - x_\varepsilon\rangle| + \varepsilon$. So, x_ε verifies (a) and (d). Now we use Proposition 1.16(a) to obtain $0 \in \partial f(x_\varepsilon) - x_0^* + \frac{\varepsilon}{\lambda}\mathbb{B}_p(0,1)^\circ + \frac{\varepsilon}{\lambda}\beta \cdot [-1, 1] \cdot x_0^*$, from which we find $y_\varepsilon^* \in \mathbb{B}_p(0,1)^\circ$ and $\lambda_\varepsilon \in [-1, 1]$ such that $x_\varepsilon^* := x_0^* + \frac{\varepsilon}{\lambda}(y_\varepsilon^* + \beta\lambda_\varepsilon x_0^*) \in \partial f(x_\varepsilon)$. Then

$$\begin{aligned}
|\langle x_\varepsilon^* - x_0^*, x_0 - x_\varepsilon\rangle| &\leq \tfrac{\varepsilon}{\lambda}|\langle y_\varepsilon^* + \beta\lambda_\varepsilon x_0^*, x_0 - x_\varepsilon\rangle| \\
&\leq \tfrac{\varepsilon}{\lambda}(|\langle y_\varepsilon^*, x_0 - x_\varepsilon\rangle| + \beta|\langle x_0^*, x_0 - x_\varepsilon\rangle|) \\
&\leq \tfrac{\varepsilon}{\lambda}(p(x_0 - x_\varepsilon) + \beta|\langle x_0^*, x_0 - x_\varepsilon\rangle|) \leq \varepsilon,
\end{aligned}$$

and (c) follows (using (a)) $|\langle x_\varepsilon^*, x_0 - x_\varepsilon\rangle| \leq |\langle x_\varepsilon^* - x_0^*, x_0 - x_\varepsilon\rangle| + |\langle x_0^*, x_0 - x_\varepsilon\rangle| \leq \varepsilon + \frac{\lambda}{\beta}$. Finally, since $x_0^* \in \partial_\varepsilon f(x_0)$ and $x_\varepsilon^* \in \partial f(x_\varepsilon)$ we get, for every $x \in X$,

$$\begin{aligned}\langle x_\varepsilon^*, x - x_0\rangle &= \langle x_\varepsilon^*, x - x_\varepsilon\rangle + \langle x_\varepsilon^* - x_0^*, x_\varepsilon - x_0\rangle + \langle x_0^*, x_\varepsilon - x_0\rangle \\ &\leq f(x) - f(x_\varepsilon) + \varepsilon + f(x_\varepsilon) - f(x_0) + \varepsilon \\ &= f(x) - f(x_0) + 2\varepsilon;\end{aligned}$$

that is, $x_\varepsilon^* \in \partial_{2\varepsilon} f(x_0)$. ☐

This Theorem allows us to obtain the counterpart of the classical statement of Brøndsted-Rockafellar's Theorem for lsc convex epi-pointed functions in locally convex spaces.

Corollary 3.9 *Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a convex lsc epi-pointed function. Then for every $x \in \operatorname{dom} f$ there exist nets $\{(x_\alpha, x_\alpha^*)\}_{\alpha \in \mathbb{A}}$ such that $x_\alpha^* \in \partial f(x_\alpha)$, $x_\alpha \to x$ and $f(x_\alpha) \to f(x)$.*

Proof. Consider a filtered family of seminorms $\mathcal{N}$ (ordered by $\rho_1 \leq \rho_2$ if and only if $\rho_1(x) \leq \rho_2(x)$ for all $x \in X$), which defines the topology on X. We also define the index set $\mathbb{A} := \mathcal{N} \times (0,1)$ associated with the partial order

$$\alpha_1 = (\rho_1, \varepsilon_1) \leq \alpha_2 = (\rho_2, \varepsilon_2) \text{ if and only if } \rho_1 \leq \rho_2 \text{ and } \varepsilon_1 \geq \varepsilon_2.$$

It is easy to see that for every $\varepsilon \in (0,1)$, $\partial_\varepsilon f(x) \cap \operatorname{int}(\operatorname{dom} f^*) \neq \emptyset$. Therefore, for every $\varepsilon \in (0,1)$ and for every continuous seminorm $\rho \in \mathcal{N}$ we can apply Theorem 3.8 to f with $\beta = 1$ and $\lambda = \sqrt{\varepsilon}$. We get that there exists $(x_{\varepsilon,\rho}, x_{\varepsilon,\rho}^*)$ in the graph of the subdifferential of f such that $\rho(x - x_{\varepsilon,\rho}) \leq \sqrt{\varepsilon}$ and $|f(x) - f(x_{\varepsilon,p})| \leq \varepsilon + \sqrt{\varepsilon}$. To prove the convergence of this net, take a neighborhood V of zero and $\delta > 0$ and let $\rho_0 \in \mathcal{N}$ and $\varepsilon_0 \in (0,1)$ such that $\mathbb{B}_{\rho_0}(0, \sqrt{\varepsilon_0}) \subseteq V$ and $\varepsilon_0 + \sqrt{\varepsilon_0} \leq \delta$. Therefore, for every $(\rho, \varepsilon) \geq (\rho_0, \varepsilon_0)$ we have that $\rho_0(x - x_{\varepsilon,\rho}) \leq \rho(x - x_{\varepsilon,\rho}) \leq \sqrt{\varepsilon} \leq \delta$ and $|f(x) - f(x_{\varepsilon,p})| \leq \varepsilon + \sqrt{\varepsilon} \leq \varepsilon_0 + \sqrt{\varepsilon_0} \leq \delta$, which implies that $x_{\varepsilon,\rho} \in x + V$ and $|f(x) - f(x_{\varepsilon,p})| \leq \delta$. ☐

Now, we apply Theorem 3.8 to obtain the maximal monotonicity of the subdifferential operator of proper lsc convex epi-pointed functions in locally convex spaces.[1] For proper lsc convex functions defined in Banach spaces, this corresponds to the famous theorem by Rockafellar, [100, Theorem A].

Theorem 3.10 *Let $f : X \to \mathbb{R} \cup \{+\infty\}$ be a convex lsc function. If either f or f^* is epi-pointed, then ∂f and ∂f^* are maximal monotone operators.*

Proof. Without lost of generality we consider that f is an epi-pointed function, that $f(0) = 0$ and that $0 \notin \partial f(0)$. We pick an $x \in X$ such that $f(2x) < f(x) < 0$ (such an element exists, since by supposing the contrary, one deduces that $f(x) \geq f(\frac{x}{2})$ for any x such that $f(x) < 0$, and we get $f(x) \geq f(\frac{x}{2^n})$ for any $n \in \mathbb{N}$, which leads us to the contradiction $f(x) \geq \liminf\limits_{n \to +\infty} f(\frac{x}{2^n}) \geq f(0)$). If $a := f(x) - f(2x)$ and $\delta \in (0, \frac{a}{a+3})$, we choose an $x^* \in$

[1]The arguments used in the proof of this result (Theorem 3.10) follows the suggestion made by one of the referees.

$\partial_{\delta^2} f(x) \cap \operatorname{int}(\operatorname{dom} f^*)$ (this choice is possible since $f(x) = f^{**}(x) = \sup\{\langle y^*, x\rangle - f^*(y^*) : y^* \in \operatorname{int}(\operatorname{dom} f^*)\}$). Then

$$\langle x^*, x\rangle = \langle x^*, 2x - x\rangle \leq f(2x) - f(x) + \delta^2 \leq \delta^2 - a < 0.$$

Define $p(x) = |\langle x^*, x\rangle|$ a continuous seminorm so that $\mathbb{B}_p(0,1) = \{x \in X : |\langle x^*, x\rangle| \leq 1\}$ and $(\mathbb{B}_p(0,1))^\circ = [-1,1]\cdot x^*$. Then Theorem 3.8 (with $\varepsilon = \delta^2$, $\lambda = \delta$ and $\beta = 0$) ensures the existence of $x_0 \in X$ and $x_0^* \in \partial f(x_0)$ such that $x - x_0 \in \delta\mathbb{B}_p(0,1)$ and $x^* - x_0^* \in \delta(\mathbb{B}_p(0,1))^\circ$. Thus, $|\langle x^* - x_0^*, x - x_0\rangle| \leq \delta^2$, $|\langle x^* - x_0^*, x\rangle| \leq \delta|\langle x^*, x\rangle| = \delta(a - \delta^2)$ and $|\langle x^*, x - x_0\rangle| \leq \delta$. In conclusion, we get

$$\begin{aligned}\langle x_0^*, x_0\rangle =& \langle x^*, x\rangle + \langle x^* - x_0^*, x - x_0\rangle + \langle x_0^* - x^*, x\rangle + \langle x^*, x_0 - x\rangle\\ \leq& \delta^2 - a + \delta^2 + \delta(a - \delta^2) + \delta < -a + \delta(a+3) < 0.\end{aligned}$$

□

We finish this section by applying Theorem 3.8 to get limiting calculus rules for the subdifferential mapping of the composition with a linear mapping and the sum of convex functions. Our proof is an adaptation of [113, Theorem 3] that uses the following lemma.

Lemma 3.11 *Let X, Z be two lcs, $f \in \Gamma_0(Z)$ be an epi-pointed function and $A \in \mathcal{L}(X,Z)$ (linear and continuous mapping). Then for every $x \in \operatorname{dom}(f \circ A)$*

$$\partial\left(f \circ A\right)(x) = \bigcap_{\eta > 0} \overline{A^*\left[\partial_\eta f(Ax) \cap \operatorname{int}(\operatorname{dom} f^*)\right]}^{w^*}.$$

Proof. By Proposition 1.3 we only need to prove that for every $\eta > 0$

$$A^*\left(\partial_\eta f(Ax)\right) \subset \overline{\left[A^*\left(\partial_\eta f(Ax) \cap \operatorname{int}(\operatorname{dom} f^*)\right)\right]}^{w^*}.$$

Indeed, if $z \in \operatorname{dom} f$, then

$$f(z) = \sup\{\langle x^*, z\rangle - f(x^*) : x^* \in \operatorname{int}(\operatorname{dom} f^*)\}$$

and $\partial_\eta f(z) \cap \operatorname{int}(\operatorname{dom} f^*) \neq \emptyset$ for every $\eta > 0$. Because $\operatorname{int}(\operatorname{dom} f^*)$ is open and dense in $\operatorname{dom} f^*$ we have that $\overline{\partial_\eta f(z) \cap \operatorname{int}(\operatorname{dom} f^*)} = \overline{\partial_\eta f(z)} = \partial_\eta f(z)$. Hence, since A^* is w^* to w^* continuous we conclude the lemma. □

Theorem 3.12 *Let X, Z be two lcs, $A \in \mathcal{L}(X,Z)$, $g \in \Gamma_0(Z)$ be an epi-pointed function, $f := g \circ A$ and $x \in \operatorname{dom} f$. Then $x^* \in \partial f(x)$ if and only if there exists a net $(z_{\mathrm{i}}, z_{\mathrm{i}}^*)_{\mathrm{i} \in I} \in Z \times Z^*$ such that $z_{\mathrm{i}}^* \in \partial g(z_{\mathrm{i}})$, $z_{\mathrm{i}} \to y = Ax$, $g(z_{\mathrm{i}}) \to g(y)$, $\langle z_{\mathrm{i}} - z, z_{\mathrm{i}}^*\rangle \to 0$ and $A^*(z_{\mathrm{i}}^*) \xrightarrow{w^*} x^*$.*

Proof. Take $x_0^* \in \partial f(x)$. Consider a filtered family of seminorms $\mathcal{N}_1$ (ordered by $\rho_Z^1 \leq \rho_Z^2$ if and only if $\rho_Z^1(z) \leq \rho_Z^2(z)$ for all $z \in Z$), which defines the topology on Z and a filtered family of seminorms $\mathcal{N}_2$ (ordered in a similar way as $\mathcal{N}_1$: $\rho_{X^*}^1 \leq \rho_{X^*}^2$ if and only if $\rho_{X^*}^1(x^*) \leq \rho_{X^*}^2(x^*)$ for all $x^* \in X^*$), which defines the weak* topology on X^*. We also define the index set $I := \mathcal{N}_1 \times \mathcal{N}_2 \times (0,1)$ ordered by $\mathrm{i}_1 = (\rho_Z^1, \rho_{X^*}^1, \varepsilon_1) \leq \mathrm{i}_2 = (\rho_Z^2, \rho_{X^*}^2, \varepsilon_2)$ if and only if $\rho_Z^1 \leq \rho_Z^2$, $\rho_{X^*}^1 \leq \rho_{X^*}^2$ and $\varepsilon_1 \geq \varepsilon_2$. Now, we take $\mathrm{i} = (\rho_Z, \rho_{X^*}, \varepsilon) \in I$ and set $U := \{z \in Z : p_Z(z) \leq 1\}$.

Choose an $\eta > 0$ such that $\sqrt{\eta} + \eta \leq \frac{\varepsilon}{2}$ and $2\sqrt{\eta} \leq \varepsilon \left(\max_{U^\circ} p_{X^*}(A^*(y^*)) + p_{X^*}(x_0^*) + 1 \right)^{-1}$. Then by Lemma 3.11 we take $z_0^* \in \partial_\eta g(z) \cap \operatorname{int}(\operatorname{dom} g^*)$ such that $p_{X^*}(A^* z_0^* - x_0^*) \leq \eta$. By Theorem 3.8 (with $\beta = 1$, $\lambda = \sqrt{\eta}$) there exists $z_{\rm i}^* \in \partial g(z_{\rm i})$ such that $p_Z(z_{\rm i} - z) \leq \sqrt{\eta}$, $z_{\rm i}^* = z_0^* + \sqrt{\eta}\,(u_{\rm i}^* + \lambda_\eta z_0^*) \in \partial g(z_{\rm i})$, $u_\eta^* \in U^\circ$, $|\langle z_{\rm i}^*, z - z_{\rm i}\rangle| \leq \eta + \sqrt{\eta}$ and $|g(z) - g(z_{\rm i})| \leq \eta + \sqrt{\eta}$. Therefore, $p_Z(z_{\rm i} - z) \leq \varepsilon$, $|\langle z_{\rm i}^*, z - z_{\rm i}\rangle| \leq \varepsilon$ and $|g(z) - g(z_{\rm i})| \leq \varepsilon$. Finally,

$$
\begin{aligned}
p_{X^*}(A^* z_{\rm i}^* - x_0^*) &\leq p_{X^*}(A^* z_0^* - x_0^*) + p_{X^*}(A^* z_{\rm i}^* - A^* z_0^*) \\
&\leq \frac{\varepsilon}{2} + \sqrt{\eta}\,(p_{X^*}(A^* u_{\rm i}^*) + p_{X^*}(A^* z_0^* - x_0^*) + p_{X^*}(x_0^*)) \\
&\leq \frac{\varepsilon}{2} + \sqrt{\eta}\,(p_{X^*}(A^* u_{\rm i}^*) + \eta + p_{X^*}(x_0^*)) \\
&\leq \frac{\varepsilon}{2} + \sqrt{\eta} \left(\max_{U^\circ} p_{X^*}(A^*(y^*)) + p_{X^*}(x_0^*) + 1 \right) \leq \varepsilon.
\end{aligned}
$$

To prove the necessity part, let $(y_{\rm i}, y_{\rm i}^*)_{{\rm i} \in I} \subset \operatorname{gph} \partial g$ a net such that $(y_{\rm i}) \to y = Ax$, $g(y_{\rm i}) \to g(y)$, $\langle y_{\rm i} - y, y_{\rm i}^* \rangle \to 0$ and $A^*(y_{\rm i}^*) \xrightarrow{w^*} x^*$. Then $\langle y - y_{\rm i}, y_{\rm i}^* \rangle \leq g(y) - g(y_{\rm i})$ for every ${\rm i} \in I$ and $y \in X$. It follows that

$$\langle z - x, A^* y_{\rm i}^* \rangle + \langle y - y_{\rm i}, y_{\rm i}^* \rangle = \langle Az - y_{\rm i}, y_{\rm i}^* \rangle \leq f(z) - g(y_{\rm i}) \quad \forall {\rm i} \in I,\ \forall z \in X.$$

Taking the limits gives $\langle z - x, x^* \rangle \leq f(z) - f(x)$, for all $z \in X$, and so $x^* \in \partial f(x)$. □

Remark 3.13 We give a simple example which shows that the above subdifferential calculus rule is not valid without the epi-pointedness assumption. Let $g : X \to \overline{\mathbb{R}}$ be a convex lsc and proper function with empty subdifferential everywhere (see [19]). Then take a continuous linear function $A : \mathbb{R} \to X$ such that $\operatorname{dom} g \cap A(\mathbb{R}) \neq \emptyset$. We easily check that the function $f = g \circ A$ is proper, convex, and lsc in $\mathbb{R}$. This implies that there exists a point $x_0 \in \mathbb{R}$ such that $\partial f(x_0) \neq \emptyset$ and Theorem 3.12 does not hold.

From the last theorem we deduce the subdifferential limiting calculus rule for the sum of convex epi-pointed functions.

Theorem 3.14 *Let $f_1, f_2 \in \Gamma_0(X)$ be two epi-ponted functions and $x \in \operatorname{dom}(f_1 + f_2)$. Then $x^* \in \partial(f_1 + f_2)(x)$ if and only if there exist two nets $(x_{k,{\rm i}}, x_{k,{\rm i}}^*)_{{\rm i} \in I} \subset X \times X^*$ such that $x_{k,{\rm i}}^* \in \partial f_k(x_{k,{\rm i}})$ $k = 1, 2$, $x_{k,{\rm i}} \to x$, $f_k(x_{k,{\rm i}}) \to f_k(x)$, $\langle x_{k,{\rm i}} - x, x_{k,{\rm i}}^* \rangle \to 0$, for $k = 1, 2$, and $(x_{1,{\rm i}}^* + x_{2,{\rm i}}^*) \xrightarrow{w^*} x^*$.*

Proof. We apply Theorem 3.12 with $Z = X \times X$, $A : X \to Z$ defined by $Ax = (x, x)$ and $g(x_1, x_2) = f_1(x_1) + f_2(x_2)$. We have $f := f_1 + f_2 = g \circ A$. The sufficiency part is immediate, taking $y = (x, x) = Ax$, $y_{\rm i} = (x_{1,{\rm i}}, x_{2,{\rm i}})$ and $y_{\rm i}^* = (x_{1,{\rm i}}^*, x_{2,{\rm i}}^*)$, for ${\rm i} \in I$. For the necessity part, we take $x^* \in \partial f(x)$. By Theorem 3.12 there exists $(y_{\rm i}, y_{\rm i}^*)_{{\rm i} \in I} \subset Z \times Z^*$ such that $y_{\rm i}^* \in \partial g(y_{\rm i})$, $y_{\rm i} \to y = Ax$, $g(y_{\rm i}) \to g(y)$, $\langle y_{\rm i} - y, y_{\rm i}^* \rangle \to 0$ and $A^*(y_{\rm i}^*) \xrightarrow{w^*} x^*$. Taking $y_{\rm i} = (x_{1,{\rm i}}, x_{2,k})$ and $y_{\rm i}^* = (x_{1,{\rm i}}^*, x_{2,{\rm i}}^*)$, by the formula $\partial g(x_1, x_2) = \partial f_1(x_1) \times \partial f_2(x_2)$, we get $x_{k,{\rm i}}^* \in \partial f_k(x_{k,{\rm i}})$ and $x_{k,{\rm i}} \to x$ for $k = 1, 2$. Suppose that $\limsup f_1(x_{1,{\rm i}}) - f_1(x) > \delta > 0$. Then $J := \{{\rm i} \in I : f_1(x_{1,{\rm i}}) - f_1(x) > \delta\}$ is a co-final set in I. It follows that $g(y_{\rm i}) - g(y) \geq \delta + f_2(x_{2,{\rm i}}) - f_2(x)$ for every ${\rm i} \in J$. Then taking the lower limits we get $0 \geq \delta$ and, hence, $f_k(x_{k,{\rm i}}) \to f_k(x)$, for $k = 1, 2$. Finally, using the following inequalities

$$f_1(x_{1,{\rm i}}) - f_1(x) \leq \langle x_{1,{\rm i}} - x, x_{1,{\rm i}}^* \rangle \leq \langle x_{1,{\rm i}} - x, x_{1,{\rm i}}^* \rangle + \langle x_{2,{\rm i}} - x, x_{2,{\rm i}}^* \rangle + f_2(x) - f_2(x_{2,{\rm i}})\ \forall {\rm i} \in I,$$

and taking the limits, we infer that $\langle x_{1,\mathrm{i}} - x, x^*_{1,\mathrm{i}}\rangle \to 0$. $\square$

3.3 Banach spaces

In this last section, we show how to recover from our previous results the classical Brøndsted-Rockafellar Theorem in the context of Banach spaces, for any proper lsc convex functions which are not necessarily epi-pointed. In the case of reflexive spaces, it is an easy exercise, using adequate perturbations of the convex function in order to obtain an epi-pointed function, in the line of Examples 1.10 and 1.12.

Proposition 3.15 *Let X be a reflexive Banach space and $f \in \Gamma_0(X)$. Consider $x_0 \in \operatorname{dom} f$, $\varepsilon \geq 0$ and $x_0^* \in \partial_\varepsilon f(x_0)$. Then there exist $(x_\varepsilon, x_\varepsilon^*) \in X \times X^*$ such that $x_\varepsilon^* \in \partial f(x_\varepsilon)$, $\|x_\varepsilon - x_0\| \leq \sqrt{\varepsilon}$ and $\|x_\varepsilon^* - x_0^*\| \leq \sqrt{\varepsilon}$. In particular,* $\operatorname{dom} f \subset \overline{\operatorname{dom} \partial f}$ *and* $\operatorname{dom} f^* \subset \overline{\operatorname{rge} \partial f}$.

Proof. Consider $x_0^* \in \partial_\varepsilon f(x_0)$. We define the function $g(w) := f(w) + \delta_{\mathbb{B}(0,M)}(w)$, where $M \geq \|x_0\| + \varepsilon + \sqrt{\varepsilon}$. It is easy to see that g is epi-pointed, $\operatorname{dom}(g^*) = X^*$ and $x_0^* \in \partial_\varepsilon g(x_0)$, hence we apply Theorem 3.8 with $\lambda = \sqrt{\varepsilon}$, and $\beta = 0$, to get the existence of $(x_\varepsilon, x_\varepsilon^*) \in X \times X^*$ such that $x_\varepsilon^* \in \partial g(x_\varepsilon))$, $\|x_0 - x_\varepsilon\| \leq \sqrt{\varepsilon}$ and $\|x_\varepsilon^* - x_0^*\| \leq \sqrt{\varepsilon}$. From the fact that $x_\varepsilon \in \operatorname{int} \mathbb{B}(0, M)$, we conclude that $x_\varepsilon^* \in \partial f(x_\varepsilon)$. $\square$

Proposition 3.16 *Let X be a reflexive Banach space and $f \in \Gamma_0(X)$. Then ∂f is a maximal monotone operator.*

Proof. Let $f \in \Gamma_0(X)$ such that for every $(x, x^*) \in \operatorname{gph} \partial f$ we have $\langle x^*, x\rangle \geq 0$. We take the function $g(w) := f(w) + \frac{1}{2}\|x\|^2$. Clearly, g is epi-pointed, and so by Theorem 3.10 ∂g is a maximal monotone operator. Moreover, for every $x \in X$, $\partial g(x) = \partial f(x) + \partial \frac{1}{2}\|\cdot\|^2(x)$, where $\partial\frac{1}{2}\|\cdot\|^2(x) = \{x^* \in X^* : \langle x^*, x\rangle = \|x\|^2 = \|x^*\|^2\}$. Then, for every $w^* \in \partial g(x)$, $\langle w^*, x\rangle \geq 0$ and by the maximality we get $0 \in \partial g(0)$. Finally, because $\partial f(0) = \partial g(0)$ we obtain $0 \in \partial f(0)$. $\square$

Proposition 3.17 *Let X be an lcs, Z be a reflexive Banach space, $A \in \mathcal{L}(X, Z)$, $g \in \Gamma_0(Z)$, $f := g \circ A$ and $x \in \operatorname{dom} f$. Then $x^* \in \partial f(x)$ if and only if there exists a net $(y_\mathrm{i}, y_\mathrm{i}^*)_{\mathrm{i}\in I} \subset Z \times Z^*$ such that $y_\mathrm{i}^* \in \partial g(y_\mathrm{i})$, $y_\mathrm{i} \to y = Ax$, $g(y_\mathrm{i}) \to g(y)$, $\langle y_\mathrm{i} - y, y_\mathrm{i}^*\rangle \to 0$ and $A^*(y_\mathrm{i}^*) \xrightarrow{w^*} x^*$.*

Proof. Consider the functions $\tilde{g} := g + \delta_{\mathbb{B}(A(x),1)}$ and $\tilde{f} := \tilde{g} \circ A$. We apply Theorem 3.12 to $\tilde{f}$ and obtain the existence of a net $(y_\mathrm{i}, y_\mathrm{i}^*)_{\mathrm{i}\in I} \in Z \times Z^*$ such that $y_\mathrm{i}^* \in \partial\tilde{g}(y_\mathrm{i})$, $y_\mathrm{i} \to y = Ax$, $g(y_\mathrm{i}) \to g(y)$, $\langle y_\mathrm{i} - y, y_\mathrm{i}^*\rangle \to 0$ and $A^*(y_\mathrm{i}^*) \xrightarrow{w^*} x^*$. Because $y_\mathrm{i} \to y = Ax$ we can suppose that $y_\mathrm{i} \in \operatorname{int} \mathbb{B}(A(x), 1)$, so we have $\tilde{g}(y_\mathrm{i}) = g(y_\mathrm{i})$ and by Proposition 1.16(a) we have $y_\mathrm{i}^* \in \partial g(y_\mathrm{i})$. $\square$

Finally we show that Brøndsted-Rockafellar's Theorem can also be obtained from Lemma 3.7.

Theorem 3.18 *Let X be a Banach space and $h : X \to \overline{\mathbb{R}}$ a convex lsc function. If $x_0^* \in \partial_{\varepsilon^2} h(x_0)$, then for every $\eta > 0$ there exist $x_\varepsilon \in X$ and $x_\varepsilon^* \in X^*$ such that $\|x_0 - x_\varepsilon\| < \varepsilon + \eta$, $\|x_0^* - x_\varepsilon^*\| < \varepsilon + \eta$ and $x_\varepsilon^* \in \partial h(x_\varepsilon)$.*

Proof. Take a sequence of positive numbers $\{\eta_n\}_{n\geq 1}$, such that $\sum_{n=1}^{\infty} \eta_n < \eta$ and $\eta_0 := \varepsilon$. We claim that if $x_n^* \in \partial_{\eta_n^2} h(x_n)$, then there exists $(x_{n+1}, x_{n+1}^*) \in X \times X^*$ such that $\|x_n - x_{n+1}\| \leq \eta_n$, $\|x_n^* - x_{n+1}^*\| \leq \eta_n$ and $x_{n+1}^* \in \partial_{\eta_{n+1}^2} h(x_{n+1})$. Take $f := h^*$ and $\rho := \eta_n \|\cdot\|$ and consider the duality pair $(X^*, w^*, X, \|\cdot\|)$. Since $x_n \in \partial_{\eta^2{}_n} f(x_n^*) \cap \operatorname{int} \operatorname{dom}(f + \rho(\cdot - x_n^*))^*$ $(\operatorname{dom}(f + \rho(\cdot - x_n^*))^* \supseteq \operatorname{dom} h + \eta_n \mathbb{B}_X(0,1))$, we apply Lemma 3.7 and we conclude that there exists $x_{n+1}^* \in X^*$ such that $\|x_{n+1}^* - x_n^*\| \leq \eta_n$ and $x_n \in \partial(f + \rho(\cdot - x_n^*))(x_{n+1}^*)$. By applying Proposition 1.16(a) in X^{**}, $x_n \in \partial f^*(x_{n+1}^*) + \eta_n \mathbb{B}_{X^{**}}$, and so there exists $x^{**} \in X^{**}$ such that $x_{n+1}^* \in \partial f^*(x^{**})$ and $\|x_n - x^{**}\| \leq \eta_n$. Finally, we apply Proposition 1.16(d) to h and $\delta_{\mathbb{B}_X(x_n,\eta_n)}$ with the duality pair $(X, \|\cdot\|, X^*, w^*)$ to get $h^* \square \sigma_{\mathbb{B}_X(x_n,\eta_n)} = (h + \delta_{\mathbb{B}_X(x_n,\eta_n)})^*$. Next, we apply Proposition 1.16(c) to h^* and $\sigma_{\mathbb{B}_X(x_n,\eta_n)} = \sigma_{\mathbb{B}_{X^{**}}(x_n,\eta_n)}$ with the duality pair $(X^*, \|\cdot\|_*, X^{**}, w^*)$ to get

$$\operatorname{epi}(h^{**} + \delta_{\mathbb{B}_{X^{**}}(x_n,\eta_n)}) = \operatorname{epi}(h + \delta_{\mathbb{B}_X(x_n,\eta_n)})^{**} = \overline{\operatorname{epi}}^{w^*}(h + \delta_{\mathbb{B}(x_n,\eta_n)}).$$

Therefore, there exist a net $(x_{\mathrm{i}}, \alpha_{\mathrm{i}})_{\mathrm{i}\in I} \in \operatorname{epi}(h + \delta_{\mathbb{B}(x_n,\eta_n)}) \subseteq X \times \mathbb{R}$ such that $x_{\mathrm{i}} \xrightarrow{w^*} x^{**}$ and $\alpha_{\mathrm{i}} \to h^*(x^{**})$, which implies the existence of an element $\mathrm{i}_0 \in I$ such that $x_{\mathrm{i}_0} \in \mathbb{B}(x_n, \eta_n)$, $\frac{\eta_{n+1}^2}{2} + h^{**}(x^{**}) > \alpha_{\mathrm{i}_0} \geq h(x_{\mathrm{i}_0})$ and $\frac{\eta_{n+1}^2}{2} + \langle x_{n+1}^*, x_{\mathrm{i}_0} \rangle > \langle x_{n+1}^*, x^{**} \rangle$ and set $x_{n+1} := x_{\mathrm{i}_0}$. Then we get

$$h(x_{n+1}) + h^*(x_{n+1}^*) \leq h^{**}(x^{**}) + h^*(x_{n+1}^*) + \frac{\eta_{n+1}^2}{2} \leq \langle x^{**}, x_{n+1}^* \rangle + \frac{\eta_{n+1}^2}{2} \leq \langle x_{n+1}^*, x_{n+1}^* \rangle + \eta_{n+1}^2,$$

so the construction of the sequences x_n and x_n^* is done. From the facts that $\|x_n - x_{n+1}\| \leq \eta_n$, $\|x_n^* - x_{n+1}^*\| \leq \eta_n$ and $\sum_{\mathrm{i}=1}^{\infty} \eta_{\mathrm{i}} < +\infty$, it follows that (x_n, x_n^*) is a Cauchy sequence. By the completeness of X and X^*, we conclude that $(x_n, x_n^*) \xrightarrow{\|\cdot\|} (x_\varepsilon, x_\varepsilon^*)$, and so $\|x_0 - x_\varepsilon\| < \varepsilon + \eta$, $\|x_0^* - x_\varepsilon^*\| < \varepsilon + \eta$. Hence, $x_\varepsilon^* \in \partial h(x_\varepsilon)$. $\square$

Chapter 4

Extensions of Zagrodny's Mean Value Theorem, integration theorems and characterization of convexity for a generalized subdifferential of functions defined in locally convex spaces

4.1 Subdifferential on locally convex spaces

Now we begin with an extension of the definition (and also a more general definition in Banach spaces) of quasi presubdifferential. This definition has been adapted from [114] and motivated by the definition of the approximate subdifferential given by Ioffe [61], where the author considers a reduction to finite dimensional spaces and also a reduction to subspaces with nice properties for the Dini-Hadammard subdifferential (called by the author weak trustworthy spaces) using the ε-enlargement of the Dini-Hadammard subdifferential. Also the ε-enlargement can be find in the extension of the Limiting/Mordukhovich subdifferential in any arbitrary Banach space (see for example [84, Chapter I]).

We will simply call this object subdifferential instead of quasi presubdifferential as in [114].

Definition 4.1 (Family of subdifferential) *Consider a directed (by inclusion) family of spaces $\mathcal{L}$ of X which cover X and a directed set $(\mathbb{I}, \preceq)$. A net of operators $\{\hat{\partial}_{\mathrm{i},L} : \mathrm{i} \in \mathbb{I}, L \in \mathcal{L}\}$ (ordered by $(\mathrm{i}_1, L_1) \leq (\mathrm{i}_2, L_2)$ iff $\mathrm{i}_1 \preceq \mathrm{i}_2$ and $L_1 \subseteq L_2$) is a family of subdifferentials for a function $f \in \overline{\mathbb{R}}^X$ if,*

1. $\hat{\partial}_{\mathrm{i},L} f(x) \subseteq X^*$ *for all* $x \in X$, $\mathrm{i} \in \mathbb{I}$ *and* $L \in \mathcal{L}$.

2. $\hat{\partial}_{\mathrm{i},L} f(x) = \emptyset$ *for all* $x \notin \operatorname{dom} f$, $\mathrm{i} \in \mathbb{I}$ *and* $L \in \mathcal{L}$.

3. *For every* $L \in \mathcal{L}$ *and every* $g : X \to \mathbb{R}$ *convex and Lipschitz, if* x_0 *is a local minimum*

of $f+g$ relative to L with $f(x_0)\in\mathbb{R}$, there exist nets $(x_\alpha,x_\alpha^*)_{\alpha\in\mathbb{D}}\in X\times X^*$ and a subnet $(\hat\partial_{\mathrm{i}_\alpha,L})_{\alpha\in\mathbb{D}}$ of $(\hat\partial_{\mathrm{i},L})_{\mathrm{i}\in\mathbb{I}}$ such that $x_\alpha\in L$, $x_\alpha^*\in\hat\partial_{\mathrm{i}_\alpha,L}f(x_\alpha)$, $x_\alpha\xrightarrow{f}x_0$, $x_\alpha^*\xrightarrow{w^*}x_0^*$ with $x_0^*\in-\partial g(x_0)+L^\perp$ and $\liminf\langle x_\alpha^*,x_0-x_\alpha\rangle\ge 0$.

When $\{\hat\partial\}=\{\hat\partial_{L,\mathrm{i}}\}$ is a singleton we simply say $\hat\partial$ is a subdiferential for the function f.

Property 3 is a kind of fuzzy controlled sum rule (see [72] and the reference therein for more details and discussions about these properties). It is not difficult to see that all the subdifferentials mentioned in the introduction satisfies Definition 4.1 when the function is defined in a Banach space with appropriate smooth properties, for example: The G-subdifferential of Ioffe and the Clarke-Rockafellar subdifferential in all Banach spaces, the Fréchet subdifferential and the Limiting/Mordukovich subdifferential in Asplund spaces, the Dini-Hadammard subdifferential in Banach spaces with Găteaux differentiable renorm, the proximal subdifferential in Hilbert spaces, etc.

Example 4.2 (Examples of subdifferentials on locally convex spaces) The convex subdifferential and the ε-subdifferential form a family of subdifferentials for proper convex lsc functions. The generalized Clarke-Rockafellar subdifferential, given by

$$\partial_C f(x)=\{x^*\in X^*:\langle x^*,h\rangle\le f^\uparrow(x;u),\ \forall h\in X\},$$

where $f^\uparrow(x;u)$ is the generalized Clarke-Rockafellar directional derivative or the lower subderivative of f at x with respect to u (see e.g. [104]), given by

$$f^\uparrow(x;v)=\limsup_{\substack{x'\xrightarrow{f}x\\ t\to0^+}}\inf_{v'\to v}\frac{f(x'+tv+)-f(x')}{t},$$

is also a subdifferential for every proper lsc function.

The approximate subdifferential, introduced by Ioffe (see, e.g. [60–62]), is given (for a point $x\in\operatorname{dom}f$) by

$$\partial_A f(x)=\bigcap_{L\in\mathcal{F}}\limsup_{u\xrightarrow{f}x}\partial^- f_{u+L}(u),$$

where $f_{u+L}:=f+I_{u+F}$, $\mathcal{F}$ denotes the collection of all finite dimensional subspaces of X, ∂^- is the Dini subdifferential and the lim sup is considered with respect to the weak* topology in X^*. So the families ∂_A and $\{\partial_L^-\}_{L\in\mathcal{F}}$ ($\partial_L^- f(x):=\partial^- f_{x+L}(x)$) are families of subdifferentials for lsc functions. Also for any admissible family $\mathcal{L}$ of weak trustworthy spaces of X (see e.g. [59, 61]), the family $\{\partial_{\varepsilon,L}^-:\varepsilon>0;L\in\mathcal{F}\}$ given by $\partial_{\varepsilon,L}^- f(x):=\partial_\varepsilon^- f_{x+L}(x)$, where ∂_ε^- is the ε-Dini subdifferential.

Given a subdifferential ∂ in the sense of [65, Definition 2.1.] and a family of subspaces $\mathcal{L}$ of X such that ∂ is trusted on $\mathcal{L}$ (see [65, Definition 2.12.]) and $\mathcal{L}$ covers X (for example $\mathcal{L}=\mathcal{F}$), then the family $\partial_L f(x):=\partial f_{x+L}(x)$ with $L\in\mathcal{F}$ is a family of subdifferentials.

The next result is a Theorem on iterated limits, also it can be understood as a diagonal argument for generalized sequences. The proof is very simple but the statement is not widely known, for this reason we prefer to present this result and make a small remark about applications to lower and upper limits of nets. This result will be crucial in our investigation, in the sense that it allows us to generalize the argument utilized in the proof of the classical result, which considers sequences in the definition of subdifferential instead of nets as in our setting.

Proposition 4.3 *[73, Chapter 2 - Theorem 4, p. 69] Let D be a directed set, let E_n be a directed set for each $n \in D$, let F be the product $D \times \Pi_{n \in D} E_n$, and for (m, f) in F let $R(m, f) = (m, f(m))$. If $S_{(m,n)}$ is a member of a topological space for each $m \in D$ and each $n \in E_m$, then $S \circ R$ converges to $\lim_m \lim_n S(m, n)$ whenever this iterated limit exists.*

Remark 4.4 It is important to recall that the previous Proposition is also valid to upper and lower limits in the sense that if $S_{(m,n)}$ are real numbers such that for every $n \in D$, $\liminf S_{(m,n)} \geq S_n$ then $\liminf S \circ R \geq \liminf S_n$. It can be easily proved using the lower topology on $\mathbb{R} \cup \{+\infty\}$ generated by the base of sets of the form $A_a := \{r : r > a\}$ with $a \in [-\infty, +\infty)$, so with this topology it is easy to prove that a net $s_\alpha \to s$ if and only if $\liminf s_\alpha \geq s$ (see e.g. [73, Chapter 3, Exercise F]).

Lemma 4.5 *Consider $C \subseteq X$ convex and closed and $p : X \to \mathbb{R}$ a continuous seminorm, define $\mathrm{d}_C^p : X \to [0, +\infty)$ by $\mathrm{d}_C^p(x) := \inf_{z \in C} p(x - z)$. Then we have that d_C is Lipschitz and convex, in addition, if there exists $\bar{z} \in C$ such that $\mathrm{d}_C^p(x) = p(x - \bar{z})$, then $\partial_c \mathrm{d}_C^p(x) \subseteq \partial_c p(x - \bar{z}) \cap N_C(\bar{z}) \subseteq \mathbb{B}_p(0, 1)^\circ \cap N_C(\bar{z})$.*

Proof. First it is clear that d_C is Lipschitz and convex, now given $x^* \in \partial \mathrm{d}_C(x)$, $y \in X$ and $\bar{z} \in C$ such that $\mathrm{d}_C^p(x) = p(x - \bar{z})$

$$\langle x^*, y - (x - \bar{z}) \rangle = \langle x^*, y + \bar{z} - x \rangle \leq \mathrm{d}_C^p(y + \bar{z}) - \mathrm{d}_C(x) \leq p(y) - p(x - \bar{z}),$$

which means, $x^* \in \partial p(x - \bar{z}) \subseteq B_p(0, 1)^\circ$. Now consider $y \in C$, then

$$\begin{aligned} \langle x^*, y - \bar{z} \rangle &= \langle x^*, y - x \rangle + \langle x^*, x - \bar{z} \rangle \leq \mathrm{d}_C^p(y) - \mathrm{d}_C^p(x) + \langle x^*, x - \bar{z} \rangle \\ &= 0 - p(x - \bar{z}) + \langle x^*, x - \bar{z} \rangle \leq 0. \end{aligned}$$

□

Two important tools are needed to develop subdifferential theory, one of them is the definition of subdifferential and the second is the existence of variational principles. In the setting of a Banach space these techniques have been well studied in the last fifty years for lsc functions, such as Ekeland's variational principle, Borwein-Preiss's variational principle, Godefroy's variational principles and smooth variational principles in Asplund spaces among many others ([14, 45, 84, 89]). Outside of Banach space it is impossible to generalize these tools for all lsc functions, for this reason the family of w-lsc epi-pointed functions appears to be promising to develop techniques of variation analysis in lcs spaces (see [30] for an extension of classical results of convex analysis in Banach spaces to epi-pointed functions in lcs spaces). The next lemma is a kind of variational principle for this class of functions in a general locally convex space.

Lemma 4.6 *An lsc proper convex function $g : X^* \to \overline{\mathbb{R}}$ is $\tau(Y,X)$-continuous at $x^* \in \operatorname{dom} g$ if and only if $g^* - x^*$ is $w(X,Y)$-infcompact. Consequently a w-lsc epi-pointed function f satisfies that $f - x^*$ is w-infcompact for every $x^* \in \operatorname{int}(\operatorname{dom} f^*)$.*

Proof. The first part is a classical result in convex analysis (see e.g. [88] and also [77, chapter 6]). The second part is an easy application of the previous result, taking into account that $f^{**} - x^*$ is $w(X,Y)$-infcompact for every $x^* \in \operatorname{int}(\operatorname{dom} f^*)$ and $f^{**} \le f$, then $\{f - x^* \le \lambda\} \subseteq \{f^{**} - x^* \le \lambda\}$. □

Lemma 4.7 *Consider f a w-lsc and epi-pointed function, $a, b \in X$, $x_0^* \in X^*$. Then there exists a continuous seminorm ρ_0 such that for all seminorm $\rho \ge \rho_0$ and $M \ge 1$ the function $H(x) := f(x) - \langle x_0^*, x\rangle + M\mathrm{d}^{\rho}_{[a,b]}(x)$ is w-infcompact. Consequently if c is an ε^2-minimun of the function H. Then for every positive lsc convex function g there exists c_ε such that $g(c_\varepsilon - c) \le \varepsilon$ and c_ε is a minimum of the function $H(\cdot) + \varepsilon g(\cdot - c)$.*

Proof. Take $y_0^* \in \operatorname{int}(\operatorname{dom} f^*)$, consider $\rho_0(x) := |\langle y_0^* - x_0^*, x\rangle|$, so for all $c \in X$

$$\begin{aligned} f(x) - \langle x_0^*, x\rangle + M\rho_0(x - c) &\ge f(x) - \langle x_0^*, x\rangle + \langle x_0^* - y_0^*, x - c\rangle \\ &= f(x) - \langle y_0^*, x\rangle - \langle x_0^* - y_0^*, c\rangle \end{aligned}$$

Therefore $f(x) - \langle x_0^*, x\rangle + M\rho_0(x - c)$ is w-infcompact, hence for all seminorm $\rho \ge \rho_0$, $f(x) - \langle x_0^*, x\rangle + M\rho(x - c)$ is w-infcompact. Now we show that $H(x) := f(x) - \langle x_0^*, x\rangle + M\mathrm{d}^{\rho}_{[a,b]}(x)$ is w-infcompact. Indeed consider $\alpha \in \mathbb{R}$ and a net $(x_n)_{n\in D} \subset \{x \in X : H(x) \le \alpha\}$, since $[a,b]$ is compact for all $n \in D$ there exists $u_n \in [a,b]$ such that $\mathrm{d}^{\rho}_{[a,b]}(x_n) = \rho(x_n - u_n)$. Moreover (taking a subnet) we can assume $u_n \to u \in [a,b]$. Now take $n \in D$ such that for all $n \ge n_0$, $\rho(u_n - u) \le 1$, then for all $n \ge n_0$, $x_n \in \Gamma := \{x \in X : f(x) - \langle x_0^*, x\rangle + M\rho(x - v) \le \beta + M\}$. Finally because Γ is compact, there exists a subnet $(z_j) \to z$ of (y_i) (therefore a subnet of (x_n)). The second part follows from the fact that the function $H(\cdot) + \varepsilon g(\cdot - c)$ is also w-lsc and infcompact, so $H(\cdot) + \varepsilon g(\cdot - c)$ attains its minimum, thus $g(c_\varepsilon - c) \le \varepsilon$. □

In the engineering thesis [92] the following extension of the Zagrodny Mean Value Theorem for w-lsc and epi-pointed function in arbitrary locally convex spaces was proved.

Proposition 4.8 *Let (X,τ) be an arbitrary lcs and consider a subdifferential $\hat{\partial}$ for a w-lsc and epi-pointed function f. Then for every $a, b \in X$ with $a \in \operatorname{dom} f$ and $a \ne b$, $r \in \mathbb{R}$ with $r \le f(b)$ and a continuous seminorm p such that $p(a - b) \ne 0$ there exists a net $(x_\alpha, x_\alpha^*)_{\beta\in\mathbb{A}} \in X \times X^*$ and such that $(x_\alpha)_{\alpha\in\mathbb{A}} \xrightarrow{f} c \in [a, b[$, $x_\alpha^* \in \hat{\partial} f(x_\alpha)$,*

(i) $r - f(a) \le \liminf\langle b - a, x_\alpha^*\rangle$.

(ii) $0 \le \liminf\langle c - x_\alpha, x_\alpha^*\rangle$.

(iii) $\frac{p(b-c)}{p(b-a)}(r - f(a)) \le \liminf\langle b - x_\alpha, x_\alpha^*\rangle$.

(iv) $p(b-a)(f(c) - f(a)) \le p(c-a)(r - f(a))$.

This first step gives us the idea to extend this result to more abstract constructions in lcs spaces. In this paper we present the following generalization of the above result. The proof of the result is an adaptation of [114], and hence, as stated by the author, an adaptation of the original work of Zagrodny [130].

Theorem 4.9 *Consider a family of subdifferentials $\{\hat{\partial}_{\mathrm{i},L} : \mathrm{i} \in \mathbb{I},\ L \in \mathcal{L}\}$ for a given proper lsc function f. Assume that one of the following conditions holds:*

(a) The topology on X is generated by a family of seminorms $(\rho_L)_{L\in\mathcal{L}}$, where for every $L \in \mathcal{L}$, (L, ρ_L) is a Banach space, and $\rho_M \leq \rho_L$ for all $M \subseteq L$.

(b) f is a w-lsc and epi-pointed function.

Then, for every $a, b \in X$ with $a \in \operatorname{dom} f$ and $a \neq b$, $r \in \mathbb{R}$ with $r \leq f(b)$ and every continuous seminorm p such that $p(a-b) \neq 0$, there exist $c \in [a, b[$ and $(x_\alpha, x_\alpha^*)_{\alpha\in\mathbb{A}} \in X \times X^*$ such that $\{(\mathrm{i}_\alpha, L_\alpha)\}_{\alpha\,\in\mathbb{A}}$ is cofinal in $\mathbb{I} \times \mathcal{L}$, $(x_\alpha)_{\alpha\in\mathbb{A}} \xrightarrow{f} c$, $x_\alpha^* \in \hat{\partial}_{\mathrm{i}_\alpha, L_\alpha} f(x_\alpha)$ with $x_\alpha \in L_\alpha$, and

(i) $r - f(a) \leq \liminf_\alpha \langle b - a, x_\alpha^* \rangle$,

(ii) $0 \leq \liminf_\alpha \langle c - x_\alpha, x_\alpha^* \rangle$,

(iii) $\frac{p(b-c)}{p(b-a)}\,(r - f(a)) \leq \liminf_\alpha \langle b - x_\alpha, x_\alpha^* \rangle$,

(iv) $\frac{p(b-a)}{p(c-a)}(f(c) - f(a)) \leq (r - f(a))$.

Moreover, if $c \neq a$, then one has $\lim_\alpha \langle b - a, x_\alpha^* \rangle = r - f(a)$.

Proof. Take $x^* \in X^*$ such that $\langle b - a, x^* \rangle = r - f(a)$, and consider the function $h = f - x^*$; hence, $h(a) \leq h(b)$. Because h is lsc, there exists $c \in [a, b[$ such that $h(c) \leq h(x)$, for all $x \in [a, b]$. Take $\bar{\lambda} \in [0, 1[$ such that $c = (1 - \bar{\lambda})a + \bar{\lambda}b$. It follows that $c - a = \bar{\lambda}(b - a)$ and $p(c - a) = \bar{\lambda}p(b - a)$; hence, $f(c) - f(a) = h(c) - h(a) + \langle c - a, x^* \rangle \leq \bar{\lambda}(r - f(a))$, and (iv) follows.

Let $\gamma < h(c)$, so that from the lower semicontinuity of h and the compactness of the closed segment $[a, b]$, we can find (a convex, closed and balanced neighborhood) $U_\gamma \in \mathcal{N}_0$ such that $\gamma < h(x)$ for all $x \in [a, b] + U_\gamma$.

In what follows, we choose ρ_0 as in Lemma 4.7 (associated with f, a, b and x^*) if f is a w-lsc and epi-pointed function, and $\rho_0 = 0$, otherwise. From the previous paragraph, there are $U_n \in \mathcal{N}_0$, $n \geq 1$, such that $U_{n+1} \subseteq U_n \subset \mathbb{B}_{p+\rho_0} = \{x \in X : p(x) + \rho_0(x) \leq 1\}$ (w.l.o.g.) and

$$h(c) - \frac{1}{n^2} < h(x) \quad \forall n \geq 1 \text{ and } \forall x \in [a, b] + U_n.$$

We denote $\rho_n := \sigma_{U_n^\circ}$ $(\geq \rho_0)$ and $\mathrm{d}^n_{[a,b]}(x) := \mathrm{d}^{\rho_n}_{[a,b]}(x)$, so that

$$h(c) \leq h(x) + \mathrm{d}^n_{[a,b]}(x) + \frac{1}{n^2}, \quad \forall x \in U := [a, b] + U_1; \tag{4.1}$$

indeed, this inequality trivially holds when $x \in [a, b] + U_n$, while for $x \in U \setminus ([a, b] + U_n)$, we have that $\mathrm{d}^n_{[a,b]}(x) \geq 1$ and, so,

$$h(x) + \mathrm{d}^n_{[a,b]}(x) \geq h(c) - 1 + 1 \geq h(c) - \frac{1}{n^2}.$$

Now, we consider the functions $H_n(x) := h(x) + \mathrm{d}^n_{[a,b]}(x) + \delta_U(x)$, $n \geq 1$, which satifies $H_n(c) \leq \inf_U H_n + \frac{1}{n^2}$, thanks to (4.1). To continue, we proceed by analyzing the alternative of the current theorem:

1. If we are in case ((a)), we may assume w.l.o.g. that the ρ_L's are such that

$$\mathbb{B}_{\rho_L} \subseteq \operatorname{int}(U_1),\ \forall L \in \mathcal{L}.$$

Then, for each $n \geq 1$ and $L \in \mathcal{L}$, by applying the Ekeland variational principle in the Banach space (L, ρ_L) we find a minimum $u \in L$ of the function $H_n(\cdot) + n^{-1}\rho_L(\cdot - u)$ on L such that $\rho_L(c - u) \leq 1/n$.

2. In the case of ((b)), we choose a family of seminorms $\{\rho_j\}_{\in \mathbb{J}}$ that generates the topology on X and such that $\mathbb{B}_{\rho_j} \subseteq \operatorname{int}(U_1)$, $\forall j \in \mathbb{J}$, where $\mathbb{J}$ is an ordered set satisfying $\rho_{j_1} \leq \rho_{j_2}$ whenever $j_1 \leq j_2$. Then, by applying Lemma 4.7, for each $n \geq 1$, $j \in \mathbb{J}$, and $L \in \mathcal{L}$, we find a minimum $u \in L$ of the function $H_n(\cdot) + n^{-1}\rho_j(\cdot - c) + \delta_L$ such that $\rho_j(c - u) \leq 1/n$.

Consequently, in each one of the two cases above there exist an ordered set $\mathbb{J}$, a cofinal set $\mathbb{D} \subseteq \mathbb{N} \times \mathbb{J} \times \mathcal{L}$ and nets $(u_{n,j,L})_{(n,j,L)\in\mathbb{D}}$, $(v_{n,j,L})_{(n,j,L)\in\mathbb{D}}$ such that

$$p_j(u_{n,j,L} - c) \leq \frac{1}{n},\ H_n(u_{n,j,L}) \leq H_n(c),\ u_{n,j,L} \text{ is a minimum of } H_n + n^{-1}p_j(\cdot - v_{n,j,L});$$

hence,

$$u_{n,j,L} \to c. \tag{4.2}$$

Moreover, since $\mathbb{B}_{\rho_j} \subseteq \operatorname{int}(U_1)$, we obtain that $(u_{n,j,L}) \subset \operatorname{int}(U)$, and this yields that each $u_{n,j,L}$ is a local minimum of the function $f - x^* + \mathrm{d}^n_{[a,b]} + n^{-1}p_j(\cdot - v_{n,j,L})$ over L. Thus, by Definition 4.1, for each $\omega := (n, j, L) \in \mathbb{D}$ there exists a net $(x_{\alpha,\omega}, x^*_{\alpha,\omega})_{\alpha\in\Lambda_\omega} \subseteq X \times X^*$, together with elements

$$v^*_\omega \in \partial\, \mathrm{d}^n_{[a,b]}(u_{n,j,L}) \text{ and } b^*_{n,j,L} \in (\mathbb{B}_{p_j})^\circ, \tag{4.3}$$

such that:

1) $x^*_{\alpha,\omega} \in \hat{\partial}_{\mathrm{i}_\alpha,L} f(x_{\alpha,\omega})$,

2) $u^*_{n,j,L} \in x^* - (v^*_{n,j,L} + \frac{1}{n}b^*_{n,j,L}) + L^\perp$,

3) $x_{\alpha,\omega} \xrightarrow{f} u_{n,j,L}$,

4) $x^*_{\alpha,\omega} \xrightarrow{w^*} u^*_{n,j,L}$,

5) $\liminf_\alpha \langle x^*_{\alpha,\omega}, u_{n,j,L} - x_{\alpha,\omega}\rangle \geq 0$.

Observe that, due to 4) and 5),

$$\begin{aligned} \liminf_\alpha \langle c - x_{\alpha,\omega}, x^*_{\alpha,\omega}\rangle &\geq \liminf_\alpha \langle u_{n,j,L} - x_{\alpha,\omega}, x^*_{\alpha,\omega}\rangle + \liminf_\alpha \langle c - u_{n,j,L}, x^*_{\alpha,\omega}\rangle \\ &\geq \liminf_\alpha \langle c - u_{n,j,L}, x^*_{\alpha,\omega}\rangle = \langle c - u_{n,j,L}, u^*_{n,j,L}\rangle. \end{aligned} \tag{4.4}$$

Let $y_{n,j,L} \in [a,b]$ be such that $\mathrm{d}^n_{[a,b]}(u_{n,j,L}) = \rho_n(u_{n,j,L} - y_{n,j,L})$; because $u_{n,j,L} \to_{n,j,L} c$ and $c \neq b$, we may suppose w.l.o.g. that $y_{n,j,L} \neq b$ for all $(n,j,L) \in \mathbb{D}$, and so we write $y_{n,j,L} = (1 - \lambda_{n,j,L})a + \lambda_{n,j,L}b$ for some $\lambda_{n,\nu,L} \in [0,1[$. Moreover, by Lemma 4.5 we have that

$$\partial\, \mathrm{d}^n_{[a,b]}(u_{n,j,L}) \subseteq \partial\, \rho_n(u_{n,j,L} - y_{n,j,L}) \cap N_{[a,b]}(y_{n,j,L}) \subseteq \mathbb{B}^\circ_{\rho_n} \cap N_{[a,b]}(y_{n,j,L}). \tag{4.5}$$

Then due to (4.3), $v^*_{n,j,L}$ satisfies

$$(\lambda - \lambda_{n,\nu,L})\langle b - a, v^*_{n,j,L}\rangle = \langle (1-\lambda)a + \lambda b - y_{n,j,L}, v^*_{n,j,L}\rangle \le 0 \text{ for all } \lambda \in [0,1],$$

which yields

$$\langle b - a, v^*_{n,j,L}\rangle \le 0; \tag{4.6}$$

moreover, when $y_{n,j,L} \neq a$, we have $\lambda_{n,\nu,L} \in]0,1[$ and the last inequality implies that

$$\langle b - a, v^*_{n,j,L}\rangle = 0. \tag{4.7}$$

In particular, since $c \in [a, b[$ we also have (recall the first inclusion in (4.5))

$$\begin{aligned}\langle c - u_{n,j,L}, v^*_{n,j,L}\rangle &= \langle c - y_{n,j,L}, v^*_{n,j,L}\rangle - \langle u_{n,j,L} - y_{n,j,L}, v^*_{n,j,L}\rangle \\ &\le \langle y_{n,j,L} - u_{n,j,L}, v^*_{n,j,L}\rangle \le -\rho_n(u_{n,j,L} - y_{n,j,L}) \le 0. \end{aligned}\tag{4.8}$$

We may assume that $a, b \in L$ for all $L \in \mathcal{L}$. Then, on the one hand, using 2) and (4.6) we get

$$\begin{aligned}\langle b - a, u^*_{n,j,L}\rangle &= \langle b - a, x^*\rangle - \langle b - a, v^*_{n,j,L}\rangle - \frac{1}{n}\langle b - a, b^*_{n,j,L}\rangle \\ &\ge \langle b - a, x^*\rangle - \frac{1}{n}p_j(b-a) = r - f(a) - \frac{1}{n}p_j(b-a),\end{aligned}\tag{4.9}$$

and so, because $\lim \frac{1}{n}p_j(b-a) = 0$ (w.l.o.g.), we deduce that that

$$\liminf_{n,j,L}\langle b - a, u^*_{n,j,L}\rangle \ge r - f(a). \tag{4.10}$$

On the other hand, by using again 2) and the fact that $a, b, u_{n,j,L} \in L$, together with (4.8), we get

$$\begin{aligned}\langle c - u_{n,j,L}, u^*_{n,j,L}\rangle &= \langle c - u_{n,j,L}, x^*\rangle - \langle c - u_{n,j,L}, v^*_{n,j,L}\rangle - \frac{1}{n}\langle c - u_{n,j,L}, b^*_{n,j,L}\rangle \\ &\ge \langle c - u_{n,j,L}, x^*\rangle - \frac{1}{n}p_j(c - u_{n,j,L}) \ge \langle c - u_{n,j,L}, x^*\rangle - \frac{1}{n^2},\end{aligned}$$

which gives us

$$\liminf_{n,j,L}\langle c - u_{n,j,L}, u^*_{n,j,L}\rangle \ge 0. \tag{4.11}$$

Also, from the inequality $H_n(u_{n,j,L}) \le H_n(c)$, we get $f(u_{n,j,L}) \le f(c) - \langle c - u_{n,j,L}, x^*\rangle$, which implies that

$$f(c) \le \liminf_{n,j,L} f(u_{n,j,L}) \le \limsup_{n,j,L} f(u_{n,j,L}) \le f(c),$$

and this together with (4.2) prove that $u_{n,j,L} \xrightarrow{f} c$.

Finally, we consider the set $\mathbb{A} = \mathbb{D} \times \Pi_{(n,j,L)\in\mathbb{D}}\Lambda(n,j,L)$ that we order by the relation

$$(\nu_1, T_1) \preceq (\nu_2, T_2) \text{ if and only if } \nu_1 \le \nu_2 \text{ and } T_1(q) \le T_1(q) \text{ for all } q \in D,$$

and define the net $\left(x_{q,T}, x^*_{q,T}\right)_{(q,T)\in\mathbb{A}}$ as $x_{q,T} := x_{T(q)}$ and $x^*_{q,T} := x^*_{T(q)}$. We are going to check that the previous net satisfies the conclusions of the theorem. By taking into account Lemma 4.3, from the paragraph above the following statements hold true:

i) $x^*_{q,T} \in \hat{\partial}_{\mathrm{i}_{T(q)},L} f(x_{q,T})$, due to 1),

ii) $x_{q,T} \xrightarrow{f}_{q,T} c$, thanks to (4.2) and 3),

iii) $\liminf_{q,T} \langle b - a, x^*_{q,T} \rangle \geq r - f(a)$, by (4.10) and 4) (see, also, Remark 4.4),

iv) $\liminf_{q,T} \langle c - x_{q,T}, x^*_{q,T} \rangle \geq 0$, by (4.4) and (4.11) (see Remark 4.4).

Moreover, since that $b - c = (1 - \bar{\lambda})(b - a)$, by using the last statements (iii)) and (iv)) we obtain, for every $p \geq p_0$,

$$\begin{aligned}
\liminf_{q,T} \langle b - x_{q,T}, x^*_{q,T} \rangle &= \liminf_{q,T} \left((1 - \bar{\lambda}) \langle b - a, x^*_{q,T} \rangle + \langle c - x_{q,T}, x^*_{q,T} \rangle \right) \\
&\geq (1 - \bar{\lambda}) \liminf_{q,T} \langle b - a, x^*_{q,T} \rangle + \liminf_{q,T} \langle c - x_{q,T}, x^*_{q,T} \rangle \\
&\geq (1 - \bar{\lambda})(r - f(a)) = \frac{p(b-c)}{p(b-a)} (r - f(a)).
\end{aligned}$$

The main statement of the theorem is proved. If $c \neq a$, then $y_{n,j,L} \neq a$ and by (4.7) we obtain that $\langle b - a, v^*_{n,j,L} \rangle = 0$, which by (4.9) gives us

$$\langle b - a, u^*_{n,j,L} \rangle \to r - f(a).$$

Then, using again Lemma 4.3 we also may suppose that $\lim_{q,T} \langle b - a, x^*_{q,T} \rangle = 0$. □

Remark 4.10 It is easy to see that, when the space X is metrizable and the sets $\mathcal{L}$ and I have countable cofinal sets, we can take sequences instead of nets. Indeed assume that there are nets (x_α, x^*_α) which satisfy the properties of the theorem above and let $(L_n)_{n\in\mathbb{N}}$ and $(\mathrm{i}_n)_{n\in\mathbb{N}}$ be countable cofinal sets in $\mathcal{L}$ and $\mathbb{I}$ respectively. Then for every $n \in \mathbb{N}$, there exists $\alpha \in \mathbb{D}$ and $x^*_\alpha \in \hat{\partial}_{L_\alpha,\mathrm{i}_\alpha} f(x_\alpha)$ such that $L_n \subseteq L_\alpha$, $\mathrm{i}_n \leq \mathrm{i}_\alpha$, $\mathrm{d}(c, x_\alpha) \leq 1/n$, $|f(x_\alpha) - f(c)| \leq 1/n$, $r - f(a) \leq \langle b - a, x^*_\alpha \rangle + 1/n$, $0 \leq \langle c - x_\alpha, x^*_\alpha \rangle + 1/n$, $\frac{p(b-c)}{p(b-a)} (r - f(a)) \leq \langle b - x_\alpha, x^*_\alpha \rangle + 1/n$, which proves the assertion.

The next corollary is the density of the domain of the subdifferential. This class of theorems are also known as Like-Brøndsted-Rockafellar Theorems. The generalization for the convex case of an epi-pointed lsc function has been proved in [30].

Corollary 4.11 *In the setting of Theorem 4.9 if $x \in \operatorname{dom} f$, then there exists $(x_\beta, x^*_\beta)_{\beta\in\mathbb{U}} \subseteq \operatorname{gph} \hat{\partial}_{L_\beta,\mathrm{i}_\beta} f$ such that $x_\beta \xrightarrow{f} x$ and $\liminf \langle x - x_\beta, x^*_\beta \rangle \geq 0$. In particular,*

$$\operatorname{dom} f \subseteq \limsup_{L,\mathrm{i}} \operatorname{dom} \hat{\partial}_{L,\mathrm{i}} f.$$

Proof. Let $x \in \operatorname{dom} f$. First assume that x is a local minimum of f, then for every $L \in \mathcal{L}$, x is a local minimum of f relative to L, so there exists a net $(x_\alpha, x^*_\alpha) \in \operatorname{gph} \partial_{L,\mathrm{i}_\alpha}$ with $\alpha \in \mathbb{D}(L)$ and $x^*_L \in L^\perp$ such that $x_\alpha \xrightarrow{f} x$ and $x^*_\alpha \xrightarrow{w^*} x^*_L$ and $\liminf \langle x - x_\alpha, x^*_\alpha \rangle \geq 0$. Then we define the set $\mathbb{U} = \mathcal{L} \times \Pi_{L\in\mathcal{L}} \mathbb{D}(L)$ and the nets $y_{L,T} = x_{T(L)}$ and $y^*_{L,T} = x^*_{T(L)}$ and applying Proposition 4.3 we get the conclusion. Now assume that x is not a local minimum, because

f is lsc we can find a net $(x_\nu)_{\nu\in\mathcal{N}} \xrightarrow{f} x$ with $f(x_\nu) < f(x)$. Then Theorem 4.9 applied to x_ν, x and $r = f(x)$ and ρ_ν, with a seminorm ρ_ν such that $\rho_\nu(x_\nu - x) \neq 0$, yields that for every ν there exists a net $(z_\alpha, z_\alpha^*)_{\alpha\in\mathbb{D}(\nu)}$ such that $z_\alpha^* \in \hat{\partial}_{L_\alpha,\mathrm{i}_\alpha} f(z_\alpha)$, $z_\alpha \xrightarrow{f} c_\nu \in [x_\nu, x[$, $0 < \frac{\rho_\nu(x-c_\nu)}{\rho_\nu(x-x_\nu)}(f(x) - f(x_\nu)) \leq \liminf\langle x - z_\alpha, z_\alpha^*\rangle$ and $f(c_\nu) \leq f(x_\nu) + \frac{\rho_\nu(c_\nu-x_\nu)}{\rho_\nu(x-x_\nu)}(f(x) - f(x_\nu)) \leq f(x)$ (because $\frac{\rho_\nu(c_\nu-x_\nu)}{\rho_\nu(x-x_\nu)} \leq 1$). So $\limsup f(c_\nu) \leq f(x)$ and by lsc of f and the fact that for every seminorm p we have $p(x - c_\nu) \leq p(x - x_\nu) \to 0$, we conclude $c_\nu \xrightarrow{f} x$. Finally taking the order set $\mathbb{U} = \mathcal{N} \times \Pi_{\nu\in\mathcal{N}}\mathbb{D}(\nu)$ and defining $w_{j,T} = z_{T(j)}$ and $w_{j,T}^* = z_{T(j)}^*$ and using Proposition 4.3 we get the conclusion. $\square$

The following statement is a classical result about the continuity and directional differentiability of convex functions over the real line (see e.g. [105, Theorems 10.1 and Theorem 24.1]).

Proposition 4.12 *Let $h \in \Gamma_0(\mathbb{R})$ be such that $\operatorname{int}(\operatorname{dom} h) \neq \emptyset$, then the following assertions hold*

i) h is continuous relative to $\operatorname{dom} h$; moreover, h is Lipschitz on every compact interval included in $\operatorname{int}(\operatorname{dom} h)$ and for all $t \in \operatorname{dom} h$.

$$h'_+(t) := \lim_{t'\downarrow t} \frac{h(t') - h(t)}{t - t} = \inf_{t'>t} \frac{h(t') - h(t)}{t' - t} \in \overline{\mathbb{R}},$$
$$h'_-(t) := \lim_{t'\uparrow t} \frac{f(t') - f(t)}{t - t} = \sup_{t'<t} \frac{h(t) - h(t)}{t' - t} \in \overline{\mathbb{R}},$$

and $h'_-(t) \leq h'_+(t)$.

ii) Let $t_1, t_2 \in \operatorname{dom} h$, $t_1 < t_2$, then $h'_+(t_1) \leq h'_-(t_2)$. Therefore h'_+ and h'_- are non-decreasing on $\operatorname{dom} h$ and for every $t \in \operatorname{dom} h$

$$h'_+(t) = \lim_{t'\downarrow t} h'_+(t') = \lim_{t\downarrow t_0} h'_-(t'), \text{ and } h'_-(t) = \lim_{t'\uparrow t} h'_+(t') = \lim_{t\uparrow t_0} h'_-(t').$$

The following (classical) result is due to Thibault-Zagrodny in the setting of Banach spaces [117, Theorem 2.1]. More general approaches, which consider a nonconvex function g can be found in [118, Theorem 3.2]. We adapt the proof of the classical statement ([117, Theorem 2.1]) to our framework.

Theorem 4.13 *In the setting of Theorem 4.9, consider a function $g \in \Gamma(X)$ and a continuous seminorm ρ. Suppose there exist $\varepsilon \geq 0$, a net $(\varepsilon_\mathrm{i})_{\mathrm{i}\in I} \downarrow \varepsilon$ and an open convex set U with $U \cap \operatorname{dom} f \neq \emptyset$ such that*

$$\hat{\partial}_{L,\mathrm{i}} f(x) \subseteq \partial g(x) + \varepsilon_\mathrm{i}\mathbb{B}_\rho(0,1)^\circ + L^\perp, \quad \forall x \in C, \ \forall \mathrm{i} \in \mathbb{I}, \forall L \in \mathcal{L}.$$

Then $C \cap \operatorname{dom} f = C \cap \operatorname{dom} g$ and for all $x \in C$, $y \in C \cap \operatorname{dom} f$ one has

$$g(x) - g(y) - \varepsilon\rho(x - y) \leq f(x) - f(y) \leq g(x) - g(y) + \varepsilon\rho(x - y). \tag{4.12}$$

First we prove the following two lemmas.

Lemma 4.14 *Let $y \in C \cap \operatorname{dom} f \cap \operatorname{dom} g$ and $x \in C \cap \operatorname{dom} g$. Consider $\gamma > \varepsilon$ and a continuous seminorm $\bar{\rho}$ such that $\bar{\rho}(x-y) \neq 0$ and $\bar{\rho} \geq p$. Consider $u := \frac{x-y}{\bar{\rho}(x-y)}$ and $h : \mathbb{R} \to \overline{\mathbb{R}}$, given by $h(t) := g(y+tu)$ and assume that $g'(y,u) \in \mathbb{R}$. Then $h \in \Gamma_0(\mathbb{R})$, h is continuous on $\operatorname{dom} h \supseteq [0, \bar{\rho}(x-y)]$, $h'_+(t) = g'(y+tu, u)$ for all $t \in \operatorname{dom} h$ and there exists $x_0 = y + t_0 u$ with $t_0 \in (0, p(x-y))$ such that*

$$f(x_0) - f(y) \leq g(x_0) - g(y) + \gamma\bar{\rho}(x_0 - y).$$

Proof. Clearly $h \in \Gamma(\mathbb{R})$ and $[0, \bar{\rho}(x-y)] \subseteq \operatorname{dom} h$, so by Proposition 4.12 h is continuous on $\operatorname{dom} h$, $h'_+(t) = g'(y+tu, u)$ for all $t \in \operatorname{dom} h$ and

$$\lim_{t \downarrow 0} g'(y+tu, u) = \lim_{t \downarrow 0} h'_+(t) = h'_+(0) = g'(y,u),$$

by our assumption $g'(y,u)$ is finite, therefore there exists $t_0 \in (0, \bar{\rho}(x-y))$ such that $g'(y + t_0 u, u) \leq g'(y,u) + \frac{1}{2}(\gamma - \varepsilon)$. Define $x_0 := y + t_0 u \in]y, x[$, hence

$$g'(x_0, u) \leq \frac{g(y+t_0 u) - g(y)}{t_0} + \frac{1}{2}(\gamma - \varepsilon) = \frac{g(x_0) - g(y)}{\bar{\rho}(x_0 - y)} + \frac{1}{2}(\gamma - \varepsilon).$$

We claim that $f(x_0) - f(y) \leq g(x_0) - g(y) + \gamma\bar{\rho}(x_0 - y)$. Indeed suppose it does not hold, which means $f(x_0) - f(y) > g(x_0) - g(y) + \gamma\bar{\rho}(x_0 - y)$. Choose $r \in \mathbb{R}$ such that $r \leq f(x_0)$ and $r - f(y) > g(x_0) - g(y) + \gamma\bar{\rho}(x_0 - y)$. By Theorem 4.9 there exists $x_\alpha \to z \in [y, x_0[$ and $x^*_\alpha \in \hat{\partial}_{\mathrm{i}_\alpha, L_\alpha} f(x_\alpha)$ such that $x_\alpha \in L_\alpha$ and $\bar{\rho}(x_0 - z)(r - f(y)) \leq \bar{\rho}(x_0 - y) \liminf \langle x_0 - x_\alpha, x^*_\alpha \rangle$. Without loss of generality we may assume that $x_0 \in L_\alpha$ for all α. Since $\bar{\rho}(x_0 - x_\alpha) \to \bar{\rho}(x_0 - z)$ we have

$$\frac{r - f(y)}{\bar{\rho}(x_0 - y)} \leq \liminf \left\langle \frac{x_0 - x_\alpha}{\bar{\rho}(x_0 - x_\alpha)}, x^*_\alpha \right\rangle.$$

Then there exists α_0 such that for all $\alpha \geq \alpha_0$

$$\left\langle \frac{x_0 - x_\alpha}{\bar{\rho}(x_0 - x_\alpha)}, x^*_\alpha \right\rangle > \frac{g(x_0) - g(y)}{\bar{\rho}(x_0 - y)} + \gamma.$$

Since $x_\alpha \to z \in [y, x_0[\subseteq C$, we may assume that $x_\alpha \in C$ for all $\alpha \geq \alpha_0$. So by our assumption we can assume that $x^*_\alpha = z^*_\alpha + \varepsilon_{\mathrm{i}_\alpha} u^*_\alpha + \lambda^*_\alpha$ with $z^*_\alpha \in \partial\, g(x_\alpha)$, $u^*_\alpha \in \mathbb{B}_\rho(0,1)^\circ$ and $\lambda^*_\alpha \in {L_\alpha}^\perp$ for all $\alpha \geq \alpha_0$. Hence

$$\frac{g(x_0) - g(x_\alpha)}{\bar{\rho}(x_0 - x_\alpha)} \geq \langle \frac{x_0 - x_\alpha}{\bar{\rho}(x_0 - x_\alpha)}, z^*_\alpha \rangle > \frac{g(x_0) - g(y)}{\bar{\rho}(x_0 - y)} + \gamma - \varepsilon_{\mathrm{i}_\alpha} \quad \forall \alpha \geq \alpha_0,$$

so taking into account the lower semi-continuity of g we get

$$\frac{g(x_0) - g(z)}{\bar{\rho}(x_0 - z)} \geq \frac{g(x_0) - g(y)}{\bar{\rho}(x_0 - y)} + \gamma - \varepsilon \geq g'(x_0, u) + \frac{1}{2}(\gamma - \varepsilon) > g'(x_0, u). \tag{4.13}$$

Since $z \in [y, x_0[$, $z = y + su$ for some $s \in [0, t_0[$ and using the convexity of h we get the following inequality

$$\frac{g(x_0) - g(z)}{\bar{\rho}(x_0 - z)} = \frac{h(t_0) - h(s)}{t_0 - s} \leq h'_-(t_0) \leq h'_+(t_0) = g'(x_0, u),$$

which contradicts (4.13). Therefor, necessarily we have that $f(x_0) - f(y) \leq g(x_0) - g(y) + \gamma\bar{\rho}(x_0 - y)$. □

Lemma 4.15 *Let $y \in C \cap \operatorname{dom} f \cap \operatorname{dom} \partial g$ and $x \in C \cap \operatorname{dom} g$. Consider $\gamma > \varepsilon$ and a continuous seminorm $\bar{\rho}$ such that $\bar{\rho}(x-y) \neq 0$ and $\bar{\rho} \geq p$. Then $f(x) - f(y) \leq g(x) - g(y) + \gamma p(x-y)$.*

Proof. Consider $u := \frac{x-y}{\bar{\rho}(x-y)}$, since $\partial g(y) \neq \emptyset$, we have that $\frac{g(x)-g(y)}{\bar{\rho}(x-y)} \geq g'(y,u) > -\infty$, which means $g'(y,u) \in \mathbb{R}$. Then by Lemma 4.14 there exist $x_0 = y + t_0 u$ with $t_0 \in (0, \bar{\rho}(x-y))$ such that

$$f(x_0) - f(y) \leq g(x_0) - g(y) + \gamma \bar{\rho}(x_0 - y).$$

Now define $T = \{t \in (0, \bar{\rho}(x-y)] : f(y+tu) - f(y) \leq g(y+tu) - g(y) + \gamma t\}$ and consider $\bar{t} := \sup T$, because f is lsc and h (the same definition as in Lemma 4.14) is continuous on $[0, \bar{\rho}(x-y)]$ (see Lemma 4.14), we obtain $\bar{t} \in T$. Moreover suppose that $\bar{t} < \bar{\rho}(x-y)$, define $\bar{y} = y + \bar{t}u$ and noticing that $u = \frac{x-\bar{y}}{\bar{\rho}(x-\bar{y})}$ and $\bar{y} \in C \cap \operatorname{dom} f \cap \operatorname{dom} g$, besides

$$\frac{h(\bar{\rho}(x-y)) - h(\bar{t})}{\bar{\rho}(x-y) - \bar{t}} = \frac{g(x) - g(\bar{y})}{\bar{\rho}(x-\bar{y})} \geq g'(\bar{y}, u) = h'_+(\bar{t}) \geq h'_+(0) = g'(y,u) > -\infty.$$

Hence $g'(\bar{y}, u) \in \mathbb{R}$, again using Lemma 4.14 there exists $x_0 = \bar{y} + s_0 u$ with $s_0 \in (0, \bar{\rho}(x-\bar{y}))$ such that

$$f(\bar{y} + s_0 u) - f(\bar{y}) \leq g(\bar{y} + s_0 u) - g(\bar{y}) + \gamma \bar{\rho}(x_0 - \bar{y}) = g(\bar{y} + s_0 u) - g(\bar{y}) + \gamma s_0,$$

then adding $f(\bar{y}) - f(y) \leq g(\bar{y}) - g(y) + \bar{t}\gamma$ we get

$$f(y + (\bar{t} + s_0)u) - f(y) \leq g(y + (\bar{t} + s_0)u) - g(y) + (\bar{t} + s_0)\gamma,$$

which contradicts the choice of $\bar{t}$. Hence $\bar{t} = \bar{\rho}(x-y)$. □

Proof of Theorem 4.13 First assume that $y \in C \cap \operatorname{dom} f \cap \operatorname{dom} \partial g$, $x \in C \cap \operatorname{dom} g$ and $x \neq y$ and choose a seminorm ρ_0 such that $\rho_0(x-y) \neq 0$, define $\bar{\rho}_n := \rho + n^{-1}\rho_0$, then by Lemma 4.15, $f(x) - f(y) \leq g(x) - g(y) + \gamma \bar{\rho}_n(x-y)$ for all $\gamma > \varepsilon$, and $n \geq 1$, therefore

$$f(x) - f(y) \leq g(x) - g(y) + \varepsilon p(x-y),\ \forall x \in C \cap \operatorname{dom} g,\ \forall y \in C \cap \operatorname{dom} f \cap \operatorname{dom} \partial g. \tag{4.14}$$

Now, taking $x \in C \cap \operatorname{dom} g$ and (if there exists) $y \in C \cap \operatorname{dom} f$ such that $x \neq y$, by Corollary 4.11 there exists $(y_\alpha, y_\alpha^*) \subseteq \operatorname{gph} \hat{\partial}_{\mathrm{i}_\alpha, L_\alpha} f$ such that $y_\alpha \xrightarrow{f} y$. Because $y \in C$, C is open and $y_\alpha \to y$, we can suppose that $y_\alpha \in C$ and $y_\alpha \neq x$ for all α; moreover from our assumptions $x_\alpha \in \operatorname{dom} f \cap \operatorname{dom} \partial g$, then by Equation (6.7)

$$f(x) - f(x_\alpha) \leq g(x) - g(x_\alpha) + \varepsilon p(x - x_\alpha).$$

From the lsc of g we conclude

$$f(x) - f(y) \leq g(x) - g(y) + \varepsilon p(x-y),\ \forall x \in C \cap \operatorname{dom} g,\ \forall y \in C \cap \operatorname{dom} f. \tag{4.15}$$

So on one hand, using Equation (6.8) and the fact that $C \cap \operatorname{dom} f \neq \emptyset$, we easily get $C \cap \operatorname{dom} g \subseteq C \cap \operatorname{dom} f$. On the other hand, take $y \in \operatorname{dom} f \cap C$, then by Corollary 4.11

there exists $(y_\alpha, y_\alpha^*) \subseteq \operatorname{gph} \hat{\partial}_{\mathrm{i}_\alpha, L_\alpha} f$ such that $y_\alpha \xrightarrow{f} y$, so by our hypothesis there exists some α_0 such that $y_\alpha \in C \cap \operatorname{dom} \partial g \subseteq C \cap \operatorname{dom} g$. Then if $y_\alpha = y$ for all $\alpha \geq \alpha_0$, we get $y \in C \cap \operatorname{dom} g$, otherwise we can assume that $y_\alpha \neq y$ for some y_α, then (6.8) implies $y \in C \cap \operatorname{dom} g$. It proves that $\operatorname{dom} f \cap C = \operatorname{dom} g \cap C$. Therefore, interchanging the roles of x and y in (6.8), we conclude (4.12). $\square$

Corollary 4.16 *In the setting of Theorem 4.13, assume that $\varepsilon = 0$ and $C = X$. Then there exists $c \in \mathbb{R}$ such that $f = g + c$.*

Proof. By Theorem 4.13 we have $\operatorname{dom} g = \operatorname{dom} f$ and $f(x) - f(y) = g(x) - g(y)$ for all x, y. Therefore fixing $x \in \operatorname{dom} f$, we have $f(y) = g(y) + f(x) - g(x)$, for all $y \in X$. $\square$

To show the potential application of Theorem 4.13 we give a short proof of two results available in the literature referent to integration theorems. It worth mentioning that the original proof of the two result does not follow the argument presented in this paper and even they cannot be covered by the original statement of [117, Theorem 2.1], therefore Theorem 4.13 allow us to include three results, which appeared to be independent, in a same framework.

The following corollary is due to Thibault [81, Theorem 1.2] (see also [25, Corollary 10.] for a nonconvex approach). In this paper the authors use the ε-subdifferential of a proper lsc convex function f to establish conditions to recover a function from its subdifferential in an arbitrary locally convex space. More precisely we have the following statement.

Corollary 4.17 *Let X be an lcs and let $f, g : X \to \mathbb{R} \cup \{+\infty\}$ be two proper lsc convex functions such that for every $x \in X$ there exists $\delta(x) > 0$ such that $\partial_\varepsilon f(x) \subseteq \partial_\varepsilon g(x)$ for all $\varepsilon < \delta(x)$. Then there exists $c \in \mathbb{R}$ such that $f = g + c$.*

Proof. Fix $x \in \operatorname{dom} f$ and consider $y \in \operatorname{dom} g$ different from x, consider L as the subspace generated by x, y and let U be a closed neighborhood of zero such that $x, y \in \operatorname{int}(U)$ and $B = L \cap U$ is compact, then the functions $\tilde{f} = f + \delta_B$ and $\tilde{g} = g + \delta_B$ are w-lsc and epi-pointed, since for every $w \in X$, $\partial \tilde{f}(w) = \bigcap_{\varepsilon>0} \overline{\partial_\varepsilon f(w) + \partial_\varepsilon \delta_B(w)}^{w^*}$ and $\partial \tilde{g}(w) = \bigcap_{\varepsilon>0} \overline{\partial_\varepsilon g(w) + \partial_\varepsilon \delta_B(w)}^{w^*}$ (see e.g. [57, Theorem 2.1]). Therefore for every $w \in X$

$$
\begin{aligned}
\partial \tilde{f}(w) &= \bigcap_{\varepsilon>0} \overline{\partial_\varepsilon f(w) + \partial_\varepsilon \delta_B(w)}^{w^*} \\
&\subseteq \overline{\partial_\varepsilon g(w) + \partial_\varepsilon \delta_B(w)}^{w^*} \\
&= \partial \tilde{g}(w).
\end{aligned}
$$

Then by Corollary 4.16 we get $x \in \operatorname{dom} \tilde{g} \subseteq \operatorname{dom} g$, $y \in \operatorname{dom} \tilde{f} \subseteq \operatorname{dom} f$ and $f(x) - f(y) = g(x) - g(y)$. $\square$

Remark 4.18 It is worth mentioning that one can apply the result above with a nonconvex function g passing to the closed convex hull or to the biconjugate function. Indeed if one have some inclusion of the form $\partial f(x) \subseteq \partial g(x)$ then necessarily $\partial f(x) \subseteq \partial g^{**}(x)$ and $\partial f(x) \subseteq \partial \overline{\operatorname{co}}\, g(x)$, because when $\partial g(x)$ is nonempty $g(x) = g^{**}(x) = \overline{\operatorname{co}}\, g(x)$ and $\partial g(x) = \partial g^{**}(x) = \partial \overline{\operatorname{co}}\, g(x)$.

The following corollary is a particular case of a more general result [33, Theorem 4.8.] (see also [25] for another approach). For the sake of simplicity in our statement we avoid the introduction of more sophisticate notions of convex and non-smooth analysis.

Corollary 4.19 *Let $f : X \to \overline{\mathbb{R}}$ be a w-lsc and epi-pointed function and let $g : X \to \mathbb{R}$ be an arbitrary function such that*

$$\partial f(x) \subseteq \partial g(x) \ \ \forall x \in X. \tag{4.16}$$

Then there exists a constant $c \in \mathbb{R}$ such that $f^{**} = g^{**} \Box \sigma_{\operatorname{dom} f^*} + c$.

Proof. Consider $x^* \in \partial f^{**}(x)$, then $x \in \partial f^*(x^*)$, by Proposition 1.14 we have that

$$\partial f^*(x^*) = \overline{\text{co}}^{w^*} \left(\partial f\right)^{-1}(x^*) + N_{\operatorname{dom} f^*}(x^*),$$

which means there exist $y, \lambda \in X$, and nets $y_{\rm i} = \sum_{k=1}^{n_{\rm i}} \alpha_{\rm i}^k w_{\rm i}^k$ such that $x = y + \lambda$, $y = \lim y_{\rm i}$, $\alpha_{\rm i} \geq 0$, $\sum_{k=1}^{n_{\rm i}} \alpha_{\rm i}^k = 1$ and $x^* \in \partial f(w_{\rm i}^k)$, by the inclusion of subdifferential yields $x^* \in \partial g(w_{\rm i}^k)$, so $w_{\rm i}^k \in \partial g^*(x^*)$, then using the fact that $\partial g^*(x^*)$ is convex and closed, we get $y \in \partial g^*(x^*)$, consequently $x^* \in \partial g^*(x^*) + N_{\operatorname{dom} f^*}(x^*) \subseteq \partial(g^* + \delta_{\overline{\operatorname{dom} f^*}})(x^*)$, so $x^* \in \partial(g^* + \delta_{\overline{\operatorname{dom} f^*}})^*(x)$. Therefore by Corollary 4.16 there exists a constant $c \in \mathbb{R}$ such that $f^{**} = (g^* + \delta_{\overline{\operatorname{dom} f^*}})^* + c$. Finally, taking $x^* \in \operatorname{int}(\operatorname{dom} f^*)$ and by Lemma 4.6, there exists $\bar{x} \in X$ such that $x^* \in \partial f(\bar{x})$. Then by the inclusion of subdifferential and the Fenchel duality, $\bar{x} \in \partial g^*(x^*)$, in particular $x^* \in \operatorname{dom} g^*$, which allow us to apply [31, Lemma 3.d] and conclude $f^{**} = g^{**} \Box \sigma_{\operatorname{dom} f^*}$. $\square$

The following theorem is due to Correa-Jofre-Thibault [35, Theorem 2.2 and 2.4] for an lsc function definite in a Banach space X. We extend this result for weakly lsc epi-pointed functions definite in an arbitrary locally convex space.

Theorem 4.20 *In the setting of Theorem 4.9, assume that ∂ is a subdifferential and consider the following statement:*

(i) f is convex.

(ii) ∂f is monotone.

(iii) $\hat{\partial} f(x) \subseteq \partial f(x)$, for all $x \in X$.

Then (ii) *and* (iii) *are equivalent and each one implies* (i). *In addition, if ∂f coincide with the convex subdifferential for every convex function f, then the three statements are equivalent.*

Proof. (ii) $\Rightarrow$ (iii), first, if $\operatorname{dom} f = \emptyset$, then $f = +\infty$, so the theorem is true, if $\operatorname{dom} f \neq \emptyset$, then take $a \in \operatorname{dom} f$, $x \in \operatorname{dom} \partial f$, a seminorm p such that $p(a - x) \neq 0$, $x^* \in \partial f(x)$ and $r \in \mathbb{R}$ with $r \leq f(x)$. Then by Theorem 4.9 there exists a net $(x_\alpha, x_\alpha^*) \in \operatorname{gph} \partial f$ and $c \in [a, b[$

such that $x_\alpha \to c$ and

$$\begin{aligned}
r - f(a) &\le \frac{p(x-a)}{p(x-c)} \liminf \langle x_\alpha^*, x - x_\alpha\rangle \\
&= \frac{p(x-a)}{p(x-c)} \liminf \left\{ \langle x_\alpha^* - x^*, x - x_\alpha\rangle + \langle x^*, x - x_\alpha\rangle \right\} \\
\text{(by the monotonicity)} \quad &\le \frac{p(x-a)}{p(x-c)} \liminf \langle x^*, x - x_\alpha\rangle = \frac{p(x-a)}{p(x-c)} \langle x^*, x - c\rangle = \langle x^*, x - a\rangle.
\end{aligned}$$

Therefore $x \in \operatorname{dom} f$ and $x^* \in \partial f(x)$. Now (iii) $\Rightarrow$ (ii) follows from the monotonicity of the ∂f. Finally (iii) $\Rightarrow$ (i), by (iii) we get $\partial f(x) \subseteq \partial f^{**}(x)$, for all $x \in X$, then by Corollary 4.16 we get $f = f^{**}$. The final implication is only a matter of using the extra hypothesis. $\square$

Chapter 5

A complete characterization of the subdifferential of convex integral functions

5.1 Introduction

Several problems in applied mathematics such calculus of variation, control theory and stochastic programing among others concern the study of integral functionals (see Definition 1.19) of the form

$$x(\cdot) \in \mathfrak{X} \to \hat{I}_f(x(\cdot)) := \int_T f(t, x(t)) \mathrm{d}\mu(t).$$

Optimization problems which include integral functions and functionals offer rich and challenging territory for variational analysis. Indeed, it is around such problems that the theory of normal integrands has traditionally been used, namely, in calculus of variations, where the underlying spaces $\mathfrak{X}$ are of Sobolev type. Models which consider integrals with respect to time are common in the study of dynamical systems and their optimal control. On the other hand, when the model is represented by random states, for instance density distributions, the problem is also formulated under the sign of integral. Applications to stochastic programing, often concern the study of expectation that are obviously defined by integration over probability spaces.

It is natural in the classical studies of calculus of variations, to consider integrands $f(t, x)$ which are continuous in t and x jointly, or indeed differentiable. Later, mathematicians considered models with integrands that are finite valued on $T \times \mathbb{R}^n$, and usually satisfying the Carathéodory condition; that is, f is assumed to be continuous in $x \in \mathbb{R}^n$ and measurable with respect to $t \in T$. In all of these cases one can easily notice that for every measurable function $x(\cdot)$, the function $t \to f(t, x(t))$ is at least measurable and, hence, (7) is well defined, using the convention adopted for the extended-real line. However,due to new mathematical models, specially the emergence of modern control theory, the integrands are forced to admit possibly infinite values, in order to include different important kinds of constraints that can efficiently be represented by these integrals. Such integrands require a distinctly new theoretical approach, where questions of measurability, meaning of the integral and the existence of *measurable selections* are prominent and are reflected in the concept of *normal integrands*.

One could hope that it might be enough simply to replace the continuity of $f(t,x)$ in x by lower semicontinuity, while maintaining the measurability of $f(t,x)$ in t. Nevertheless, it is not enough to ensure the measurability of $t \to f(t,x(t))$. Indeed, consider $T=[0,1]$ and $\mathcal{A}$ as the Lebesgue measurable sets in $[0,1]$, and consider a non-measurable set D in $[0,1]$. Then define the function $f(t,x)=0$ if $t=x\in D$ and $f(t,x)=1$ otherwise. The measurability and lower semicontinuity of the function f hold trivially in this case. However, considering $x(t)=t$, one can verify the lack of measurability of the function $t \to f(t,x(t))$. This example shows that although lower semicontinuity in x is certainly right, the assumption of measurability in t for each fixed x is not adequate.

The way out of this impasse was found by Rockafellar [99] in the concept of *normal convex integrands*, which were defined by Rockafellar [99] as (for an equivalent way see Section 1.4): *We shall call a convex integrand $f := f(t,x)$ normal if f is proper and lower semi-continuous in x for each t, and if, further, there exists a countable collection U of measurable functions u from T to $\mathbb{R}^n$ having the following properties:* ***(a)*** *for each $u \in U$, $f(t,u(t))$ is measurable in t;* ***(b)*** *for each t, $U_t \cap \operatorname{dom} f_t$ is dense in $\operatorname{dom} f_t$, where $U_t=\{u(t) : u \in U\}$.*

The notation of convex normal integrands was the link that allowed to connect the theories of Measurable multifunctions and subdifferentials. The preservation of the measurability of multifunctions under a broad variety of operations, including countable intersections, countable unions, sum of measurable multifunctions, and *Painlevé-Kuratowski limits*, made this theory very popular in various problems of applied mathematics during the last four decades.

Among many important properties in the theory of measurable multifunctions, one of the most useful, at least for the author's experience from the beginning to the end of this thesis, is *Castaing's representation* of a measurable multifunction. This mathematical result comes from Castaing's thesis [21] (see, also, Castaing-Valadier' book [22] for more details about this fascinating theory) and basically express the possibility of getting some measurable selection in every measurable multifunction (see Proposition 1.21).

Castaing's representation theorem is intrinsically related to the possibility of extending the definition of integration of a function to a set-valued mapping, considering their measurable selections which are also integrable. For our purpose, this corresponds to the possibility of representing formulas for the subdifferential of integral functionals, in a similar way as the *Leibniz integral rule.*

Some of the classical studies about this class of functionals can be found in Castaing-Valadier [22], Ioffe-Levin [66], Ioffe-Tikhomirov [67], Levin [79], Rockafellar [99, 101, 102]. Other recent works about this class of functionals are Borwein-Yao [12], Ioffe [64], Lopez-Thibault [80] among others. A summary of the elementary theory of measurability and integral functionals in finite dimension can be found in [103, 106], and for infinite dimensional spaces it can be found in [22, 58, 127, 128].

The aim of this Chapter is to bring formulas for the ε-subdifferential of the convex integral functional I_f; this particular case is also known as the continuous sum. A well-known formula, given by Ioffe-Levin [66], shows that under certain continuity assumptions, the following formula holds for the subdifferential of I_f:

$$\partial I_f(x) = \int_T \partial f(t,x)\mathrm{d}\mu(t) + N_{\mathrm{dom}_f}(x), \text{ for all } x \in \mathbb{R}^n, \tag{5.1}$$

where the set $\int_T \partial f(t,x)\mathrm{d}\mu(t)$ is understood in the sense of (1.6), that is to say, as the set of points of the form $\int_T x(t)\mathrm{d}\mu(t)$ where x is an integrable function such that $x(t) \in \partial f(t,\cdot)(x)$ for almost all $t \in T$ (ae for short). One can compare (5.1) with its discrete counterpart, which declares that for every two convex lsc functions f_1, f_2 such that f_1 is continuous at some point of the domain of f_2 one gets $\partial(f_1+f_2)(x) = \partial f_1(x) + \partial f_2(x)$ for all $x \in \mathbb{R}^n$. So, a reasonable idea is to give similar formulas as the ones given in the discrete case by Hiriart-Urruty and Phelps without qualification conditions (see Proposition 1.2). Hence, it feels natural to think in a generalization of (8) as

$$\partial I_f(x) = \bigcap_{\eta>0} \mathrm{cl}\left\{ \int_T \partial_\eta f(t,x)\mathrm{d}\mu(t) + N_{\mathrm{dom}_f}(x) \right\} \text{ for all } x \in \mathbb{R}^n. \tag{5.2}$$

However, such a mathematical expression does not hold without any qualification conditions; indeed, one can even find counterexamples where the set $\int_T \partial_\varepsilon f(t,x)\mathrm{d}\mu(t)$ is empty and the integrand $f_t(\cdot)$ is smooth at the point of interest (see Example 5.12).

The above example motivates us to use larger sets than $\int_T \partial_\varepsilon f(t,x)\mathrm{d}\mu(t)$ to generalize (8). With this in mind we provide general formulas for the ε-subdifferential of convex integral functionals defined in an arbitrary locally convex space (see 5.9). Later, we derive many corollaries and simplifications under classical qualification conditions. Finally, in Section 5.7 we provide a characterization of the subdifferential of an integral functional by means of limits of integrable selections of ∂f_t.

5.2 Preliminary results

The next definition corresponds to the notion of decomposability in locally convex Suslin spaces [22, Definition 3, Chapter VII].

Definition 5.1 (i) *Assume that* $(T,\mathcal{A}) = (\mathbb{N},\mathcal{P}(\mathbb{N}))$. *A vector space* $L \subset X^T$ *is said to be decomposable if*

$$c_{00}(X) := \{(x_n) : \exists k_0 \in \mathbb{N} \textit{ such that } x_k = 0, \forall k \geq k_0\} \subset L.$$

(ii) *Assume that* $(T,\mathcal{A}) \neq (\mathbb{N},\mathcal{P}(\mathbb{N}))$. *A vector space* L *of weakly integrable functions in* X^T *is said to be decomposable if for every* $u \in L$, *every weakly integrable function* $f \in X^T$ *such that* $f(T)$ *is relatively compact, and every finite-measure set* $A \in \mathcal{A}$, *we have that* $f\mathbb{1}_A + u\mathbb{1}_{A^c} \in L$.

The specification of the decomposability above to the underlying σ-Algebra $(T,\mathcal{A})$ makes sense, since both definitions may not coincide. For instance, if $X = \mathbb{R}$ and μ is a finite measure over $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$, then the space $L = c_{00}(X)$ is obviously decomposable in the sense of Definition 5.1(i), but not with respect to Definition 5.1(ii). Indeed, the decomposability of L in the sense of Definition 5.1(ii) would imply that $\ell^\infty \subseteq L$. Typical decomposable spaces in the literature are the spaces $\mathrm{L}^p(T, X)$.

We shall extensively use the following result, which characterizes the Fenchel conjugate of $\hat{I}_f$. The first part of it, corresponding to the case when (X, X^*) is a dual pair of Suslin spaces, can be found in [22, Theorem VII-7]. In the second part, we obtain a similar representation of the conjugate of $\hat{I}_f$ when $(T, \mathcal{A}) = (\mathbb{N}, \mathcal{P}(\mathbb{N}))$.

Proposition 5.2 *Let $L(T, X)$ and $L(T, X^*)$ be two vector spaces of weakly integrable functions from T to X and X^*, resp., such that $L(T, X)$ is decomposable and the function $t \to \langle v(t), u(t)\rangle$ is integrable for every $(u, v) \in L(T, X) \times L(T, X^*)$. If $f : T \times X \to \overline{\mathbb{R}}$ is a normal integrand such that $\hat{I}_f(u_0) < \infty$ for some $u_0 \in L(T, X)$, and either (X, X^*) is a dual pair of Suslin spaces, or $(T, \mathcal{A}) = (\mathbb{N}, \mathcal{P}(\mathbb{N}))$, then for all $v \in L(T, X^*)$*

$$\hat{I}_{f^*}(v) = \sup_{v \in L(T,X)} \int_T (\langle u(t), v(t)\rangle - f(t, u(t))\, \mathrm{d}\mu(t),$$

Proof. First, we may suppose w.l.o.g. that $\hat{I}_f(u_0) \in \mathbb{R}$; because, for otherwise, $\hat{I}_f(u_0) = -\infty$ and the conclusion holds trivially. So, the proof in the first case of Suslin spaces follows from [22, Theorem VII-7]. For the proof in the second case of $(T, \mathcal{A}) = (\mathbb{N}, \mathcal{P}(\mathbb{N}))$, we denote

$$\begin{aligned} \delta(n) :=& \langle u_0(n), v(n)\rangle - f(n, u_0(n)), \\ \alpha :=& \sup \left\{ \int_T (\langle u(t), v(t)\rangle - f(t, u(t))\, \mathrm{d}\mu(t) : v \in L(T, X) \right\}. \end{aligned}$$

and consider the sequence $(x_k) \in L(\mathbb{N}, X)$ defined for $n > k$ by $x_k(n) = u_0(n)$, and for $n \le k$ by

$$x_k(n) := w_n,$$

where $w_n \in X$ is any vector such that $\langle w_n, v(n)\rangle - f(n, w_n) \ge \max\{f^*(n, v(n)) - \frac{1}{k}, \delta(n)\}$ when $f^*(n, v(n)) < +\infty$, and $\langle w, v(n)\rangle - f(n, w) \ge \max\{k, \delta(n)\}$, otherwize. Then, for every $k, k_0 \in \mathbb{N}$, with $k_0 < k$,

$$\begin{aligned} \alpha &\ge \int_{n \le k_0} \langle x_k(n), v(n)\rangle - f(n, x_k(n))\mathrm{d}\mu(n) + \int_{n > k_0} \langle x_k(n), v(n)\rangle - f(n, x_k(n))\mathrm{d}\mu(n) \\ &\ge \int_{n \le k_0} \langle x_k(n), v(n)\rangle - f(n, x_k(n))\mathrm{d}\mu(n) + \int_{n > k_0} \delta(n)\mathrm{d}\mu(n). \end{aligned}$$

Then, taking the limit on k we get $\alpha \ge \int_{n \le k_0} f^*(n, v(n))\mathrm{d}\mu(n) + \int_{n > k_0} \delta(n)\mathrm{d}\mu(n)$, and the inequality $\alpha \ge \int_{\mathbb{N}} f^*(n, v(n))\mathrm{d}\mu(n)$ follows as k_0 goes to $+\infty$. This finishes the proof because the converse inequality $\alpha \le \int_{\mathbb{N}} f^*(n, v(n))\mathrm{d}\mu(n)$ holds trivially. □

Remark 5.3 t is worth mentioning that in the above result one can always use $L(T, X^*) = \{0\}$, then the above result gives an interchange between infimum and the sign of integral.

The second well-known result gives a representation of the Fenchel conjugate of $\hat{I}_f$ in $L^\infty(T, X)$. This result was first proved in [102, Theorem 1] for the case $X = \mathbb{R}^n$, and in [101, Theorem 4] when X an arbitrary separable reflexive Banach space.

Theorem 5.4 *Let X be a separable reflexive Banach space, and $f : T \times X \to \mathbb{R} \cup \{+\infty\}$ be a normal convex integrand. Assume that the integral functional $\hat{I}_f$ defined on $\mathrm{L}^\infty(T, X)$ is finite at some point in $\mathrm{L}^\infty(T, X)$, and that $\hat{I}_{f^*}$ is finite at some point in $\mathrm{L}^1(T, X)$. Then the Fenchel conjugate of $\hat{I}_f$ on $(\mathrm{L}^\infty(T, X))^*$ is given by, for every $u^* = \ell^* + s^*$ with $\ell^* \in \mathrm{L}^1(T, X^*)$ and $s^* \in \mathrm{L}^{sing}(T, X)$,*

$$(\hat{I}_f)^*(u^*) = \int\limits_T f^*(t, \ell^*(t))\mathrm{d}\mu(t) + \sigma_{\operatorname{dom} \hat{I}_f}(s^*).$$

A straightforward application of the above theorem gives us a representation of the subdifferential of integrand functionals. The proof can be found (for $\varepsilon = 0$) in [102, Corollary 1B] for the finite-dimesional case, and in [80, Proposition 1.4.1.] for arbitrary separable reflexive Banach spaces. The proof of the general case $\varepsilon \geq 0$ is similar, and is given here for completeness.

Proposition 5.5 *With the assumptions of Theorem 5.4, for every $u \in \mathrm{L}^\infty(T, X)$ and $\varepsilon \geq 0$, one has that $u^* = \ell^* + s^* \in \partial_\varepsilon \hat{I}_f(u)$ (with $\ell^* \in \mathrm{L}^1(T, X^*)$ and $s^* \in \mathrm{L}^{sing}\ (T, X)$) if and only if there exists an integrable function $\varepsilon_1 : T \to [0, +\infty)$ and a constant $\varepsilon_2 \geq 0$ such that*

$$\ell^*(t) \in \partial_{\varepsilon_1(t)} f(t, u(t)) \text{ ae, } s^* \in \mathrm{N}^{\varepsilon_2}_{\operatorname{dom} \hat{I}_f}(u), \text{ and } \int_1 \varepsilon_1(t)\mathrm{d}\mu(t) + \varepsilon_2 \leq \varepsilon.$$

Proof. Take $u^* = \ell^* + s^*$ in $\partial_\varepsilon \hat{I}_f(u)$; hence, $u \in \operatorname{dom} \hat{I}_f$. Then by Theorem 5.4 and the definition of ε-subdifferentials we have

$$\int\limits_T \left(f(t, u(t)) + f^*(t, \ell^*(t)) - \langle \ell^*(t), u(t)\rangle\right) \mathrm{d}\mu(t) + \Big(\sigma_{\operatorname{dom} \hat{I}_f}(s^*) - \langle s^*, u\rangle\Big) \leq \varepsilon.$$

Hence, we conclude by setting $\varepsilon_1(t) := f(t, u(t)) + f^*(t, \ell^*(t)) - \langle \ell^*(t), u(t)\rangle$ (≥ 0) and $\varepsilon_2 := \sigma_{\operatorname{dom} \hat{I}_f}(s^*) - \langle s^*, u\rangle$ (≥ 0). □

The next result, also given in [102, Theorem 2], will be used in the proof of Theorem 5.9.

Theorem 5.6 *Let $f : T \times \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a normal convex integrand. Assume that $\bar{u} \in \mathrm{L}^\infty(T, \mathbb{R}^n)$, and that for some $r > 0$ the function $f(\cdot, \bar{u}(\cdot) + x)$ is integrable for every x such that $\|x\| < r$. Then there is some u^* in $\mathrm{L}^1(T, \mathbb{R}^n)$ such that $u^* \in \operatorname{dom} \hat{I}_{f^*}$. Moreover, $\hat{I}_f$ is continuous (in the $\mathrm{L}^\infty(T, \mathbb{R}^n)$-norm) at u wherever $\|u - \bar{u}\|_\infty < r$.*

In the two following lemmas we consider a finite-dimensional Banach space subspace F of X, endowed with its dual F^* with norms $\|\cdot\|_F$ and $\|\cdot\|_{F^*}$, respectively, together with the continuous linear projection $P : X \to F$, whose dual is denoted by P^*.

Lemma 5.7 *There exists a constant $M \geq 0$ and a neighborhood $W \in \mathcal{N}_0$ (only depending on P and F)) such that for every integrable function $u^*(\cdot) : T \to F^*$ the composite function $P^* \circ u^*(\cdot)$ is integrable and satisfies $\sigma_W(u^*(t) \circ P) \leq M\|u^*(t)\|_{F^*}$.*

Proof. Since $P : X \to F$ is a continuous linear mapping, there exist $M \geq 0$ and neighborhood $W \in \mathcal{N}_0$ such that $\|P(x)\|_F \leq M_1 \sigma_{W^\circ}(x)$ for all $x \in X$ and, hence,

$$\sigma_W(u^*(t) \circ P) = \sup_{x \in W} \langle u^*(t), P(x) \rangle \leq M_1 \sup_{y \in B_F(0,1)} \langle u^*(t), y \rangle = M_1 \|u^*(t)\|_{F^*}.$$

We are done since the function $P^* \circ u^*(\cdot)$ inherits the measurabilty from u^*. $\square$

Lemma 5.8 *Assume that both X and X^* are Suslin and let $u^* : T \to X^*$ be a weak*-measurable function. Then*

$$\mathfrak{G} := \{(t, x^*, y^*, v^*) \in T \times X^* \times X^* \times F^* \mid x^* + y^* + P^*(v^*) = u^*(t)\} \in \mathcal{A} \otimes \mathcal{B}(X^* \times X^* \times F^*).$$

Consequently, given measurable multifunctions $C_1, C_2 : T \rightrightarrows X^*$, $C_3 : T \rightrightarrows F^*$, the multifunction $C : T \rightrightarrows X^* \times X^* \times F^*$ defined as

$$(x^*, y^*, z^*) \in C(t) \Leftrightarrow (x^*, y^*, z^*) \in C_1(t) \times C_2(t) \times C_3(t) \text{ and } u^*(t) = x^* + y^* + P^*(z^*),$$

is measurable.

Proof. Consider the functions $g(t, x^*, y^*, v^*) = x^* + y^* + P^*(v^*) - u^*(t)$ and $h(x^*, y^*, v^*) = x^* + y^* + P^*(v^*))$, we claim that h is $(\mathcal{A} \otimes \mathcal{B}(X^* \times X^* \times F^*), \mathcal{B}(X))$-measurable. First assume that u^* is a simple function, that is, there exist a measurable partition of T, $\{T_i\}_{i=1,..n}$ and elements $x_i^* \in X$ such that $u^*(t) = \sum x_i^* \mathbb{1}_{T_i}(t)$, then it easy to see that for every open set U on X^*, $h^{-1}(U) = \bigcup_{i=1}^n T_i \times \varphi^{-1}(B + x_i^*) \in \mathcal{A} \otimes \mathcal{B}(X^* \times X^* \times F^*)$ so h . Now if u^* is measurable, then by [22, Theorem III.36] u^* is the limit of a a sequence of simple functions u_n^*, so considering a countable dense set D on X we can write

$$\mathfrak{G} = \bigcap_{v \in D} \bigcap_{i \geq 1} \bigcup_{j \in \mathbb{N}} \bigcap_{k \geq j} \left\{(t, x^*, y^*, v^*) \in T \times X^* \times X^* \times F^* \mid |\langle x^* + y^* + P^*(v^*) - u_k^*(t), v \rangle| < 1/i\right\},$$

therefore $\mathfrak{G} \in \mathcal{A} \otimes \mathcal{B}(X^* \times X^* \times F^*)$. $\square$

5.3 Characterization via ε-subdifferentials

In this section we characterize the sudifferential of the convex function I_f, by means of the ε-subdifferentials of the functions $f(t, \cdot)$, $t \in T$. As before, X and X^* are two lcs paired in duality, and let $f : T \times X \to \overline{\mathbb{R}}$ be such that for every finite dimensional space F of X, $f_{|F} : T \times F \to \overline{\mathbb{R}}$ a is convex normal integrand.

We start with the main result of this section.

Theorem 5.9 *For every $\varepsilon \geq 0$ we have the formulas*

$$\partial_\varepsilon I_f(x) = \bigcap_{L\in\mathcal{F}(x)} \bigcup_{\substack{\varepsilon=\varepsilon_1+\varepsilon_2\\ \varepsilon_1,\varepsilon_2\geq 0\\ \ell\in\mathcal{I}(\varepsilon_1)}} \left\{ \int_T \partial_{\ell(t)}(f_t + \delta_{\mathrm{aff}\{L\cap \mathrm{dom}\, I_f\}})(x)\mathrm{d}\mu(t) + N^{\varepsilon_2}_{\mathrm{dom}\, I_f\cap L}(x) \right\} \tag{5.3}$$

$$= \bigcap_{L\in\mathcal{F}(x)} \bigcup_{\ell\in\mathcal{I}(\varepsilon)} \left\{ \int_T \partial_{\ell(t)}(f_t + \delta_{L\cap \mathrm{dom}\, I_f})(x)\mathrm{d}\mu(t) \right\}. \tag{5.4}$$

Where $\mathcal{F}(x) := \{V \subseteq X : V$ is finite dimensional linear space and $x \in V\}$ and $\mathcal{I}(\eta) := \{\ell \in \mathrm{L}^1(T, \mathbb{R}_+) : \int_T \ell(t)\mathrm{d}\mu(t) \leq \eta\}$.

Proof. First suppose $x = 0$. Take $x^* \in \partial_\varepsilon I_f(0)$ and choose $L \in \mathcal{F}(0)$ and define $F = \mathrm{span}\{L\cap \mathrm{dom}\, I_f\} = \mathrm{span}\{\mathrm{e_i}\}_{\mathrm{i}=1}^n$ (where $\mathrm{e_i} \in L\cap\mathrm{dom}\, I_f$ is a basis of F). Consider a continuous projection $P : X \to F$ and $\hat{f} : T \times F \to \overline{\mathbb{R}}$ the restriction of f to F, we have $x^* \in \partial_\varepsilon(I_f + \delta_F)(0)$, hence define $y^* = x^* \circ P$ and $N^* = x^* - y^*$. Now $y^* \in \partial_\varepsilon(I_f + \delta_F)(0)$, $y^*_{|F} \in \partial_\varepsilon(I_{\hat{f}})(0)$ and $N^* \in F^\perp$.

Because $\mathrm{dom}\, I_{\hat{f}} = F \cap \mathrm{dom}\, I_f \neq \emptyset$ and $\mathrm{span}\{\mathrm{dom}\, I_f \cap F\} = \mathrm{span}\{\mathrm{dom}\, I_f \cap L\} = F$ is a finite dimensional subspace we have that $I_{\hat{f}}$ is continuous on $\mathrm{ri}_F\, \mathrm{dom}\, I_{\hat{f}}$, that is to say, there exist $\eta > 0$ and $x_0 \in \mathrm{dom}\, I_f \cap L$ such that $x_0 + \eta\, \mathrm{co}\{\pm \mathrm{e_i}\} \subseteq \mathrm{dom}\, I_{\hat{f}}$. Hence if $h \in F$ belongs to $\eta\, \mathrm{co}\{\pm\mathrm{e_i}\}$ we have that $f(\cdot, x_0 + h)$ is integrable, so applying Theorem 5.6 we have $\hat{I}_{\hat{f}}$ is continuous in a neighborhood of x_0 (in $\mathrm{L}^\infty(T, F)$) and the hypotheses of Theorem 5.4 are satisfies. Then we can apply the composition rule to $I_{\hat{f}}$, so $\partial_\varepsilon I_{\hat{f}}(0) = A^*(\partial_\varepsilon \hat{I}_{\hat{f}}(0))$ (see Proposition 1.4), where $A : X \to \mathrm{L}^\infty(T, F)$ is given by $A(h) = h\mathbb{1}_T$ and $A^*(u^* + v^*)(h) = \int_T \langle u^*(t), h\rangle + v(h\mathbb{1}_T)$, where $u^* \in \mathrm{L}^1(T, F^*)$ and $v^* \in \mathrm{L}^{\mathrm{sing}}(T, F)$. Then there are $\alpha^* \in \mathrm{L}^1(T, F^*)$ and $\beta^* \in \mathrm{L}^{\mathrm{sing}}(T, F)$ such that $y^*_{|F}(h) = y^*(h) = \int_T\langle \alpha^*(t), h\rangle \mathrm{d}\mu(t) + \beta^*(h\mathbb{1}_T)$ for all $h \in F$.

Moreover by Proposition 5.5 there exist $\varepsilon_1, \varepsilon_2 \geq 0$ and $\ell \in \mathcal{I}(\varepsilon_1)$ such that $\alpha(t) \in \partial_{\ell(t)} \hat{f}_t(0)$ μ-ae and $\sigma_{\mathrm{dom}\, \hat{I}_{\hat{f}}}(\beta) \leq \varepsilon_2$, so we define $z^* \in \mathrm{L}^1(T, X^*)$ and $\lambda^* \in X^*$ by $z^*(t) = P^*(\alpha^*(t)) = \alpha^*(t) \circ P$ and $\lambda^*(x) = \beta(P(x)\mathbb{1}_T)$ respectively.

From Lemma 5.7 $z^* \in \mathrm{L}^1(T, X^*)$. Now from the fact that $\alpha(t) \in \partial_{\ell(t)} \hat{f}_t(0)$ ae and the definition of z^* we have $z^*(t) \in \partial_{\ell(t)}(f_t + \delta_{\mathrm{span}\{\mathrm{dom}\, I_f\cap L\}})(0)$. Since $A(\mathrm{dom}\, I_{\hat{f}}) \subseteq \mathrm{dom}\, \hat{I}_{\hat{f}}$ we conclude that $\lambda^* \in N_{\mathrm{dom}\, I_f\cap L}(0)$. So we get

$$x^* \in \bigcup_{\substack{\varepsilon=\varepsilon_1+\varepsilon_2\\ \varepsilon_1,\varepsilon_2\geq 0\\ \ell\in\mathcal{L}_{\varepsilon_1}}} \left\{ \int_T \partial_{\ell(t)}(f_t + \delta_{\mathrm{aff}\{L\cap\mathrm{dom}\, I_f\}})(0)\mathrm{d}\mu(t) + N^{\varepsilon_2}_{\mathrm{dom}\, I_f\cap L}(0) \right\}.$$

Finally using the fact that $\bigcap\limits_{L\in\mathcal{F}(x)} \partial_\varepsilon(I_f + \delta_L)(0) = \partial_\varepsilon I_f(0)$, we obtain the first equality of equation 5.3 for $x = 0$.

In the general case consider $g(t,y) = f(t,y+x)$ is easy to verify $\operatorname{epi} g_t = \operatorname{epi} f_t - (x,0)$. Then

$$\partial I_f(x) = \partial I_g(0) = \bigcap_{L\in\mathcal{F}(0)} \bigcup_{\substack{\varepsilon=\varepsilon_1+\varepsilon_2\\ \varepsilon_1,\varepsilon_2\geq 0\\ \ell\in\mathcal{I}(\varepsilon_1)}} \left\{ \int_T \partial_{\ell(t)}(g_t + \delta_{\operatorname{aff}\{L\cap \operatorname{dom} I_g\}})(0)\mathrm{d}\mu(t) + N^{\varepsilon_2}_{\operatorname{dom} I_g \cap L}(0) \right\}.$$

Then if we suppose that $L \in \mathcal{F}(x)$, we have $\partial_{\ell(t)}(g_t+\delta_{\operatorname{aff}\{L\cap \operatorname{dom} I_g\}})(0) = \partial_{\ell(t)}(f_t+\delta_{\operatorname{aff}\{L\cap \operatorname{dom} I_f\}})(x)$ and $N^{\varepsilon_2}_{\operatorname{dom} I_g\cap L}(0) = N^{\varepsilon_2}_{\operatorname{dom} I_f\cap L}(x)$.

To prove Formula (5.4) consider $x^* \in \partial_\varepsilon I_f(x)$ and $L \in \mathcal{F}(x)$, then there exists $\varepsilon_1, \varepsilon_2 \geq 0$ such that $\varepsilon = \varepsilon_1 + \varepsilon_2$, integrable functions $\ell \in \mathcal{I}(\varepsilon_2)$, $y(t) \in \partial_{\ell(t)}(f_t + \delta_{\operatorname{aff}\{L\cap \operatorname{dom} I_f\}})(x)$ and $\lambda^* \in N^{\varepsilon_2}_{\operatorname{dom} I_f\cap L}$ such that $x^* = \int_T y(t)\mathrm{d}\mu(t) + \lambda^*$, so taking $\lambda^*(t) := c(t)\lambda^*$ and $\ell_2(t) = \varepsilon_2 c(t)$, where $c(t) > 0$ and $\int_T c(t)\mathrm{d}\mu(t) = 1$ we get:

$$y^*(t) + \lambda^*(t) \in \partial_{\ell(t)+\ell_2(t)}(f_t + \delta_{\operatorname{aff}\{L\cap \operatorname{dom} I_f\}})(x)$$

and $\int\limits_T \ell(t) + \ell_2(t) = \varepsilon$.

Finally the right side of (5.4) is trivially included in $\partial_\varepsilon I_f(x)$. $\square$

Remark 5.10 It is worth mentioning that theorem above also holds if instead of the set $\mathcal{F}(x)$ we take some subfamily of finite dimensional $\tilde{\mathcal{L}} \subseteq \mathcal{F}(x)$ such that $\bigcap\limits_{n\in\mathbb{N}} \partial_\varepsilon(I_f + \delta_{L_n})(x) = \partial_\varepsilon I_f(x)$, for example if, the space X is separable, or more general if, $\operatorname{epi} I_f$ is separable, we can take $(x_{\mathrm{i}}, \alpha_{\mathrm{i}})_{\mathrm{i}\geq 1}$ dense in $\operatorname{epi} I_f$, so we define $L_n := \operatorname{span}\{x, x_{\mathrm{i}}\}_{\mathrm{i}}^n$ and it is easy to see that $\bigcap\limits_{n\in\mathbb{N}} \partial_\varepsilon(I_f + \delta_{L_n})(x) = \partial_\varepsilon I_f(x)$.

Remark 5.11 The reader can notice that in Theorem 5.9 we can consider the hypotheses for every finite dimensional subspace F of X, $f_{|F}$ is normal integrand and $\overline{\operatorname{co}}_F I_f = I_{\overline{\operatorname{co}}_F f}$, we have the formulas 5.3 and 5.4 change to

$$\partial_\varepsilon I_f(x) \subseteq \bigcap_{L\in\mathcal{F}(x)} \bigcup_{\substack{\varepsilon=\varepsilon_1+\varepsilon_2\\ \varepsilon_1,\varepsilon_2\geq 0\\ \ell\in\mathcal{I}(\varepsilon_1)}} \left\{ \int_T \partial_{\ell(t)+m_{x,F}(t)}(f_t + \delta_{\operatorname{aff}\{L\cap \operatorname{dom} I_f\}})(x)\mathrm{d}\mu(t) + N^{\varepsilon_2}_{\operatorname{dom} I_f\cap L}(x) \right\},$$

$$\partial_\varepsilon I_f(x) \subseteq \bigcap_{L\in\mathcal{F}(x)} \bigcup_{\ell\in\mathcal{I}(\varepsilon)} \left\{ \int_T \partial_{\ell(t)+m_{x,F}(t)}(f_t + \delta_{L\cap \operatorname{dom} I_f})(x)\mathrm{d}\mu(t) \right\},$$

respectively, where $m_{x,F}(t) := f(t,x) - \overline{\operatorname{co}}_F f(t,x)$ is the modulus of convexity over F. Moreover we notice that if $\partial_\varepsilon I_f(x) \neq \emptyset$, then $\int\limits_T m_{x,L}(t)\mathrm{d}\mu(t) \leq \varepsilon$. Indeed take $F \in \mathcal{F}(x)$, on the one hand for every $t \in T$, $f(t,x) \geq \overline{\operatorname{co}}(f(t,\cdot)_{|F})(x) \geq \overline{\operatorname{co}} f(t,x)$. On the other hand if $\partial_\varepsilon I_f(x) \neq \emptyset$ we have $I_f(x) \leq \overline{\operatorname{co}} I_f(x) + \varepsilon$ and by the hypothesis $I_{\overline{\operatorname{co}} f}(x) = \overline{\operatorname{co}} I_f(x)$, so $I_f(x) \leq I_{\overline{\operatorname{co}} f}(x) + \varepsilon$, that implies, $\int\limits_T f(t,x) - \overline{\operatorname{co}}(f(t,\cdot)_{|F})(x)\mathrm{d}\mu(t) \leq \varepsilon$. In particular if f is a convex normal integrand we have $m_x(\cdot) = 0$. We do not put this in the principal statement of Theorem 5.9 because this formulas can be deduced from the Formulas 5.3 and 5.4.

The next example justifies the use of the normal cone inside the integral sign, even if the date is smooth and the space is finite dimensional.

Example 5.12 Consider the function $f(x) := \frac{b}{a}x + b + \delta_{[-\eta,\eta]}$ $a, b, \eta > 0$, then we have $\partial_\varepsilon f(0) = \left[-\frac{\varepsilon}{\eta} + \frac{b}{a}, \frac{\varepsilon}{\eta} + \frac{b}{a}\right]$. Indeed,

$$\begin{aligned}\alpha \in \partial_\varepsilon f(0) &\Leftrightarrow \alpha \cdot x \leq f(x) - f(0) + \varepsilon \ \ \forall x \in [-\eta, \eta]\\ &\Leftrightarrow \alpha \cdot x \leq \frac{b}{a}x + b - b + \varepsilon \ \ \forall x \in [-\eta, \eta]\\ &\Leftrightarrow (\alpha - \frac{b}{a}) \cdot x \leq \varepsilon \ \ \forall x \in [-\eta, \eta].\end{aligned}$$

Then, it is clear that the last inequality holds if and only if $\alpha \in \left[-\frac{\varepsilon}{\eta} + \frac{b}{a}, \frac{\varepsilon}{\eta} + \frac{b}{a}\right]$. Using the previous function one constructs a normal integrand f $:]0,1] \times \mathbb{R} \to [0, +\infty]$ by $f(t,x) := \frac{b(t)}{a(t)}x + b(t) + I_{[-\eta(t),\eta(t)]}$, with $b(t) = \frac{1}{\sqrt{t}} + 1$, $a(t) = \delta(t) = t$. Hence we compute

$$I_f(x) = \begin{cases} \displaystyle\int_0^1 (1 + \frac{1}{\sqrt{t}})\mathrm{d}t & \text{if } x = 0,\\ \quad +\infty & \text{if } x \neq 0.\end{cases}$$

That implies $\partial I_f(0) = \mathbb{R}$, but $\int_T \partial_\varepsilon f(t, 0) = \emptyset$ for $\varepsilon < 1$.

Remark 5.13 The only thing one needs to simplify formulas of Theorem 5.9 is a kind of *Qualification Condition* (QC) (see for example [132, Theorem 2.8.3]), that allow us separate the subdifferential of the sum $\partial_\varepsilon(f_t + \delta_{\mathrm{aff}\, L \cap \mathrm{dom}\, I_f})(x) = \partial_\varepsilon f_t(x) + \partial \delta_{\mathrm{aff}\, L \cap \mathrm{dom}\, I_f}(x)$, for instance:

(i) $X = \mathbb{R}^n$ and $\mathrm{ri}(\mathrm{dom}\, f_t) \cap \mathrm{ri}(\mathrm{dom}\, I_f) \neq \emptyset$.

(ii) *Attouch-Brézis* X Banach and $\mathbb{R}_+(\mathrm{dom}\, f_t - \mathrm{aff}(\mathrm{dom}\, I_f \cap L))$ is a closed subspace for every $L \in \mathcal{F}(x)$.

(iii) X lcs and f_t continuous at some point of $\mathrm{dom}\, I_f$.

(iv) *Fenchel-Rockafellar* for every $L \in \mathcal{F}(x)$ and every $U \in \mathcal{N}_0$ there exist $\lambda > 0$ and $V \in \mathcal{N}_0$ such that $V \cap \mathrm{span}\left\{ \mathrm{dom}\, f_t - \mathrm{aff}(\mathrm{dom}\, I_f \cap L)\right\} \subseteq \{f_t \leq r\} \cap B - \mathrm{aff}(\mathrm{dom}\, I_f \cap L)$.

(v) $\left(f + \delta_{\mathrm{aff}(\mathrm{dom}\, I_f \cap L)}\right)^*(x^*) = \min\left\{f^*(y^*) : x^* - y^* \in \mathrm{aff}(\mathrm{dom}\, I_f \cap L)^\circ\right\}$ for every $x^* \in X^*$ and every $L \in \mathcal{F}(x)$.

Corollary 5.14 *In the setting of Theorem 5.9 suppose one* qualification condition *holds (for example one of the remark above) in a measurable set T_{QC}, then:*

$$\partial I_f(x) = \bigcap_{L \in \mathcal{F}(x)} \left\{ \int_{T_{QC}} \left(\partial f_t(x) + \partial \delta_{\mathrm{aff}(\mathrm{dom}\, I_f \cap L)}(x)\right) \mathrm{d}\mu(t) + \int_{T_{QC}^c} \partial(f_t + \delta_{\mathrm{aff}\{L \cap \mathrm{dom}\, I_f\}})(x)\mathrm{d}\mu(t) \right.$$
$$\left. + N_{L \cap \mathrm{dom}\, I_f}(x) \right\} \tag{5.5}$$

Moreover if T is finite we have:

$$\partial\left(\sum_{i\in T} f_i\right)(x) = \bigcap_{\substack{\varepsilon_i>0, i\in T_{QC}^c \\ \sum \varepsilon_i=\varepsilon}} \mathrm{cl}\left\{\sum_{i\in T_{QC}} \partial f_i(x) + \partial_{\varepsilon_i} f_i(x)\right\} \tag{5.6}$$

Proof. Equation (5.5) is direct from Theorem 5.9. Now we proof the second formula fix $V \in \mathcal{N}_0$, by (

$$\begin{aligned}\partial\left(\sum_{i\in T} f_i\right)(x) &\subseteq \sum_{i\in T_{QC}} \partial f_i(x) + N_{L\cap \mathrm{dom}\, I_f}(x) + \sum_{i\in T_{QC}^c} \partial(f_t + \delta_{\mathrm{aff}\{L\cap \mathrm{dom}\, I_f\}})(x) + N_{L\cap \mathrm{dom}\, I_f}(x)\\ &\subseteq \sum_{i\in T_{QC}} \partial f_i(x) + N_{L\cap \mathrm{dom}\, I_f}(x) + \sum_{i\in T_{QC}^c} \partial_{\varepsilon_i} f_t(x) + N_{L\cap \mathrm{dom}\, I_f}(x) + N_{L\cap \mathrm{dom}\, I_f}(x),\end{aligned}$$

here we have applied Theorem 3.14 to upper-estimate

$$\partial(f_t + \delta_{\mathrm{aff}\{L\cap \mathrm{dom}\, I_f\}})(x) = \bigcap_{\delta>0} \mathrm{cl}^{w^*}\left\{\partial_\delta f_t(x) + N_{L\cap \mathrm{dom}\, I_f}(x)\right\},$$

then we can take $V_i \in \mathcal{N}_0$, $i \in T_{QC}^c$ such that $\sum\limits_{i\in T_{QC}^c} V_i \subseteq V$, on the other hand because $N_{L\cap \mathrm{dom}\, I_f}(x) + N_{L\cap \mathrm{dom}\, I_f}(x) + N_{L\cap \mathrm{dom}\, I_f}(x) = N_{L\cap \mathrm{dom}\, I_f}(x)$ we conclude that

$$\partial\left(\sum_{i\in T} f_i\right)(x) \subseteq \sum_{i\in T_{QC}} \partial f_i(x) + \sum_{i\in T_{QC}^c} \partial_{\varepsilon_i} f_t(x) + N_{L\cap \mathrm{dom}\, I_f}(x) + V,$$

Moreover if $\partial f_i(x) \neq \emptyset$ we have $N_{\mathrm{dom}\, I_f\cap L} = \left[\mathrm{cl}\left(\sum_{i\in T_{QC}} \partial f_i(x) + \sum_{i\in T_{QC}^c} \partial_{\varepsilon_i} f_t(x) + L^\perp\right)\right]_\infty$ (see [54, Lemma 11]), therefore if $L^\perp \subseteq V$

$$\partial\left(\sum_{i\in T} f_i\right)(x) \subseteq \sum_{i\in T_{QC}} \partial f_i(x) + \sum_{i\in T_{QC}^c} \partial_{\varepsilon_i} f_t(x) + V + V,$$

and consequently

$$\begin{aligned}\partial\left(\sum_{i\in T} f_i\right)(x) &\subseteq \bigcap_{\substack{\varepsilon_i\geq 0, i\in T_{QC}^c \\ \sum \varepsilon_i=\varepsilon}} \bigcap_{V\in\mathcal{N}_0} \left[\sum_{i\in T_{QC}} \partial f_i(x) + \sum_{i\in T_{QC}^c} \partial_{\varepsilon_i} f_t(x) + V + V\right]\\ &= \bigcap_{\substack{\varepsilon_i\geq 0, i\in T_{QC}^c \\ \sum \varepsilon_i=\varepsilon}} \mathrm{cl}\left\{\sum_{i\in T_{QC}} \partial f_i(x) + \sum_{i\in T_{QC}^c} \partial_{\varepsilon_i} f_t(x)\right\}.\end{aligned}$$

□

Example 5.15 The main feature of the finite parametrized case given in Equation (5.6) is that the characterization of $\partial I_f(x)$ does not involve the normal cone $N_{\mathrm{dom}\, I_f\cap L}(x)$. This fact is specific to this finite case and cannot be true in general, even for smooth data functions

f_t with $\int \partial f(t,x)\mathrm{d}\mu(t) \neq \emptyset$. For example, consider the Lebesgue measure on $]0,1]$ and the integrand $f : T \times \mathbb{R} \to \mathbb{R}$ given by $f(t,x) = x^2/t$. Then we obtain that $I_f = \delta_{\{0\}}$ and, so, $\partial f(t,0) = \{0\}$, while $\partial I_f(0) = \mathbb{R}$. The same example can be adapted to construct similar counterexamples to Equation (5.6) for countable measure over the measurable space $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$.

Let us illustrate a possible applications of the Corollary above using the relative interior, but first we need the following lemma (to avoid the use of the classical sum rule).

Lemma 5.16 *Let $f \in \Gamma_0(X)$. Assume that f is continuous at $x \in \operatorname{dom} f$ relative to $\operatorname{aff}(\operatorname{dom} f)$, let L be an finite dimensional affine subspace which contains x, then for every $x^* \in X^*$*

$$\big(f+\delta_L\big)^*(x^*) = \min\big\{f^*(y^*) + \delta_L^*(x^*-y^*) : y^* \in X^*\big\}.$$

Proof. Because the inequality $\leq$ always holds, we prove the opposite. First suppose $x=0$, define $W = \operatorname{lin}(\operatorname{dom} f)$ and $Z := \operatorname{lin}\big(\operatorname{dom} f - L\big)$, it is easy to see that there exists a finite dimensional subspace $\tilde{L} \subseteq L$ such that $Z = W \oplus \tilde{L}$. From the fact that $\operatorname{dom} f \subseteq W$ one see easily that for every $z^* \in W^*$ and every continuous linear extension $\tilde{z}^* \in X^*$ of z^*,

$$f^*(\tilde{z}^*) = (f_{|Z})^*(\tilde{z}^*_{|Z}) = (f_{|W})^*(z^*). \tag{5.7}$$

Since $0 \in \operatorname{dom} f \cap L$ we have $f + \delta_L$ is proper, then $\big(f+\delta_L\big)^* = \operatorname{cl}^{w^*}\{f^* \Box \delta_L^*\}$. Take $x^* \in X^*$, then the previous equality allows us take nets $x^*_{1,\mathrm{i}}, x^*_{2,\mathrm{i}}$, $\mathrm{i} \in I$ such that $x^*_{1,\mathrm{i}} + x^*_{2,\mathrm{i}} =: w^*_{\mathrm{i}} \rightharpoonup x^*$ and $f^*(x^*_{1,\mathrm{i}}) + \delta_L^*(x^*_{2,\mathrm{i}}) \to \big(f+\delta_L\big)^*(x^*)$, since $\delta_L^* = \delta_{L^\perp}$ we can assume that $x^*_{2,\mathrm{i}} \in L^\perp$. Now consider $\|\cdot\|_{\tilde{L}}$ a norm in $\tilde{L}$ and B the closed unit ball in $\tilde{L}$ respect to this norm, set $M := \sup\{\|w^*_{\mathrm{i}}\|_{\tilde{L}} : \mathrm{i} \in I\} < +\infty$ and consider P_W and $P_{\tilde{L}}$ continuous projections from Z to W and $\tilde{L}$ respectively, because f is continuous (relative to W) at 0 there exists a real number $r \geq \sup\{f(0), f^*(x_{1,\mathrm{i}}) : \mathrm{i} \in I\}$ such that $\{f \leq r\}$ is a neighborhood of 0 in W, then $U := P_W^{-1}(\{f \leq r\}) \cap P_{\tilde{L}}^{-1}(B)$ is a neighborhood of 0 relative to Z, moreover for every $\mathrm{i} \in I$ and every $u \in U$

$$\begin{aligned}
\langle x^*_{1,\mathrm{i}}, u\rangle &= \langle x^*_{1,\mathrm{i}}, P_W(u)\rangle + \langle x^*_{1,\mathrm{i}}, u - P_W(u)\rangle \\
&= \langle x^*_{1,\mathrm{i}}, P_W(u)\rangle - f(P_W(u)) + f(P_W(u)) + \langle x^*_{1,\mathrm{i}}, P_{\tilde{L}}(u)\rangle \\
&\leq f^*(x^*_{1,\mathrm{i}}) + f(P_W(u)) + \langle x^*_{1,\mathrm{i}} + x^*_{2,\mathrm{i}}, P_{\tilde{L}}(u)\rangle \\
&= f^*(x^*_{1,\mathrm{i}}) + f(P_W(u)) + \langle w^*_{\mathrm{i}}, P_{\tilde{L}}(u)\rangle \leq 2r + M.
\end{aligned}$$

Then by Banach-Alaoglu-Bourbaki's theorem [46, Theorem 3.37] there exists a subnet

$$(x^*_{1,\beta})_{|Z} \rightharpoonup x^*_1 \in Z^*,$$

which implies $(x^*_{2,\beta})_{|Z} \rightharpoonup x^*_2 \in Z^*$ and $\langle x^*_2, h\rangle = 0$ for all $h \in L$. Then using 5.7 and $\delta_L^* = \delta_{L^\perp}$ we get

$$(f_{|Z})^*(x^*_1) \leq \liminf (f_{|Z})^*((x^*_{2,\beta})_{|Z}) = \liminf f^*(x^*_{2,\beta}) \leq \lim\big(f^*(x^*_{2,\beta}) + \delta_L^*(x^*_{2,\beta})\big) = (f+\delta_L)^*(x^*).$$

Then take some extension $\tilde{x}^*_1, \tilde{x}^*_2 \in X^*$ of x^*_1 and x^*_2 respectively, define $\tilde{x}^* = \tilde{x}^*_1 + \tilde{x}^*_2$ and $y^* := \tilde{x}^*_1 + x^* - \tilde{x}^*$, so $y^* + x^*_2 = x^*$ and y^* is an extension of x^*_1, then using 5.7 we get

$f^*(y^*) + \delta_L^*(x_2^*) = (f_{|Z})^*(x_1^*) + \delta_L^*(x_2^*) \leq (f + \delta_L)^*(x^*)$. Finally if x is an arbitrary point, take $g(\cdot) := f(\cdot + x)$ and the $L - x$ as the affine subspace, then by the previous

$$\big(g + \delta_{L-x}\big)^*(x^*) = \min\big\{g^*(y^*) + \delta_{L-x}^*(x^* - y^*) : y^* \in X^*\big\},$$

since $\big(g + \delta_{L-x}\big)^*(x^*) = \big(f + \delta_L\big)^*(x^*) - \langle x^*, x\rangle$, $g^*(y^*) = f^*(y^*) - \langle y^*, x\rangle$ and $\delta_{L-x}^*(x^* - y^*) = \delta_L^*(x^* - y^*) - \langle x^* - y^*, x\rangle$ we conclude the result. □

Corollary 5.17 *Let* $\{g_i\}_{i=1}^k, \{g_1\}_{i=k+1}^p \subseteq \Gamma_0(X)$ *such that* $\bigcap\limits_{i\leq k} \mathrm{ri}_{\mathrm{aff}(\mathrm{dom}\, g_i)}(\mathrm{dom}\, g_i) \cap \bigcap\limits_{j\geq k+1} \mathrm{dom}\, g_j \neq \emptyset$ *and for every* $j \leq k$, g_j *is continuous on* $\mathrm{ri}_{\mathrm{aff}(\mathrm{dom}\, g_i)}(\mathrm{dom}\, g_i)$. *Then*

$$\partial\Big(\sum_{i=1}^p g_i\Big)(x) = \bigcap_{\substack{\varepsilon_i > 0, i>k\\ \sum \varepsilon_i = \varepsilon}} \mathrm{cl}\Big\{\sum_{i\leq k} \partial\, g_i(x) + \sum_{i>k} \partial_{\varepsilon_i}\, g_i(x)\Big\}. \tag{5.8}$$

Proof. Take $x \in X$ and $x_0 \in \bigcap\limits_{i\leq k} \mathrm{ri}_{\mathrm{aff}(\mathrm{dom}\, g_i)}(\mathrm{dom}\, g_i) \cap \bigcap\limits_{j\geq k+1} \mathrm{dom}\, g_j \neq \emptyset$, define $g = \sum\limits_{i=1}^p g_i$, then we apply Lemma 5.16 to g_i and $L := \mathrm{aff}(\mathrm{dom}(g) \cap F)$ with $F \in \mathcal{F}(x)$ such that $x_0 \in F$, therefore 5.13 (e) is satisfied, then by Corollary 5.14 we conclude the formula. □

Remark 5.18 We notice that to prove (5.8) it is only necessary that for every $j \leq k$ there exists $x_j \in \mathrm{ri}(\mathrm{dom}\, g_j) \cap \bigcap\limits_{i\neq j} \mathrm{dom}\, g_i \neq \emptyset$ such that g_j is continuous at x_j relative to $\mathrm{aff}(\mathrm{dom}\, g_j)$. Obviously both statement are equivalent when $\mathrm{aff}(\mathrm{dom}\, g_i)$ is closed. It is worth nothing that when there exist some j such that

$$\big[\partial\, g_j(x)\big]_\infty \bigcap -\big[\sum_{i\neq j} \partial_{\varepsilon_i}\, g_i(x)\big]_\infty = N_{\mathrm{dom}\, g_j}(x) \bigcap -N_{\mathrm{dom}\, \sum\limits_{i\neq j} g_i}(x) = \{0\},$$

for example if g_j is continuous at some point of $\mathrm{dom}(\sum\limits_{i\neq j} \partial\, g_i))$, then the formula above is simplified to

$$\partial\Big(\sum_{i=1}^p g_i\Big)(x) = \partial\, g_j(x) + \bigcap_{\substack{\varepsilon_i > 0, i>k\\ \sum \varepsilon_i = \varepsilon}} \mathrm{cl}\Big\{\sum_{\substack{i\leq k\\ i\neq j}} \partial\, g_i(x) + \sum_{i>k} \partial_{\varepsilon_i}\, g_i(x)\Big\}.$$

5.4 Suslin spaces or discrete measure space

In this section we give more sharp characterizations of the ε-subdifferential of I_f under the cases where either X, X^* are Suslin spaces, or $(T, \mathcal{A}) = (\mathbb{N}, \mathcal{P}(\mathbb{N}))$. These settings indeed permit the use of mesurable selection theorems, which gives us more control on the integration of the multifunctions $\partial_{\varepsilon(t)} f_t(x)$ and $N^\varepsilon_{\mathrm{dom}\, I_f \cap L}(x)$. We recall that $f : T \times X \to \mathbb{R} \cup \{+\infty\}$ is a given normal convex integrand, and $(T, \mathcal{A}, \mu)$ is a complete σ-finite measure space. The function $I_f : X \to \mathbb{R} \cup \{+\infty\}$ is defined as

$$I_f(x) = \int_T f(t, x) \mathrm{d}\mu(t).$$

Then the following corollary makes sharper the characterization given in Theorem 5.9 by using only the ε-subdifferential of the f_t's.

In the next corollary we apply approximated sum rule (see [56]) to obtain simplifications of the formulas presented in Theorem 5.9

Corollary 5.19 *We suppose that either X and X^* are Suslin spaces or $(T,\Sigma) = (\mathbb{N}, \mathcal{P}(\mathbb{N}))$. Then for every $x \in X$ and $\varepsilon \geq 0$ we have*

$$\partial_\varepsilon I_f(x) = \bigcap_{L\in\mathcal{F}(x)} \bigcup_{\substack{\varepsilon_1,\varepsilon_2\geq 0\\ \varepsilon=\varepsilon_1+\varepsilon_2\\ \ell\in\mathcal{I}(\varepsilon_1)\\ \eta\in L^1(T,(0,+\infty))}} \bigcap_{\eta\in L^1(T,(0,+\infty))} \operatorname{cl}\left\{\int_T \left(\partial_{\ell(t)+\eta(t)} f_t(x) + N^{\varepsilon_2}_{\operatorname{dom} I_f\cap L}(x)\right) \mathrm{d}\mu(t)\right\},$$

where the closure is taking in the strong topology $\beta(X^*, X)$.

Proof. We only need to prove the inclusion " $\subset$ " in which we suppose that $x = 0$. Take $F \in \mathcal{F}(0)$, $\varepsilon > 0$ and $L := \operatorname{span}\{F \cap \operatorname{dom} I_f\}$ $(= \operatorname{aff}\{F \cap \operatorname{dom} I_f\})$, say $L = \operatorname{span}\{\mathrm{e_i}\}_1^p$ with $\{\mathrm{e_i}\}_1^p$ being linearly independent so that $\operatorname{co}\{\pm \mathrm{e_i}\}_{\mathrm{i}=1}^p$ is the united closed ball in L with respect to a norm $\|\cdot\|_L$ (on L). Let $P : X \to L$ be a continuous projection, $M \geq 0$, and $W \in \mathcal{N}_0$ as in Lemma 5.7.Given $\delta > 0$, we pick an integrable function $\gamma : T \to (0, +\infty)$ such that

$$M \int_T \gamma(t) \mathrm{d}\mu \leq \delta, \tag{5.9}$$

and define the measurable multifunctions U, $V : T \rightrightarrows L^*$ as

$$U(t) := \{x^* \in X^* : |\langle x^*, \mathrm{e_i}\rangle| \leq \gamma(t),\ \mathrm{i} = 1, ..., p\},\ V(t) := \{x^* \in L^* : |\langle x^*, \mathrm{e_i}\rangle| \leq \gamma(t),\ \mathrm{i} = 1, ..., p\}.$$

Now take $x^* \in \partial_\varepsilon I_f(0)$ and fix a positive measurable function η . By formula 5.3 in Theorem 5.9 there exist $\varepsilon_1, \varepsilon_2 \geq 0$ with $\varepsilon_1 + \varepsilon_2 = \varepsilon$, $\ell \in \mathcal{I}(\varepsilon_1)$, an integrable selection $x^*_{L,\varepsilon}(\cdot)$ of the multifunction $t \to \partial_{\ell(t)} (f_t + \delta_L)(0)$, and $\lambda^* \in N^{\varepsilon_2}_{\operatorname{dom} I_f\cap L}(0)$ such that $x^* = \int_T x^*_{L,\varepsilon}(t) \mathrm{d}\mu(t) + \lambda^*$. Since $U(t) \in \mathcal{N}_0$, by the Hiriart-Urruty-Phelps sum rule (see Proposition 1.2) we have that, for ae $t \in T$,

$$\begin{aligned} x^*_{L,\varepsilon}(t) \in \partial_{\ell(t)} (f_t + \delta_L)(0) &\subset \partial_{\ell(t)+\eta(t)} f_t(0) + L^\perp + U(t) \\ &\subset \partial_{\ell(t)+\eta(t)} f_t(0) + N_{\operatorname{dom} I_f\cap L}(0) + P^*(V(t)). \end{aligned}$$

We define the multifunction $G : T \rightrightarrows X^* \times X^* \times F^*$ as

$$(y^*, w^*, v^*) \in G(t) \Leftrightarrow \left\{ \begin{array}{r} y^* \in \partial_{\ell(t)+\eta(t)} f(t, 0),\ w^* \in N_{\operatorname{dom} I_f\cap F}, \text{ and } v^* \in V(t), \\ x^*_{L,\varepsilon}(t) = y^* + w^* + P^*(v^*). \end{array} \right.$$

If X, X^* are Suslin spaces, then by Lemma 5.8 G is measurable, and so, by Theorem 1.21 it admits a mesurable selection $(y^*(\cdot), w^*(\cdot), v^*(\cdot))$. This also obviously holds when $(T, \Sigma) = (\mathbb{N}, \mathcal{P}(\mathbb{N}))$. Thus, by Lemma 5.7 the function $u^*(t) := v^*(t) \circ P$ is integrable and we get

$$\sigma_W(u^*(t)) \leq M \max_{\mathrm{i}=1,...,p} \langle v^*(t), \mathrm{e_i}\rangle \leq M\gamma(t) \text{ for ae } t \in T.$$

Consequently, the function $y^* + w^* = x^*_{L,\varepsilon}(\cdot) - u^*(\cdot)$ is strongly integrable and we have (recall (5.9))

$$\begin{aligned}
\sigma_W(x^* - \int\limits_T (y^*(t) + w^*(t))\mathrm{d}\mu(t) - \lambda^*) &= \sigma_W(\int_T x^*_{L,\varepsilon}(t)\mathrm{d}\mu(t) - \int\limits_T (y^*(t) + w^*(t))\mathrm{d}\mu(t)) \\
&= \sigma_W(\int_T u^*(t)\mathrm{d}\mu(t)) \\
&\leq \int_T \sigma_W(u^*(t))\mathrm{d}\mu(t) \\
&\leq M\int_T \gamma(t)\mathrm{d}\mu \leq \delta;
\end{aligned}$$

that is,

$$x^* - \int\limits_T (y^*(t) + w^*(t))\mathrm{d}\mu(t) - \lambda^* \in \delta W^\circ,$$

and due to the arbitrariness of δ we obtain

$$x^* \in \mathrm{cl}^{\beta(X^*,X)} \int\limits_T \left(\partial_{\ell(t)+\eta(t)} f_t(0) + N_{\mathrm{dom}\, I_f \cap L}(0)\right) \mathrm{d}\mu(t) + N^{\varepsilon_2}_{\mathrm{dom}\, I_f \cap L}(0).$$

Finally, since $N^{\varepsilon_2}_{\mathrm{dom}\, I_f \cap L}(0) \subset \int\limits_T N^{\varepsilon_2}_{\mathrm{dom}\, I_f \cap L}(0)\mathrm{d}\mu(t)$ we conclude that

$$x^* \in \mathrm{cl}^{\beta(X^*,X)} \int\limits_T \left(\partial_{\ell(t)+\eta(t)} f_t(0) + N^{\varepsilon_2}_{\mathrm{dom}\, I_f \cap L}(0)\right) \mathrm{d}\mu(t).$$

□

The next result is a finite-dimensional-like characterization of the subdifferential of I_f. Recall that a closed affine subspace $A \subset X$ is said to have a continuous projection if there exists a continuous projection from X to A.

Theorem 5.20 *Let X, X^* and T be as in Theorem 5.19. If I_f is continuous on $\mathrm{ri}(\mathrm{dom}\, I_f) \neq \emptyset$ and $\overline{\mathrm{aff}}(\mathrm{dom} I_f)$ has a continuous projection, then*

$$\partial I_f(x) = \bigcap_{\eta \in L^1(T,(0,+\infty))} \mathrm{cl}^{w^*} \left\{ (w)\text{-}\int\limits_T \left(\partial_{\eta(t)} f_t(x) + N_{\mathrm{dom}\, I_f}(x)\right) \mathrm{d}\mu(t) \right\}.$$

Proof. Because the inclusion " $\supset$ " is immediate we only need to prove the other inclusion " $\subset$ " when $x = 0$; hence, $F := \overline{\mathrm{aff}}(\mathrm{dom} I_f)$ is a closed subspace of X. Let $x^* \in \partial I_f(0)$, $\eta \in L^1(T, (0, +\infty))$, and $V := \{h^* \in X^* : |\langle h^*, \mathrm{e_i}\rangle| \leq 1,\ \ \mathrm{i} = 1, ..., p\}$ for some $\{\mathrm{e_i}\}_{\mathrm{i}=1}^p \subset X$. By the current assumption, we take $x_0 \in \mathrm{ri}(\mathrm{dom}\, I_f)$ and a continuous projection $P : X \to F$. Define $L = \mathrm{span}\{\mathrm{e_i}, P(\mathrm{e_i}), x_0\}_{\mathrm{i}=1}^p$ and $W(t) := \{h^* \in X^* : \max\{|\langle h^*, \mathrm{e_i}\rangle|, |\langle h^*, P(\mathrm{e_i})\rangle|, |\langle h^*, x_0\rangle|\} \leq \varepsilon(t),\ \mathrm{i} = 1, ..., p\}$, where $\varepsilon(\cdot)$ is any positive integrable function with values on $(0, 1)$ and

$\int_T \varepsilon \mathrm{d}\mu \leq 1/2$. Then $L^\perp + W(t) \subseteq W(t) \subseteq V$. Because $L \cap \mathrm{ri}(\mathrm{dom}\, I_f) \neq \emptyset$ we have (see, e.g., Corollary 5.17)

$$N_{\mathrm{dom}\, I_f \cap L}(0) = \mathrm{cl}^{w^*}\left(L^\perp + N_{\mathrm{dom}\, I_f}(0)\right). \tag{5.10}$$

By Theorem 5.19 there exists a (strong) integrable selection $y^*(t) \in \partial_{\eta(t)} f_t(0) + N_{\mathrm{dom}\, I_f \cap L}(0) \subset \partial_{\eta(t)} f_t(0) + N_{\mathrm{dom}\, I_f}(0) + W(t)$, due to (5.10), such that

$$x^* - \int_T y^* \mathrm{d}\mu \in V. \tag{5.11}$$

Also, by the measurability of multifunctions $\partial_{\eta(\cdot)} f_\cdot(0)$, $N_{\mathrm{dom}\, I_f}(0)$, and $W(\cdot)$ (see, e.g., [66]), there exists a (weakly) measurable selection $z^*(\cdot)$ of $\partial_{\eta(\cdot)} f_\cdot(0) + N_{\mathrm{dom}\, I_f}(0)$ such that $y^*(t) - z^*(t) \in W(t)$ for ae t (the existence of such a selection is guaranteed for Suslin spaces by the representation theorem of Castaing, while it is straightforward in the discrete case). Let us verify that the function $z^*(\cdot) \circ P$ is weakly integrable: Given $U \in \mathcal{N}_0$ such that $x_0 + P(U) \subset \mathrm{ri}(\mathrm{dom}\, I_f)$ (using the the continuity of P) we have, for every $y \in U$,

$$\begin{aligned}
\langle z^*(t) \circ P, y\rangle &= \langle z^*(t), Py\rangle \\
&= \langle z^*(t), x_0 + Py\rangle - \langle z^*(t), x_0\rangle \\
&\leq f(t, x_0 + P(y)) - f(t, 0) + \eta(t) - \langle z^*(t), x_0\rangle + \sigma_{N_{\mathrm{dom}\, I_f}(0)}(x_0 + P(y)) \\
&\leq f(t, x_0 + P(y)) - f(t, 0) + \eta(t) + |\langle z^*(t) - y^*(t), x_0\rangle| + |\langle y^*(t), x_0\rangle| \\
&\leq f(t, x_0 + P(y)) - f(t, 0) + \eta(t) + \varepsilon(t) + |\langle y^*(t), x_0\rangle|,
\end{aligned} \tag{5.12}$$

and the weak integrability of $z^*(\cdot) \circ P$ follows, as

$$\begin{aligned}
\int_T |\langle z^*(t) \circ P, y\rangle| \,\mathrm{d}\mu(t) \leq \int_T |f(t, x_0 + P(y))| \,\mathrm{d}\mu(t) + \int_T |f(t, x_0 - P(y))| \,\mathrm{d}\mu(t) - I_f(0) \\
+ \int_T (\eta(t) + \varepsilon(t) + |\langle y^*(t), x_0\rangle|) \mathrm{d}\mu(t) < +\infty.
\end{aligned}$$

Moreover, (5.12) implies that $\int_T z^* \circ P \mathrm{d}\mu$ is uniformally bounded on a neighborhood of zero, so that $\int_T z^* \circ P \mathrm{d}\mu \in X^*$. Finally, we have

$$z^*(t) \circ P = z^*(t) + z^*(t) \circ P - z^*(t) \in \partial_{\eta(t)} f_t(0) + N_{\mathrm{dom}\, I_f}(0) + F^\perp = \partial_{\eta(t)} f_t(0) + N_{\mathrm{dom}\, I_f}(0),$$

and (recall (5.11))

$$x^* - \int_T z^* \circ P \mathrm{d}\mu = x^* - \int_T y^* \mathrm{d}\mu + \int_T (y^* - z^*) \mathrm{d}\mu - \int_T (z^* \circ P - z^*) \mathrm{d}\mu \in V + V + F^\perp,$$

so that (observing that $F^\perp \subset \int\limits_T N_{\mathrm{dom}\, I_f}(x) \mathrm{d}\mu(t)$)

$$\begin{aligned}
x^* &\in (w) - \int\limits_T \left(\partial_{\eta(t)} f_t(x) + N_{\mathrm{dom}\, I_f}(x)\right) \mathrm{d}\mu(t) + V + V + F^\perp \\
&\subseteq (w) - \int\limits_T \left(\partial_{\eta(t)} f_t(x) + N_{\mathrm{dom}\, I_f}(x)\right) \mathrm{d}\mu(t) + V + V.
\end{aligned}$$

Hence, by intersecting over V we get

$$x^* \in \mathrm{cl}^{w^*}\left\{ (w)\text{-} \int\limits_T \left(\partial_{\eta(t)} f_t(x) + N_{\mathrm{dom}\, I_f}(x)\right) \mathrm{d}\mu(t) \right\},$$

which gives the desired inclusion due to the arbitrariness of function η. $\square$

5.5 Qualification conditions

Now we are going to present simplifications of Formulas 5.3 and 5.4 under classical conditions. The proof of the following two corollaries is similar, but it differs in the argument to present the existence of measurable selections.

Corollary 5.21 *Let X be Asplund, and assume that for every finite-dimensional subspace $F \subset X$, the function $f_{|_F} : T \times F \to \mathbb{R} \cup \{+\infty\}$ is a convex normal integrand, and assume that I_f has a continuity point. If $x \in X$ is such f_t is continuous at x for almost every t, then for every $\varepsilon \geq 0$*

$$\partial_\varepsilon I_f(x) = \bigcup_{\substack{\varepsilon_1+\varepsilon_2=\varepsilon\\ \varepsilon_1,\varepsilon_2\geq 0}} \bigcap_{\gamma>0} \mathrm{cl}^{w^*}\left(\bigcup_{\ell\in\mathcal{I}(\varepsilon_1+\gamma)} \left\{ (w) - \int_T \partial_{\ell(t)} f_t(x) \mathrm{d}\mu(t) \right\} \right) + N^{\varepsilon_2}_{\mathrm{dom}\, I_f}(x). \tag{5.13}$$

Moreover, if $\varepsilon = 0$, then

$$\partial I_f(x) = \mathrm{cl}^{w^*}\left(\left\{ (w) - \int_T \partial f_t(x) \mathrm{d}\mu(t) \right\} \right) + N_{\mathrm{dom}\, I_f}(x). \tag{5.14}$$

Proof. W.l.o.g. we suppose that $x = 0$ and $\mu(T) < +\infty$. We divide the proof in tree steps.

Step 1: We show in this step that for every $\ell \in \mathcal{I}(\varepsilon_1)$, $\varepsilon_1 \geq 0$, $t \rightrightarrows \partial_{\ell(t)} f_t(0)$ is a w^*-measurable multifunction with w^*-compact and convex values. Indeed, the continuity assumption of the f_t's ensures that the non-empty set $\partial_{\ell(t)} f_t(0)$ is w^*-compact and convex, as well as $\sigma_{\partial_{\ell(t)} f_t(0)}(u) = \inf\limits_{\lambda>0} \frac{f(t,0+\lambda u)-f(t,0)+\ell(t)}{\lambda}$ for all $u \in X$. Hence, the function $t \to \sigma_{\partial_{\ell(t)} f(t,0)}(u)$ is measurable, and so is the multifunction $t \rightrightarrows \partial_{\ell(t)} f_t(0)$.

Step 2: We have that $\partial_\varepsilon I_f(0) \subseteq \mathrm{cl}^{w^*}\left(\bigcup_{\substack{\varepsilon=\varepsilon_1+\varepsilon_2\\ \varepsilon_1,\varepsilon_2\geq 0\\ \ell\in\mathcal{I}(\varepsilon_1)}} \left\{ (w) - \int_T \partial_{\ell(t)} f_t(0) \mathrm{d}\mu(t) + N^{\varepsilon_2}_{\mathrm{dom}\, I_f\cap L}(0) \right\} \right)$ for every fixed $L \in \mathcal{F}(0)$. To prove this we take $x_0 \in \mathrm{int}(\mathrm{dom}\, I_f) \cap L$ pick $x^* \in \partial_\varepsilon I_f(0)$. By Theorem 5.9 and the continuity of $f(t,\cdot)$ there are $\varepsilon_1, \varepsilon_2 \geq 0$ with $\varepsilon = \varepsilon_1 + \varepsilon_2$ and $\ell \in \mathcal{I}(\varepsilon_1)$ such that

$$x^* \in \int_T \left(\partial_{\ell(t)} f_t(0) + \mathrm{span}\{\mathrm{dom}\, I_f \cap L\}^\perp \right) \mathrm{d}\mu(t) + N^{\varepsilon_2}_{\mathrm{dom}\, I_f\cap L}(0).$$

Hence, there exist an integrable function $x^*_L(t) \in \partial_{\ell(t)} f(t,0) + \mathrm{span}\{\mathrm{dom}\, I_f \cap L\}^\perp$ ae, and $\lambda^* \in N^{\varepsilon_2}_{\mathrm{dom}\, I_f\cap L}(0)$ such that $x^* = \int_T x^*_L \mathrm{d}\mu + \lambda^*$. Now, define the multifunction $G : T \rightrightarrows X^*$ as $G(t) := \{y^* \in \partial_{\ell(t)} f(t,0) : \langle y^*, x_0\rangle = \sigma_{\partial_{\ell(t)} f(t,0)}(x_0)\}$. By [20, Lemma 4.3] G is w^*-measurable multifunction (with w^*-compact convex valued), and so by Proposition 1.22 there exists a w^*-measurable selection $x^*(\cdot)$ of G; moreover, we have that $\langle x^*_L(t), x_0\rangle \leq \sigma_{\partial_{\ell(t)} f(t,0)+\mathrm{span}\{\mathrm{dom}\, I_f\cap L\}^\perp}(x_0) = \sigma_{\partial_{\ell(t)} f(t,0)}(x_0) = \langle x^*(t), x_0\rangle$. By the continuity of I_f we choose

$r > 0$ such $x_0 + \mathbb{B}(0, r) \subset \operatorname{int}(\operatorname{dom} I_f)$. Then, for every $v \in \mathbb{B}(0, r)$

$$\begin{aligned}\langle x^*(t), v\rangle \leq & f(t, x_0 + v) - f(t, 0) + \ell(t) - \langle x^*(t), x_0\rangle \\ = & f(t, x_0 + v) - f(t, 0) + \ell(t) - \sigma_{\partial_{\ell(t)} f(t,0)}(x_0) \\ \leq & f(t, x_0 + v) - f(t, 0) + \ell(t) - \langle x^*_L(t), x_0\rangle,\end{aligned}$$

and so,

$$\int_T |\langle x^*(t), v\rangle| \mathrm{d}\mu(t) \leq \int_T (\max\{f(t, x_0 + v), f(t, x_0 - v)\} - f(t, 0) + \ell(t) - \langle x^*_L(t), x_0\rangle) \, \mathrm{d}\mu(t) < +\infty;$$

that is, $x^*(\cdot)$ is Gelfand integrable (X is Banach). This last inequality implies that

$$C := \bigcup_{\substack{\varepsilon = \varepsilon_1 + \varepsilon_2 \\ \varepsilon_1, \varepsilon_2 \geq 0 \\ \ell \in \mathcal{I}(\varepsilon_1)}} \left\{ (w) - \int_T \partial_{\ell(t)} f_t(0) \mathrm{d}\mu(t) + N^{\varepsilon_2}_{\operatorname{dom} I_f \cap L}(0) \right\} \neq \emptyset.$$

Because $\langle x^*, x_0\rangle = \langle \int_T x^*_L(t)\mathrm{d}\mu(t), x_0\rangle + \langle \lambda^*, x_0\rangle \leq \langle \int_T x^*(t)\mathrm{d}\mu(t) + \lambda^*, x_0\rangle$, from the arbitrariness of $x^* \in \partial_\varepsilon I_f(0)$ and x_0 in $\operatorname{int}(\operatorname{dom} I_f) \cap L$ we get

$$\sigma_{\partial_\varepsilon I_f(0)}(x_0) \leq \sigma_C(x_0) \text{ for every } x_0 \in \operatorname{int}(\operatorname{dom} I_f) \cap L,$$

which also implies by usual arguments that $\sigma_{\partial_\varepsilon I_f(0)}(u) \leq \sigma_C(u)$ for all $u \in X$, and the desired relation holds.

Step 3: We complete the proof of the theorem. We show now that (5.13) and (5.14). We take $x^* \in \partial_\varepsilon I_f(0)$, so that by Step 2 there are nets of numbers $\varepsilon_{1,L,V}, \varepsilon_{2,L,V} \geq 0$ with $\varepsilon_{1,L,V} + \varepsilon_{2,L,V} = \varepsilon$ and $\ell_{L,V} \in \mathcal{I}(\varepsilon_{1,L,V})$ and vectors $x^*_{L,V} \in (w) - \int_T \partial_{\ell_{0L,V}(t)} f_t(0)\mathrm{d}\mu(t)$ and $\lambda^*_{L,V} \in N^{\varepsilon_{2,L,V}}_{\operatorname{dom} I_f \cap L}(0)$, indexed by $(L, V) \in \mathcal{F}(0) \times \mathcal{N}_0(w^*)$ such that $x^* = \lim x^*_{L,V} + \lambda^*_{L,V}$. We may assume that $\varepsilon_{1,L,V} \to \varepsilon_1$ and $\varepsilon_{2,L,V} \to \varepsilon_2$, with $\varepsilon_1 + \varepsilon_2 = \varepsilon$. Now, by the continuity of I_f at x_0 there exist some $r > 0$ such that for every $v \in B(0, r)$ and for every $L \ni x_0$ (w.l.o.g.)

$$\begin{aligned}\langle x^*_{L,V}, v\rangle \leq & I_f(x_0 + v) - I_f(0) + \varepsilon_{1,L,V} - \langle x^*_{L,V}, x_0\rangle \\ \leq & I_f(x_0 + v) - I_f(0) + \varepsilon_{1,L,V} - \langle x^*, x_0\rangle + \langle \lambda^*_{L,V}, x_0\rangle + 1 \\ \leq & I_f(x_0 + v) - I_f(0) + \varepsilon - \langle x^*, x_0\rangle + 1.\end{aligned}$$

Therefore, we may suppose that $(x^*_{L,V})$ w^*-converges to some $y^* \in X^*$ and that $(\lambda^*_{L,V})$ w^*-converges to some $\nu^* \in X^*$; hence, $\nu^* \in N^{\varepsilon_2}_{\operatorname{dom} I_f}(0)$. Finally, if $\varepsilon = 0$, then the conclusion follows. Otherwise, if $\varepsilon > 0$, for every $\gamma > 0$ we obtain that $\ell_{L,V} \in \mathcal{I}(\varepsilon_1 + \gamma)$ for a co-final family of index L, V, and so the desired inclusion follows. □

The next corllary is a generalization of [66], see also [50, 51, 78] for other versions. Before we need to make a remark on the relation between the continuity of I_f and the f_t's.

Remark 5.22 *Assume that either X is Suslin or $(T, \Sigma) = (\mathbb{N}, \mathcal{P}(\mathbb{N}))$. Due to the relation*

$$\operatorname{int}(\overline{\operatorname{dom} I_f}) \subset \operatorname{int}(\overline{\operatorname{dom} f_t}) \text{ for ae } t \in T,$$

the continuity hypothesis used in Corollary 5.23 is equivalent to the continuity of the functions I_f and $f(t, \cdot)$, $t \in T$, at some common point. In particular, in the finite-dimensional setting, the continuity of I_f alone ensures the continuity of the f_t's on the interior of their domains.

Corollary 5.23 *Assume in the setting of Theorem 5.19 that each one of the functions I_f and $f(t,\cdot)$, $t \in T$, is continuous at some point. Then for every $x \in X$ and $\varepsilon \geq 0$.*

$$\partial_\varepsilon I_f(x) = \bigcup_{\substack{\varepsilon=\varepsilon_1+\varepsilon_2\\ \varepsilon_1,\varepsilon_2\geq 0\\ \ell\in\mathcal{I}(\varepsilon_1)}} \left\{ (w) - \int\limits_T \partial_{\ell(t)} f_t(x) \mathrm{d}\mu(t) + N^{\varepsilon_2}_{\operatorname{dom} I_f}(x) \right\}.$$

Proof. We fix $x \in X$ and $\varepsilon \geq 0$, and choose a common continuity point x_0 of I_f and the f_t's (see Remark 5.22). The right-hand side is straightforwardly included in $\partial_\varepsilon I_f(x)$, and so we focus on the opposite inclusion. W.l.o.g. we may assume that $x = 0$, $\partial_\varepsilon \neq \emptyset$, $I_f(0) = 0$, as well as $\mu(T) = 1$. Take $x^* \in \partial_\varepsilon I_f(0)$, anf fix a sequence of positive functions $(\eta_n)_n \subset L^\infty(T,\mathbb{R})$ which converges to zero. By Theorem 5.19, there exists a net of integrable functions $w^*_{n,L,V}(t) \in \partial_{\ell_{n,L,V}(t)+\eta_n(t)} f_t(0) + N^{\varepsilon_{n,L,V,2}}_{\operatorname{dom} I_f\cap L}(0)$, with $n \in \mathbb{N}$, $L \in \mathcal{F}(0)$, $V \in \mathcal{N}_0$ and $\ell_{n,L,V} \in \mathcal{I}(\varepsilon_{n,L,V,1})$ such that

$$x^* = \lim_{n,L,V} \int_T w^*_{n,L,V}(t) \mathrm{d}\mu(t) \text{ and } \varepsilon_{n,L,V,1} + \varepsilon_{n,L,V,2} = \varepsilon. \tag{5.15}$$

Next, as in the proof of Theorem 5.19, we find measurable functions $x^*_{n,L,V}$ and $\lambda^*_{n,L,V}$ such that $x^*_{n,L,V}(t) \in \partial_{\eta_n(t)+\ell_{n,L,V}(t)} f(t,0)$ and $\lambda^*_{n,L,V}(t) \in N^{\varepsilon_{n,L,V,2}}_{\operatorname{dom} I_f\cap L}(0)$ for ae, with $w^*_{n,L,V}(t) = x^*_{n,L,V}(t) + \lambda^*_{n,L,V}(t)$. To simplify the notation, we just write $w^*_{n,\mathrm{i}}(\cdot)$, $x^*_{n,\mathrm{i}}(\cdot)$, $\varepsilon_{n,\mathrm{i},1}$, $\varepsilon_{n,\mathrm{i},2}$, $\lambda^*_{n,\mathrm{i}}(\cdot)$ and $\varepsilon_{n,\mathrm{i}} := \ell_{n,\mathrm{i}} + \eta_n$, $\mathrm{i} \in I := \mathcal{F}(0) \times \mathcal{N}_0$, where $\mathbb{N}\times I$ is endowed with the partial order "$\preceq$" given by $(n_1, L_1, V_1) \preceq (n_2, L_2, V_2)$ iff $n_1 \leq n_2$, $L_1 \subset L_2$ and $V_1 \supset V_2$.

The rest of the proof is devided into three steps.

Step 1: W.l.o.g. on n, i, there exists $U \in \mathcal{N}_0$ such that

$$m := \sup_{v\in U, n\in\mathbb{N}, \mathrm{i}\in I} \int\limits_T \langle x^*_{n,\mathrm{i}}(t), v\rangle \mathrm{d}\mu(t) < +\infty \text{ and } m_x := \sup_{n,\mathrm{i}} \int\limits_T |\langle x^*_{n,\mathrm{i}}(t), x\rangle| \mathrm{d}\mu(t) < +\infty \;\; \forall x \in X. \tag{5.16}$$

Indeed, we choose $U \in \mathcal{N}_0$ such that $\sup\limits_{v\in V} I_f(x_0 + v) < +\infty$. Then for every $n \in \mathbb{N}$, $\mathrm{i} \in I$ and $v \in U$

$$\begin{aligned}\langle x^*_{n,\mathrm{i}}(t), v\rangle &\leq f(t, x_0 + v) - f(t, 0) - \langle x^*_{n,\mathrm{i}}(t), x_0\rangle + \varepsilon_{n,\mathrm{i}}(t)\\ &= f(t, x_0 + v) - f(t, 0) - \langle w^*_{n,\mathrm{i}}(t) - \lambda^*_{n,\mathrm{i}}(t), x_0\rangle + \varepsilon_{n,\mathrm{i}}(t)\\ &\leq f(t, x_0 + v) - f(t, 0) - \langle w^*_{n,\mathrm{i}}(t), x_0\rangle + \varepsilon_{n,\mathrm{i},2} + \varepsilon_{n,\mathrm{i}}(t).\end{aligned} \tag{5.17}$$

But, by (5.15) and the definition of $\varepsilon_{n,\mathrm{i}}$ ($\varepsilon_{n,\mathrm{i}} = \ell_{n,\mathrm{i}} + \eta_n$), we may suppose that for all n and i,

$$-\int_T \langle w^*_{n,\mathrm{i}}(t), x_0\rangle + \int_T \varepsilon_{n,\mathrm{i}}(t)\mathrm{d}\mu(t) + \varepsilon_{n,\mathrm{i},2} \leq -\langle x^*, x_0\rangle + \varepsilon + \int_T \eta_n \mathrm{d}\mu + \frac{1}{2} \leq -\langle x^*, x_0\rangle + \varepsilon + 1, \tag{5.18}$$

and so (5.17) leads to $\sup\limits_{v\in U, n\in\mathbb{N}, \mathrm{i}\in I} \int_T \langle x^*_{n,\mathrm{i}}(t), v\rangle \mathrm{d}\mu(t) < +\infty$, which is the first part of (5.16). Now, we define the sets $T^+_{n,\mathrm{i},v} := \{t \in T : \langle x^*_{n,\mathrm{i}}(t), v\rangle \geq 0\}$, $T^-_{n,\mathrm{i},v} := \{t \in T : \langle x^*_{n,\mathrm{i}}(t), v\rangle < 0\}$,

$n \in \mathbb{N}$, $\mathrm{i} \in I$ and $v \in U$. Then, using (5.17),

$$\begin{aligned}
\int_T |\langle x^*_{n,\mathrm{i}}(t), v\rangle| \mathrm{d}\mu(t) &= \int_{T^+_{n,\mathrm{i},v}} \langle x^*_{n,\mathrm{i}}(t), v\rangle \mathrm{d}\mu(t) - \int_{T^-_{n,\mathrm{i},v}} \langle x^*_{n,\mathrm{i}}(t), v\rangle \mathrm{d}\mu(t) \\
&\leq \int_{T^+_{n,\mathrm{i},v}} \left\{ f(t, x_0+v) - f(t,0) - \langle w^*_{n,\mathrm{i}}(t), x_0\rangle + \varepsilon_{n,\mathrm{i},2} + \varepsilon_{n,\mathrm{i}} \right\} \mathrm{d}\mu(t) \\
&\quad + \int_{T^-_{n,\mathrm{i},v}} \left\{ f(t, x_0-v) - f(t,0) - \langle w^*_{n,\mathrm{i}}(t), x_0\rangle + \varepsilon_{n,\mathrm{i},2} + \varepsilon_{n,\mathrm{i}} \right\} \mathrm{d}\mu(t) \\
&= \int_{T^+_{n,\mathrm{i},v}} f(t, x_0+v) \mathrm{d}\mu(t) + \int_{T^-_{n,\mathrm{i},v}} f(t, x_0-v) \mathrm{d}\mu(t) \\
&\quad - \langle x^*, x_0\rangle + \varepsilon + 1 \quad \text{(by (5.18))} \\
&\leq \int_T |f(t, x_0+v)| \mathrm{d}\mu(t) + \int_T |f(t, x_0-v)| \mathrm{d}\mu(t) - \langle x^*, x_0\rangle + \varepsilon + 1 < +\infty,
\end{aligned}$$

and the second part in (5.16) follows since U is absorbent.

Step 2: There exist $\varepsilon_1, \varepsilon_2, \varepsilon_3 \geq 0$ with $\varepsilon_1 + \varepsilon_2 + \varepsilon_3 \leq \varepsilon$, neighborhood $U \in \mathcal{N}_0$, $\lambda_1^* \in N^{\varepsilon_3}_{\operatorname{dom} I_f}(0)$, linear functions $F_1 : X \to L^1(T, \mathbb{R})$ and $F_2 : X \to L^{\mathrm{sing}}(T, \mathbb{R})$, together with elements $\ell \in L^1(T, \mathbb{R}_+)$ and $s \in L^{\mathrm{sing}}(T, \mathbb{R})$ such that (w.l.o.g. on n and i):

(i) $\lambda_1^* = \lim_{n,\mathrm{i}} \int_T \lambda^*_{n,\mathrm{i}}(t) \mathrm{d}\mu(t)$.

(ii) For every $x \in X$, $(\langle x^*_{n,\mathrm{i}}(\cdot), x\rangle)_j \subset L^1(T, \mathbb{R}) \subset L^\infty(T, \mathbb{R})^*$ and $\langle x^*_{n,\mathrm{i}}(\cdot), x\rangle \to F_1(x) + F_2(x)$ wrt to the w^*-topology in $L^\infty(T, \mathbb{R})^*$.

(iii) $(\varepsilon_{n,\mathrm{i}}(\cdot)) \subset L^1(T, \mathbb{R}) \subset L^\infty(T, \mathbb{R})^*$ and $\varepsilon_{n,\mathrm{i}}(\cdot) \longrightarrow \ell + s$ wrt to the w^*-topology in $L^\infty(T, \mathbb{R})^*$.

(iv) $\sup\limits_{v \in U} \int\limits_T F_1(v) \mathrm{d}\mu(t) < +\infty$, $\sup\limits_{v \in U} F_2(v)(\mathbb{1}_T) < +\infty$.

(v) $\varepsilon_1 := \int_T \ell(t) \mathrm{d}\mu(t)$, $\varepsilon_2 := s(\mathbb{1}_T) \geq 0$.

(vi) For every $A \in \mathcal{A}$

$$\int_A F_1(x) \mathrm{d}\mu(t) \leq \int_A f(t,x) \mathrm{d}\mu(t) - \int_A f(t,0) \mathrm{d}\mu(t) + \int_A \ell(t) \mathrm{d}\mu(t), \quad \text{for all } x \in X, \tag{5.19}$$

$$F_2(x)(\mathbb{1}_T) \leq s(\mathbb{1}_T), \quad \text{for all } x \in \operatorname{dom} I_f. \tag{5.20}$$

Consider U, m and $(m_x)_{x \in X}$ as in the previous step, and denote by B the unit ball in the dual space of $L^\infty(T, \mathbb{R})$. From (5.16) and the definition of x^*, we obtain the existence of $\lambda_1^* \in X^*$ such that (w.l.o.g.) $\lambda_1^* = \lim_{n,\mathrm{i}} \int_T \left\langle \lambda^*_{n,\mathrm{i}}(t), \cdot \right\rangle \mathrm{d}\mu(t)$. Moreover, given an $x \in \operatorname{dom} I_f$ we write, since $\lambda^*_{n,\mathrm{i}}(t) \in N^{\varepsilon_{n,\mathrm{i},2}}_{\operatorname{dom} I_f \cap L}(0)$ and $x \in L$ (for L large enough),

$$\langle \lambda_1^*, x\rangle = \lim_{n,\mathrm{i}} \int_T \left\langle \lambda^*_{n,\mathrm{i}}(t), x \right\rangle \mathrm{d}\mu(t) \leq \lim_{n,\mathrm{i}} \varepsilon_{n,\mathrm{i},2} =: \varepsilon_3,$$

and so $\lambda_1^* \in N^{\varepsilon_3}_{\operatorname{dom} I_f}(0)$; hence, (i) follows. Next, by Thychonoff's theorem the space $\mathfrak{X} := \prod\limits_{x\in X}(m_x B, w^*((\mathrm{L}^\infty(T,\mathbb{R}))^*, \mathrm{L}^\infty(T,\mathbb{R})))$ is a compact space with respect to the product topology, and so w.l.o.g. we may assume that the net $(\langle x^*_{n,\mathrm{i}}(\cdot), x\rangle)_{x\in X} \in \mathfrak{X}$, $(n,\mathrm{i}) \in \mathbb{N}\times I$, converges to some $(F(x))_{x\in X}$, where $F : X \to (\mathrm{L}^\infty(T,\mathbb{R}))^*$ is linear function. Using the decomposition $(\mathrm{L}^\infty(T,\mathbb{R}))^* = \mathrm{L}^1(T,\mathbb{R}) \oplus \mathrm{L}^{\mathrm{sing}}(T,\mathbb{R})$, for every $x \in X$ we write $F(x) = F_1(x) + F_2(x)$, where $F_1 : X \to \mathrm{L}^1(T,\mathbb{R})$ and $F_2 : X \to \mathrm{L}^{\mathrm{sing}}(T,\mathbb{R})$ are two linear functions, and (ii) follows. Similarly, since $(\varepsilon_{n,\mathrm{i}}(\cdot))$ is bounded in $\mathrm{L}^1(T,\mathbb{R})$ we may assume that it converges to some $l + s$, with $l \in \mathrm{L}^1(T,\mathbb{R})$ and $s \in \mathrm{L}^{\mathrm{sing}}(T,\mathbb{R})$, such that for all $G \in \mathcal{A}$

$$\int_G \ell(t)\mathrm{d}\mu(t)+s(\mathbb{1}_G) = \lim_{n,\mathrm{i}} \int_G \varepsilon_{n,\mathrm{i}}(t)\mathrm{d}\mu(t) = \lim_{n,\mathrm{i}} \int_G (\ell_{n,\mathrm{i}}(t)+\eta_n(t))\mathrm{d}\mu(t) = \lim_{n,\mathrm{i}} \int_G \ell_{n,\mathrm{i}}(t)\mathrm{d}\mu(t) \le \varepsilon-\varepsilon_3. \tag{5.21}$$

Fix $x \in X$. Since $F_2(x)$, $s \in \mathrm{L}^{\mathrm{sing}}(T,\mathbb{R})$, there exists a sequence of measurable sets $T_n(x)$ such that $\mu(T\backslash \bigcup T_n(x)) = 0$ and

$$F_2(x)(g\mathbb{1}_{T_n(x)}) = 0, \;\; s(g\mathbb{1}_{T_n(x)}) = 0 \text{ for all } n \in \mathbb{N} \text{ and } g \in \mathrm{L}^\infty(T,\mathbb{R}).$$

Thus, by replacing in (5.21) the set G by $T_k(x)$, $k \ge 1$, and $T \setminus \cup_{1\le k\le n} T_k(x)$, respectively, and taking the limit on k, (iii) and (v) follow.

Now, for every $v \in U$, $G \in \mathcal{A}$, $n \in \mathbb{N}$ and $(n,\mathrm{i}) \in \mathbb{N}\times I$ (recall (5.16))

$$\int\limits_G \langle x^*_{n,\mathrm{i}}(t), x\rangle \mathrm{d}\mu(t) \le \int\limits_G f(t,x)\mathrm{d}\mu(t) - \int\limits_G f(t,0)\mathrm{d}\mu(t) + \int\limits_G \varepsilon_{n,\mathrm{i}}(t)\mathrm{d}\mu(t),$$

$$\int\limits_G \langle x^*_{n,\mathrm{i}}(t), v\rangle \mathrm{d}\mu(t) \le m.$$

So, by taking the limit we get

$$\int\limits_G F_1(x)\mathrm{d}\mu(t) + F_2(x)(\mathbb{1}_G) \le \int\limits_G f(t,x)\mathrm{d}\mu(t) - \int\limits_G f(t,0)\mathrm{d}\mu(t) + \int_G l(t)\mathrm{d}\mu(t) + s(\mathbb{1}_G), \tag{5.22}$$

$$\int\limits_G F_1(v)\mathrm{d}\mu(t) + F_2(v)(\mathbb{1}_G) \le m. \tag{5.23}$$

In particular, for $A \in \mathcal{A}$ and $G_n = A \cap T_n(x)$ we get

$$\int\limits_{G_n} F_1(x)\mathrm{d}\mu(t) \le \int\limits_{G_n} f(t,x)\mathrm{d}\mu(t) - \int\limits_{G_n} f(t,0)\mathrm{d}\mu(t) + \int\limits_{G_n} l(t)\mathrm{d}\mu(t),$$

$$\int\limits_{G_n} F_1(v)\mathrm{d}\mu(t) = \int\limits_{G_n} F_1(v)\mathrm{d}\mu(t) + F_2(v)(\mathbb{1}_{G_n}) \le m$$

which as $n \to \infty$ gives us

$$\int\limits_A F_1(v)\mathrm{d}\mu(t) \le m, \tag{5.24}$$

and

$$\int_A F_1(x)\mathrm{d}\mu(t) \le \int_A f(t,x)\mathrm{d}\mu(t) - \int_A f(t,0)\mathrm{d}\mu(t) + \int_A l(t)\mathrm{d}\mu(t),$$

yielding the first part in (5.19). Now, for $G_n = T\backslash \bigcup_{\mathrm{i}=1}^{n} T_{\mathrm{i}}(x)$ we have that $F_2(x)(\mathbb{1}_{G_n}) = F_2(x)(\mathbb{1}_T)$, $s(\mathbb{1}_{G_n}) = s(\mathbb{1}_T)$, and (for $v \in U$ and $x \in \operatorname{dom} I_f$)

$$\int_{G_n} F_1(v)\mathrm{d}\mu(t), \int_{G_n} F_1(x)\mathrm{d}\mu(t), \int_{G_n} f(t,x)\mathrm{d}\mu(t), \int_{G_n} f(t,0)\mathrm{d}\mu(t), \int_{G_n} l(t)\mathrm{d}\mu(t) \to_n 0,$$

and so, (5.22) and (5.23) yield

$$\begin{aligned} F_2(x)(\mathbb{1}_T) &\le s(\mathbb{1}_T), \\ F_2(v)(\mathbb{1}_T) &\le m, \end{aligned} \tag{5.25}$$

and we get ((vi)). Relation ((iv)) follows from (5.24) and (5.25).

Step 3: Let ℓ, s, $\varepsilon_1, \varepsilon_2$, ε_3, $U \in \mathcal{N}_0$, $\lambda_1^* \in N^{\varepsilon_3}_{\operatorname{dom} I_f}(0)$, F_1, and F_2 be as in step 2. We show the existence of a weakly integrable function $y^* : T \to X^*$ such that $y^*(t) \in \partial_{\ell(t)} f_t(0)$ ae, and $x^* - \int_T y^* \mathrm{d}\mu \in N^{\varepsilon_2}_{\operatorname{dom} I_f}(0)$.

Assume first that $(T, \mathcal{A}) = (\mathbb{N}, \mathcal{P}(\mathbb{N}))$. In this case, on the one hand we take $y^*(t) := F_1(\cdot)(t)$, $t \in T$. Then, for every $x \in X$, by (5.19) we have that for all $t \in T$

$$\langle y^*(t), x\rangle = F_1(x)(t) = (\mu(t))^{-1} \int_{\{t\}} F_1(x)\mathrm{d}\mu(t) \le f(t,x) - f(t,0) + \ell(t),$$

which, by taking into account the continuity assumption on $f(t,\cdot)$, shows that $y^*(t) \in \partial_{\ell(t)} f_t(0)$. Also, since that $\int_T |\langle y^*, x\rangle| \,\mathrm{d}\mu = \int_T |F_1(x)(t)| \,\mathrm{d}\mu(t) < +\infty$, for all $x \in X$, and (by (iv))

$$\sup_{v\in U} \int_T \langle y^*, v\rangle \,\mathrm{d}\mu = \sup_{v\in U} \int_T F_1(v)\mathrm{d}\mu(t) < +\infty,$$

it follows that $y^* := \int_T y^* \mathrm{d}\mu \in (w)\text{-}\int_T \partial_{\ell(t)} f_t(x)\mathrm{d}\mu(t)$. On the other hand, we take $\lambda^* := \lambda_1^* + \lambda_2^*$, with $\lambda_2^* := F_2(\cdot)(\mathbb{1}_T)$ ($\in X^*$, by (iv)), so that for all $x \in \operatorname{dom} I_f$ (using (5.20))

$$\langle \lambda_2^*, x\rangle = F_2(x)(\mathbb{1}_T) \le s(\mathbb{1}_T) = \varepsilon_2.$$

Hence, $\lambda^* \in N^{\varepsilon_3}_{\operatorname{dom} I_f}(0) + N^{\varepsilon_2}_{\operatorname{dom} I_f}(0) \subset N^{\varepsilon_2+\varepsilon_3}_{\operatorname{dom} I_f}(0)$, and so we get

$$\begin{aligned} x^* &= \lim_{n,L,V} \int_T (x^*_{n,L,V}(t) + \lambda^*_{n,L,V}(t))\mathrm{d}\mu(t) \\ &= \lim_{n,L,V} \int_T x^*_{n,L,V}(t)\mathrm{d}\mu(t) + \lim_{n,L,V} \lambda^*_{n,L,V}(t)\mathrm{d}\mu(t) \\ &= y^* + \lambda_2^* + \lambda_1^* \in (w) - \int_T \partial_{\ell(t)} f_t(x)\mathrm{d}\mu(t) + N^{\varepsilon_2+\varepsilon_3}_{\operatorname{dom} I_f}(0), \end{aligned}$$

which ensures the desired inclusion.

We treat now the case when X, X^* are Suslin spaces. We choose a countable set D such that $X = \{\lim x_n : (x_n)_{n\in\mathbb{N}} \subset D\}$. Equivalently, we can take an at most countable family of linearly independent vectors $\{\mathrm{e_i}\}_{\mathrm{i}=1}^{\infty}$ such that $L := \mathrm{span}\{\mathrm{e_i}\}_{\mathrm{i}=1}^{\infty} \supseteq D$ and $\mathrm{L}_n := \mathrm{span}\{\mathrm{e_i}\}_{\mathrm{i}=1}^{n} \ni x_0$ for all $n \geq 1$ (recall that x_0 is a common continuity point of I_f and the f_t's). As in the previous discrete case, we take $\lambda_2^* = F_2(\cdot)(\mathbb{1}_T)$. So, by argyuing as above we obtain that $\lambda^* := \lambda_1^* + \lambda_2^* \in N_{\mathrm{dom}\, I_f}^{\varepsilon_2+\varepsilon_3}(0)$. Next, we consider a sequence of functions $(b_n)_n$ such that each b_n is in the class of equivalence $F_1(\mathrm{e}_n)$, and define for every $t \in T$ a linear function $y_t^* : L \to \mathbb{R}$ as $\langle y_t^*, z\rangle = \sum_{\mathrm{i}=1}^{n} \alpha_{\mathrm{i}} b_{\mathrm{i}}(t)$ $(\in F_1(\sum_{\mathrm{i}=1}^{n} \alpha_{\mathrm{i}} \mathrm{e}_n)(t) = F_1(z)(t))$, where $z = \sum_{\mathrm{i}=1}^{n} \alpha_{\mathrm{i}} \mathrm{e_i}$, $\alpha_{\mathrm{i}} \in \mathbb{R}$. We notice that for every $z \in L$, $t \to \langle y_t^*, z\rangle$ is measurable. Now, given $z \in \mathrm{L}_{\mathbb{Q}} := \bigoplus_{\mathrm{i}=1}^{\infty} \mathbb{Q}\mathrm{e_i}$, we define $T_z := \{t \in T \mid \langle y_t^*, z\rangle \leq f(t,z) - f(t,0) + \ell(t)\}$ and $\tilde{T} := \bigcap_{z\in \mathrm{L}_{\mathbb{Q}}} T_z$; hence, $\mu(T\backslash\tilde{T}) = 0$, by equation (5.19). Now, because $\mathrm{int}\,\mathrm{dom}\, f_t \cap \mathrm{L}_{\mathbb{Q}} \neq \emptyset$ and f_t is continuous on $\mathrm{int}\,\mathrm{dom}\, f_t$, it follows that

$$\langle y_t^*, z\rangle \leq f(t,z) - f(t,0) + l(t) \quad \text{for all } t \in \tilde{T} \text{ and } z \in L;$$

in particular, y_t^* is a continuous linear functional on L for every $t \in \tilde{T}$. Now, by the Hahn-Banach theorem, we can extend y_t^* to a continuous linear functional on X, denoted by $y^*(t)$, such that

$$\langle y^*(t), z\rangle \leq f(t,z) - f(t,0) + l(t) \quad \text{for all } t \in \tilde{T} \text{ and } z \in L.$$

We notice that $y^*(\cdot)$ is weakly measurable, because for every $x \in X$, since $D \subset L$ there exists a sequence of element $x_n \in L$ such that $x_n \to x$ and, hence, $\langle y^*(t), x\rangle = \lim_n \langle w_t^*, x_n\rangle$ is measurable as is each function $t \to \langle w_t^*, x_n\rangle$. Moreover, by the continuity of f_t on $\mathrm{int}\,\mathrm{dom}\, f_t$, the last inequality above holds on X, and this gives us $y^*(t) \in \partial_{l(t)} f_t(0)$ for all $t \in \tilde{T}$. Now, by the continuity of I_f and using similar arguments as those in the proof of Theorem 5.20 or Corollary 5.21, it is not difficult to show that $y^*(t)$ is weakly integrable and that $y^* := \int_T y^*(t)\mu(t)$ defines a continuous linear operator on X, showing that $y^* \in (w)\text{-}\int_T \partial_{\ell(t)} f_t(x) \mathrm{d}\mu(t)$. Whence $x^* = y^* + \lambda^* \in (w)\text{-}\int_T \partial_{\ell(t)} f_t(x)\mathrm{d}\mu(t) + N_{\mathrm{dom}\, I_f}^{\varepsilon_2+\varepsilon_3}(0)$. The proof of the corollary is finished.

□

Corollary 5.24 *Assume that X is a separable Banach space. If $\hat{I}_f$ is bounded above on some neighborhood with respect to $(\mathrm{L}^{\infty}(T,X), \|\cdot\|)$ of some constant function $x_0(\cdot) \equiv x_0 \in X$, then for all $x \in X$ and $\varepsilon \geq 0$ we have that*

$$\partial_\varepsilon I_f(x) = \bigcup_{\substack{\varepsilon=\varepsilon_1+\varepsilon_2 \\ \varepsilon_1,\varepsilon_2\geq 0 \\ \ell\in\mathcal{I}(\varepsilon_1)}} \left\{ \int_T x^* \mathrm{d}\mu : x^* \in \mathrm{L}^1_{w^*}(T, X^*),\ x^*(t) \in \partial_{\ell(t)} f_t(x) \ ae \right\} + N_{\mathrm{dom}\, I_f}^{\varepsilon_2}(x).$$

If, in addition, X is reflexive, then

$$\partial_\varepsilon I_f(x) = \bigcup_{\substack{\varepsilon=\varepsilon_1+\varepsilon_2 \\ \varepsilon_1,\varepsilon_2\geq 0 \\ \ell\in\mathcal{I}(\varepsilon_1)}} \int_T \partial_{\ell(t)} f_t(x)\mathrm{d}\mu(t) + N_{\mathrm{dom}\, I_f}^{\varepsilon_2}(x).$$

Proof. First, observe that the current continuity assumption of $\hat{I}_f$ implies that both I_f and the functions f_t, for ae $t \in T$, are continuous at x_0. Then, according to Corollary 5.23, to prove the first part we only need to verify that, for every $x \in \operatorname{dom} I_f$, $\varepsilon_1 \geq 0$ and $\ell \in \mathcal{I}(\varepsilon_1)$,

$$(w) - \int_T \partial_{\ell(t)} f_t(x) \mathrm{d}\mu(t) \subset \left\{ \int_T x^* \mathrm{d}\mu : x^* \in \mathrm{L}^1_{w^*}(T, X^*),\ x^*(t) \in \partial_{\ell(t)} f_t(x) \text{ ae} \right\}. \tag{5.26}$$

Take $x^* := \int_T x^* \mathrm{d}\mu$ for a w^*-measurable function $x^*(\cdot)$ such that $x^*(t) \in \partial_{\ell(t)} f_t(x)$ ae. By Proposition 1.21, for each function $\alpha \in \mathrm{L}^1(T, \mathbb{R}_+)$ there exists a measurable function $x : T \to X$ such that $\|x(t)\| \leq 1$ and

$$\langle x^*(t), x(t) \rangle \geq \|x^*(t)\| - \alpha(t) \text{ ae};$$

hence, $\int_T \|x^*(t)\| \mathrm{d}\mu(t) \leq \int_T \langle x^*(t), x(t) \rangle \mathrm{d}\mu(t) + \int_T \alpha(t) \mathrm{d}\mu(t)$. Since, by the continuity of $\hat{I}_f$ at x_0 there are $\delta, M > 0$ such that

$$\begin{aligned} \delta \int_T \langle x^*(t), x(t) \rangle \mathrm{d}\mu(t) &\leq \int_T f_t(\delta x(t) + x_0) \mathrm{d}\mu(t) - I_f(x) + \int_T \langle x^*(t), x - x_0 \rangle \mathrm{d}\mu(t) + \varepsilon_1 \\ &\leq I_f(x_0) - I_f(x) + 1 + \int_T \langle x^*(t), x - x_0 \rangle \mathrm{d}\mu(t) + \varepsilon_1 \leq M, \end{aligned}$$

we obtain that $\int_T \|x^*(t)\| \mathrm{d}\mu(t) \leq \delta^{-1} M + \int_T \alpha(t) \mathrm{d}\mu(t) < +\infty$, and (5.26) holds.

Finally, the last statement follows because $\mathrm{L}^1_{w^*}(T, X^*) = \mathrm{L}^1(T, X^*)$. □

The next example shows that the second formula of Corollary 5.24 could not be valid if we drop the continuity of $\hat{I}_f$.

Example 5.25 Consider $(T, \mathcal{A}) = (\mathbb{N}, \mathcal{P}(\mathbb{N}))$ and $X = \ell^2$, and let $(\mathrm{e}_n)_n$ be the canonical basis of ℓ^2, and μ be the finite measure given by $\mu(\{n\}) = 2^{-n}$. Define the integrand $f : \mathbb{N} \times \ell^2 \to \mathbb{R}$ as $f(n, x) := 2^n (\langle \mathrm{e}_n, x \rangle)^2$, so that $I_f(x) = \int_{\mathbb{N}} f(n, x) \mathrm{d}\mu(n) = \sum_{n \in \mathbb{N}} x_n^2 = \|x\|^2$. Then I_f is differentiable on X with $\nabla I_f(x) = 2x$, and for all $n \geq 1$ we have that $\partial f_n(x) = \{\nabla f_n(x)\} = \{2^{n+1} \langle x, \mathrm{e}_n \rangle \mathrm{e}_n\}$, so that

$$\nabla I_f(x) = \int_{\mathbb{N}} 2^{n+1} \langle x, \mathrm{e}_n \rangle \mathrm{e}_n \mathrm{d}\mu(n) = 2 \sum_{n \in \mathbb{N}} \langle x, \mathrm{e}_n \rangle \mathrm{e}_n = 2x,$$

which is the result of Corollary 5.23. On the other side, the value $\int_{n \in \mathbb{N}} \|\nabla f_n(x)\| \mathrm{d}\mu(n) = 2 \sum_{n \in \mathbb{N}} |\langle x, \mathrm{e}_n \rangle|$ could not be finite for all $x \in \ell^2$ (consider, for instance, $x = (1/n)_{n \geq 1}$), which means that $(2^{n+1} \langle x, \mathrm{e}_n \rangle \mathrm{e}_n)_n \notin \mathrm{L}^1(\mathbb{N}, \ell^2)$.

The following is an easy consequence of Corollaries 5.21 and 5.23:

Corollary 5.26 *Assume that either X is Asplund, X and X^* are Suslin, or $(T, \mathcal{A}) = (\mathbb{N}, \mathcal{P}(\mathbb{N}))$. If $x \in X$ is a common continuity point of both I_f and the f_t's, then I_f is Gâteaux-differentiable at x if and only if f_t is Gâteaux-differentiable at x for ae $t \in T$.*

The last result was given [121, Corollary 2.11] when $(T,\mathcal{A}) = (\mathbb{N},\mathcal{P}(\mathbb{N}))$. Concerning the Fréchet-differentiability, in the same referred result the authors proved one implication (the Fréchet-differentiability of the sum implies the one of the data functions), and let the other implication as an open problem (see [121, Question 2.12, page 1146]). The following example answers this question in the negative.

Example 5.27 Consider $(T,\mathcal{A}) = (\mathbb{N},\mathcal{P}(\mathbb{N}))$ and $X = \ell^1$, and let $(\mathrm{e}_n)_n$ be the canonical basis of ℓ^1, and μ be the finite measure given by $\mu(\{n\}) = 1$. Define the integrand $f : \mathbb{N}\times\ell^1 \to \mathbb{R}$ as $f(n,x) := |\langle \mathrm{e}_n, x\rangle|^{1+1/n}$, so that $I_f(x) = \sum |\langle \mathrm{e}_n, x\rangle|^{1+1/n} < +\infty$. Since each f_n is a Fréchet-differentiable convex function such that $\nabla f_n(x) = (1+\frac{1}{n})|\langle x, \mathrm{e}_n\rangle|^{1/n}\mathrm{e}_n$, according to Corollary 5.26 I_f is Gâteaux-differentiable on ℓ^1, with Gâteaux-differential equal to $\sum \nabla f_n(x) := \int_{\mathbb{N}} \nabla f_n(x)\mathrm{d}\mu(n) = \sum(1+\frac{1}{n})|\langle x,\mathrm{e}_n\rangle|^{1/n}\mathrm{e}_n$ (by Corollary 5.14). Thus, if I_f would be Fréchet-differentiable at $x = 0$, then we would have

$$\frac{I_f(n^{-1}\mathrm{e}_n) - I_f(0) - n^{-1}\langle \nabla I_f(0), \mathrm{e}_n\rangle}{n^{-1}} = nn^{-1-\frac{1}{n}} = n^{-\frac{1}{n}} \to 1,$$

which is a contradiction.

Remark 5.28 In the previous Corollary we only use the separability in step 3 (step 2 also can be deduced from Corollary 5.14 and using the hypothesis *for every finite dimensional space F of X, $f_{|_F} : T\times F \to \overline{\mathbb{R}}$ is a convex normal integrand* instead of f is a convex normal integrand), so step 2 gives a generalization of [66, Theorem 2]. Indeed by Corollary 5.14 we have for every $L \in \mathcal{F}(0)$, $x^* = \int_T y_L^* + \lambda_L^*$, where y_L^* is an integrable selection of $\partial f_t(0) + L^\perp$ and $\lambda_L^* \in \mathrm{N}_{\operatorname{dom} I_f\cap L}(0)$, so for every $L \in \mathcal{F}(0)$, there are (not necessary measurable) functions $w_L^*(t) \in \partial f_t(0)$ and $\beta_L^*(t) \in L^\perp$. So we notice that for every $u \in L$, $\langle y^*(t), u\rangle = \langle w_L^*(t), u\rangle$ ae, which implies, $\langle w_L^*(\cdot), u\rangle$ is measurable for every $u \in L$. Then similar estimation to Step 1 are valid. To proceed with step 2, we must be cautious because $\langle w_L^*(\cdot), u\rangle$ belongs to $\mathrm{L}^1(T,\mathbb{R})$ only for $u \in L$, so one idea to skip the inconvenience is define a net in $\mathfrak{X}$ as following, $\langle z_L^*(t), x\rangle = \langle w_L^*(\cdot), u\rangle$ if $u \in L$ and $\langle z_L^*(t), x\rangle = 0$ otherwise, with this the conclusion of step 2 follows from the same argument. Furthermore, in [86] the authors use the following definitions to extend results of the integral representation of the Clarke and Limiting/Mordukhovich sub-differential in nonseparable Banach spaces, we extracted this lines from the same article. *Denote by $\mathcal{N}(\mu)$ the null ideal of $\mathcal{A}$, i.e., $\mathcal{N}(\mu) = \{\, N \in \mathcal{A} : \mu(N) = 0\}$. A measure μ on σ is κ-additive if for every pairwise disjoint family $\mathcal{E} \subseteq \mathcal{A}$ with $|\mathcal{E}| < \kappa$ we have $\bigcup\mathcal{E} \in \mathcal{A}$ and $\mu(\bigcup\mathcal{E}) = \sum\limits_{A\in\mathcal{E}} \mu(A) := \sup\limits_{\mathcal{O}\subseteq\mathcal{E}} \sum\limits_{A\in\mathcal{O}} \mu(A)$. The additivity $\kappa(\mu)$ of μ is the largest cardinal of κ for which μ is κ-additive, or it is ∞ if μ is κ-additive for every κ. We denote* $\operatorname{den}(X)$ *the density of the Banach space X, i.e., the smallest cardinal of the form $|D|$, where D is a dense subset of X. A useful representation is* $\kappa(\mu) = \min\{|\mathcal{E}| : \mathcal{E} \subseteq \mathcal{N}, \bigcup\mathcal{E} \notin \mathcal{A}\}$. In the context of a pair (X, X^*) of locally convex spaces we can adapt the previous definition as follow. We define the relation seq-den$(X, w) < \kappa$ if and only if there exists a set $D \subseteq X$ satisfies that $|D| < \kappa$ and for every $x \in X$ there exists a generalized sequence $(x_{\mathrm{i}})_{\mathrm{i}\in I} \subseteq D$ such that $|I| < \kappa$ and $x_{\mathrm{i}} \xrightarrow{w} x$. With this Hypothesis the proof of Step 3 (in Theorem 5.23) is still valid (without the separability assumption and using the hypothesis *for every finite dimensional space F of X, $f_{|_F} : T\times F \to \overline{\mathbb{R}}$ is a convex normal integrand* instead of f is a convex normal integrand) outside of a separable space. Indeed in step 3 we have to choose a set D with the above density property and then define, as in Theorem 5.23, $L_{\mathbb{Q}} = \bigoplus\limits_{\mathrm{e}\in D} \mathbb{Q}\mathrm{e}$, hence the proof of step

3 follows similarly.

In the following, we extend Corollaries 5.21 and 5.23 to the nonconvex Lipschitz case. The resulting results are known for both the case of a separable Banach spaces or $(T, \mathcal{A}) = (\mathbb{N}, \mathcal{P}(\mathbb{N}))$ ([23, Theorem 2.7.2]), and the case of Asplund spaces ([86]).

Proposition 5.29 (Clarke-Murdokovich-Sagara) *Assume that either X is Asplund, X is a separable Banach space, or $(T, \mathcal{A}) = (\mathbb{N}, \mathcal{P}(\mathbb{N}))$. Let integrand $f : T \times X \to \overline{\mathbb{R}}$ and $x \in X$ be such that:*

(a) There exists $K \in \mathrm{L}^1(T, \mathbb{R}_+)$ and $\delta > 0$ such that for every $y, z \in B(x, \delta)$, $t \to f(t, y)$ is measurable and $|f_t(y) - f_t(z)| \leq K(t)\|y - z\|$ ae $t \in T$.

(b) (When X is Asplund) For every $u \in X$, the function $t \to f_t^\circ(x; u)$ is measurable. Then we have that

$$\partial_C I_f(x) \;\subseteq\; \mathrm{cl}^{w^*}\left((w) - \int\limits_T \partial_C f_t(x) \mathrm{d}\mu(t) \right).$$

Where the closure operator can be omitted if $\mathrm{den}(X) < \kappa(\mu)$ *(see Remark 5.28).*

Proof. By taking into account Fatou's lemma, we have that $I_f^\circ(x; u) \leq \int_T f_t^\circ(x; u)\mathrm{d}\mu(t)$ for all $u \in X$, and so $\partial_C I_f(x) = \partial I_f^\circ(x; 0) \subset \partial I_{f^\circ(x;\cdot)}(0)$. Since $(t, u) \to f_t^\circ(x; u)$ is a Carathéodory map, for every finite-dimensional subspace $F \subset X$, the mapping $(f_\cdot^\circ(x; \cdot))_{|F} : T \times F \to \mathbb{R} \cup \{+\infty\}$ is a convex normal integrand. Hence, because $I_{f^\circ(x;\cdot)}$ and $f_t^\circ(x; \cdot)$ are continuous everywhere, the two desired formulas follow by applying Corollaries 5.21 and 5.23, respectively. □

The final result correspond to the closeness of the integral of a multifunction with closed convex values.

Corollary 5.30 (Measurable selection and closeness of the integral of multifunction) *Let $C : T \rightrightarrows X^*$ uniformly integrable w^*-measurable multifunction (i.e., there exists an integrable function $K \in \mathrm{L}^1(T, \mathbb{R})$ and a continuous seminorm $\rho : X \to \mathbb{R}$ such that for all $t \in T$ and all $x \in X$ $\sigma_{C(t)}(x) \leq K(t)\rho(x)$) with convex and w^*-closed values. If den$(X) < k(\mu)$, then $(w) - \int\limits_T C(t)\mathrm{d}\mu(t)$ is non-empty and w^*-closed.*

Proof. Take the integrand $f(t, x) := \sigma_{C(t)}(x)$, then for every finite dimensional space F of X, $f_{|F} : T \times F \to \overline{\mathbb{R}}$ is a convex normal integrand. So by the Hypotheses the integrand function I_f, $f_t(\cdot)$ are continuous (which implies in particular $\partial I_f(0) \neq \emptyset$). Then by Theorem 5.23 and Remark 5.28 we have that $\partial I_f(0) = (w) - \int\limits_T \partial f(t, 0)\mathrm{d}\mu(t) = (w) - \int\limits_T C(t)\mathrm{d}\mu(t)$. □

5.6 Approach using conjugate functions

We investigate now the representation of the ε-normal set to $\operatorname{dom} I_f$ in terms of the data functions f_t. We suppose that $f : T \times X \to \overline{\mathbb{R}}$ is a normal integrand defined on a locally convex space X, such that for some $x_0^* \in L^1_{w^*}(T, X^*)$ and $\alpha \in L^1(T, \mathbb{R})$ it holds

$$f(t,x) \geq \langle x_0^*(t), x\rangle + \alpha(t) \text{ for all } x \in X \text{ and } t \in T. \tag{5.27}$$

In what follows, we suppose that either X, X^* are Suslin or $(T, \mathcal{A}) = (\mathbb{N}, \mathcal{P}(\mathbb{N}))$. We recall that the continuous infimal convolution of the f_t^*'s is the function $\oplus \int_T f^*(t,\cdot)\mathrm{d}\mu(t) : X^* \to \overline{\mathbb{R}}$ given by (see)

$$\begin{aligned}\left(\oplus \int_T f^*(t,\cdot)\mathrm{d}\mu(t)\right)(x^*) \quad &: = \oplus \int_T f^*(t,x^*)\mathrm{d}\mu(t) \\ &: = \inf\{ \int_T f^*(t,x^*(t))\mathrm{d}\mu(t) \mid \begin{array}{c} x^*(\cdot) \in L^1_{w^*}(T,X^*) \\ \text{and } \int_T x^*(t)\mathrm{d}\mu(t) = x^* \end{array} \},\end{aligned}$$

with the convention that $\inf_\emptyset := +\infty$. We also recall the notation $\mathrm{cl}^{w^*}(h)$, which refers to the w^*-closure of a function $h : X^* \to \mathbb{R} \cup \{+\infty\}$. We shall need the following Lemma, which can be found in [56, Lemma 1.1].

To finish this section we investigate how represent $N_{\operatorname{dom} I_f}(x)$ in terms of the data functions f_t. An straightforward application of Theorem 5.2 give us the representation of the ε-subdifferential of I_f and $\hat{I}_f$. For this propose we need the following Lemma, which can be found in [56, Lemma 1.1].

Lemma 5.31 *Let $h : X^* \to \overline{\mathbb{R}}$ be a convex function. Then for all $r \in \mathbb{R}$*

$$\{x^* \in X^* : \mathrm{cl}^{w^*}(h)(x^*) \leq r\} = \bigcap_{\delta > 0} \mathrm{cl}\{x^* \in X^* : h(x^*) < r + \delta\}.$$

Moreover, if $r > \inf_{X^*} h$, then

$$\{x^* \in X^* : \mathrm{cl}^{w^*}(h)(x^*) \leq r\} = \mathrm{cl}\{x^* \in X^* : h(x^*) < r\}.$$

Theorem 5.32 *If $\overline{\mathrm{co}} I_f = I_{\overline{\mathrm{co}} f}$, then*

$$(I_f)^*(x^*) = \mathrm{cl}^{w^*}\left(\oplus \int_T f^*(t,\cdot)\mathrm{d}\mu(t)\right)(x^*), \text{ for all } x^* \in X^*,$$

and, for all $x \in X$, and $\varepsilon \geq 0$,

$$\partial_\varepsilon I_f(x) = \bigcap_{\varepsilon_1 > \varepsilon} \mathrm{cl}^{w^*}\left(\bigcup_{\ell \in \mathcal{I}(\varepsilon_1)} \left\{\int_T x^* \mathrm{d}\mu : x^* \in \mathrm{L}^1_{w^*}(T, X^*),\ x^*(t) \in \partial_{\ell(t)} f_t(x) \ ae\right\}\right). \tag{5.28}$$

In addition, if $\varepsilon > I_f(x) - \overline{\text{co}} I_f$*, then*

$$\partial_\varepsilon I_f(x) = \text{cl}^{w^*}\left(\bigcup_{\ell\in\mathcal{I}(\varepsilon)}\left\{\int_T x^*\text{d}\mu : x^* \in \text{L}^1_{w^*}(T,X^*),\ x^*(t)\in\partial_{\ell(t)}f_t(x)\ ae\right\}\right).$$

Proof. It easy to see that the space $L^1_{w^*}(T,X^*)$ is decomposable and that f^* is a convex integrand function such that $f^*(t,x_0^*(t)) \leq -\alpha(t)$ ae, by (5.27). As well, denoting $\varphi := \oplus\int_T f^*(t,\cdot)\text{d}\mu(t)$, we verify that, for all $x\in X$,

$$\begin{aligned}\varphi^*(x) &= \sup_{\lambda^*\in X^*}\sup_{x^*\in L^1_{w^*}(T,X^*),\ \int_T x^*\text{d}\mu=\lambda^*}\{\langle\lambda^*,x\rangle - \int_T f^*(t,x^*(t))\text{d}\mu(t)\}\\ &= \sup_{x^*\in L^1_{w^*}(T,X^*)}\int_T(\langle x^*(t),x\rangle - f^*(t,x^*(t))\text{d}\mu(t),\end{aligned}$$

and, so, according to Theorem 5.2 and Moreau's envelope theorem,

$$\varphi^*(x) = I_{f^{**}}(x) = I_{\overline{\text{co}}f}(x) = \overline{\text{co}}I_f.$$

Consequently, the convexity of the continuous infimal convolution yields

$$\text{cl}^{w^*}(\varphi) = \varphi^{**} = (\overline{\text{co}}I_f)^* = (I_f)^*. \tag{5.29}$$

Now, we assume that $\partial_\varepsilon I_f(x)\neq\emptyset$. Then, for all $\varepsilon_1 > \varepsilon$ one has (by (5.29)) that

$$\inf_{x^*\in X^*}\{\varphi(x^*) - \langle x^*,x\rangle + I_f(x)\} = I_f(x) - (\overline{\text{co}}I_f)(x) \leq \varepsilon < \varepsilon_1,$$

and so, using Lemma 5.31,

$$\begin{aligned}\partial_{\varepsilon_1}I_f(x) &= \left\{x^*\in X^* : \text{cl}^{w^*}(\varphi)(x^*) + I_f(x) - \langle x^*,x\rangle \leq \varepsilon_1\right\}\\ &= \left\{x^*\in X^* : \text{cl}^{w^*}(\varphi - x)(x^*) + I_f(x) \leq \varepsilon_1\right\}\\ &= \text{cl}^{w^*}\left\{x^*\in X^* : \varphi(x^*) + I_f(x) < \langle x^*,x\rangle + \varepsilon_1\right\}.\end{aligned}$$

Observe that for each $x^*\in X^*$ satisfying $\varphi(x^*) + I_f(x) < \langle x^*,x\rangle + \varepsilon_1$, there exists some $x^*(\cdot)\in L^1_{w^*}(T,X^*)$ such that $\int_T f(t,x)\text{d}\mu(t) + \int f^*(t,x^*(t)) \leq \int\langle x^*(t),x\rangle\text{d}\mu(t) + \varepsilon_1$. Hence, by defining $\ell(t) := f(t,x) + f^*(t,x^*(t)) - \langle x^*(t),x\rangle$; hence, $\ell\in\mathcal{I}(\varepsilon_1)$ and $x^*(t)\in\partial_{\ell(t)}f(t,x)$ ae. The proof is finished since the other inclusion is straightforward. □

Remark 5.33 It is important to mention that result of Proposition above also holds when the space X is changed by a space of weakly integrable functions from T to X such that for every $(u,v)\in L(T,X)\times L(T,X^*)$ the scalar function $t\to\langle v(t),u(t)\rangle$ is integrable and the linear function $x^*\in LD(T,X^*)\to z^*(\cdot) := \int_T\langle x^*(t),\cdot\rangle\text{d}\mu(t)\in L(T,X)^*$ is well defined, where $LD(T,X^*)$ is a decomposable space of weakly integrable functions. In this case the (5.28) is changed by:

$$\partial_\varepsilon I_f(x) = \bigcap_{\delta>\varepsilon}\text{cl}\left\{\bigcup_{\ell\in\mathcal{I}(\delta)}\int_T\partial_{\ell(t)}f(t,x)\text{d}\mu(t)\right\}.$$

Where

$$\int_T \partial_{\ell(t)} f(t,x(t))\mathrm{d}\mu(t) = \{\int_T \langle x^*(t), \cdot\rangle \mathrm{d}\mu(t) : x^*(t) \in \partial_{\ell(t)} f(t,x(t)) \text{ ae and } x^* \in LD(T,X^*)\}.$$

Moreover, $(I_{\overline{\mathrm{co}}\, f})^*(x^*) = \overline{\varphi}(x^*)$, where

$$\varphi(x^*) := \inf\{\int_T f^*(t, v^*(t))\mathrm{d}\mu(t) \mid v^* \in LD(T,X^*) \text{ and } \int_T v^* = x^*\}.$$

We also obtain a characterization of the epigraph of the function I_f^* :

Corollary 5.34 *Assume that I_f^* is proper. If $\overline{\mathrm{co}} I_f = I_{\overline{\mathrm{co}} f}$, then we have that*

$$\operatorname{epi} I_f^* = \mathrm{cl}^{w^*}\left\{ \left(\int_T x^*\mathrm{d}\mu, \int_T \alpha\mathrm{d}\mu\right) : \begin{array}{c} x^* \in \mathrm{L}^1_{w^*}(T,X^*), \\ \alpha \in \mathrm{L}^1(T,\mathbb{R}), \\ (x^*(t),\alpha(t)) \in \operatorname{epi} f_t^* \ ae \end{array} \right\}.$$

Proof. We denote E the set between parenthesis in the equation above. Take $(x^*,\alpha) \in E$. Then, using again the notation $\varphi := \oplus \int_T f^*(t,\cdot)\mathrm{d}\mu(t)$, we obtain that $\varphi(x^*) \le \alpha$, and so by Theorem 5.32 $(x^*,\alpha) \in \operatorname{epi} I_f^*$. Hence, the lower semicontinuity of I_f^* yields the inclusion $\mathrm{cl}^{w^*}(E) \subset \operatorname{epi} I_f^*$. To prove the other inclusion, we take $(x^*,\alpha) \in \operatorname{epi} I_f^*$, and fix $\varepsilon > 0$ and $V \in \mathcal{N}_{x^*}(w^*)$ together with $\gamma(\cdot) \in \mathrm{L}^1(T,\mathbb{R}_+)$ such that $\int_T \gamma \mathrm{d}\mu = 1$. Then by Theorem 5.32 there exists $x^*(\cdot) \in \mathrm{L}^1_{w^*}(T,X^*)$ such that (w.l.o.g.) $\int_T x^*\mathrm{d}\mu \in V$ and

$$-\infty < (I_f)^*(x^*) - 1 = \mathrm{cl}^{w^*}(\varphi)(x^*) - 1 \le \varphi\left(\int_T x^*\mathrm{d}\mu\right) \le \int_T f^*(t,x^*(t))\mathrm{d}\mu(t) \le \alpha + \varepsilon$$

Thus, if we denote $\beta(t) := f^*(t,x^*(t)) + \gamma(t)\left(\alpha + \varepsilon - \int_T f^*(t,x^*(t))\mathrm{d}\mu(t)\right)$, then $\int_T \beta\mathrm{d}\mu = \alpha + \varepsilon$ and we get $(x^*(t),\beta(t)) \in \operatorname{epi} f_t$ and $(\int_T x^*\mathrm{d}\mu, \int_T \beta\mathrm{d}\mu) \in E$. Thus, from the arbitrariness of $\varepsilon > 0$ and V we deduce that $(x^*,\alpha) \in \mathrm{cl}^{w^*}(E)$. □

Remark 5.35 It is worth mention that if (5.27) is satisfied with some $x^* \in \mathrm{L}^p_{w^*}(T,X^*)$ (resp. $x^* \in \mathrm{L}^p(T,X^*)$) and $\alpha \in \mathrm{L}^s(T,\mathbb{R})$, then the epigraph of I_f^* can be expressed as

$$\operatorname{epi} I_f^* = \mathrm{cl}^{w^*}\left\{ \left(\int_T x^*\mathrm{d}\mu, \int_T \alpha\mathrm{d}\mu\right) : \begin{array}{c} x^* \in \mathrm{L}^p_{w^*}(T,X^*), \\ \alpha \in \mathrm{L}^s(T,\mathbb{R}), \\ (x^*(t),\alpha(t)) \in \operatorname{epi} f_t^* \text{ ae} \end{array} \right\}$$

$$\left(\text{resp. } = \mathrm{cl}^{w^*}\left\{ \left(\int_T x^*\mathrm{d}\mu, \int_T \alpha\mathrm{d}\mu\right) : \begin{array}{c} x^* \in \mathrm{L}^p(T,X^*), \\ \alpha \in \mathrm{L}^s(T,\mathbb{R}), \\ (x^*(t),\alpha(t)) \in \operatorname{epi} f_t^* \text{ ae} \end{array} \right\}.\right.$$

We are now in position to give the desired representation of the ε-normal set to $\operatorname{dom} I_f$.

Proposition 5.36 *Assume that f is convex normal integrand. Then for every $x \in \operatorname{dom} I_f$ and $\varepsilon \geq 0$ we have that*

$$\begin{aligned} N^{\varepsilon}_{\operatorname{dom} I_f}(x) &= \{x^* \in X^* : (x^*, \langle x^*, x\rangle + \varepsilon) \in \operatorname{epi}(\sigma_{\operatorname{dom} I_f})\} \\ &= \{x^* \in X^* : (x^*, \langle x^*, x\rangle + \varepsilon) \in (\operatorname{epi}(I_f)^*)_\infty\} \\ &= \left\{x^* \in X^* : (x^*, \langle x^*, x\rangle + \varepsilon) \in \left[\operatorname{cl}^{w^*} \mathcal{E}\right]_\infty\right\} \\ &= \left\{x^* \in X^* : (x^*, \langle x^*, x\rangle + \varepsilon) \in \left[\overline{\operatorname{co}}^{w^*} \mathcal{G}\right]_\infty + \{0\} \times [0, \varepsilon]\right\}. \end{aligned}$$

Where

$$\mathcal{E} := \left\{ \Big(\int_T x^* \mathrm{d}\mu, \int_T \alpha \mathrm{d}\mu\Big) : \begin{array}{c} x^* \in \mathrm{L}^1_{w^*}(T, X^*), \\ \alpha \in \mathrm{L}^1(T, \mathbb{R}), \\ (x^*(t), \alpha(t)) \in \operatorname{epi} f_t^* \ ae \end{array} \right\} \tag{5.30}$$

$$\mathcal{G} := \left\{ \Big(\int_T x^* \mathrm{d}\mu, \int_T \alpha \mathrm{d}\mu\Big) : \begin{array}{c} x^* \in \mathrm{L}^1_{w^*}(T, X^*), \\ \alpha \in \mathrm{L}^1(T, \mathbb{R}), \\ (x^*(t), \alpha(t)) \in \operatorname{gph} f_t^* \ ae \end{array} \right\} \tag{5.31}$$

Proof. For the first two equalities see [54, Lemma 5.], while the third one is given by Corollary 5.34. So, we only have to prove the fourth equality, or equivalently, the inclusion "$\subseteq$ ". On the one hand, we have that

$$\operatorname{cl}^{w^*} \mathcal{E} = \operatorname{cl}^{w^*} \left(\overline{\operatorname{co}}^{w^*} \mathcal{G} + \{0\} \times \mathbb{R}_+\right). \tag{5.32}$$

Indeed, to see the last inclusion, take $(x^*, \alpha) \in \mathcal{E}$ and let $(x^*(t), \alpha(t)) \in\in \operatorname{epi} f_t^* = \operatorname{gph} f_t^* + \{0\} \times \mathbb{R}_+$ ae such that $x^*(\cdot) \in \mathrm{L}^1_{w^*}(T, X^*)$, $\alpha(\cdot) \in \mathrm{L}^1(T, \mathbb{R})$, and $(x^*, \alpha) = (\int_T x^* \mathrm{d}\mu, \int_T \alpha \mathrm{d}\mu)$. Then, since $(I_f)^*$ is proper, we have that $\int_T f_t^*(x^*(t)) \mathrm{d}\mu(t) \in \mathbb{R}$ and so, writing

$$(x^*(t), \alpha(t)) = (x^*(t), f_t^*(x^*(t))) + (0, \alpha(t) - f_t^*(x^*(t))) \in \operatorname{gph} f_t^* + \{0\} \times \mathbb{R}_+,$$

we get that $(x^*, \alpha) \in \mathcal{G} + \{0\} \times \mathbb{R}_+$, besides by the convexity of f_t^*'s $\mathcal{G} + \{0\} \times \mathbb{R}_+ \subseteq \mathcal{E}$

On the other hand,since ($(I_f)^*$ is proper) we have that

$$\begin{aligned} \left[\overline{\operatorname{co}}^{w^*} \mathcal{G}\right]_\infty \cap (-[\{0\} \times \mathbb{R}_+]_\infty) &\subseteq \left[\operatorname{cl}^{w^*} \mathcal{E}\right]_\infty \cap (\{0\} \times \mathbb{R}_-) \\ &= (\operatorname{epi}(I_f)^*)_\infty \cap (\{0\} \times \mathbb{R}_-) = \{(0,0)\}, \end{aligned}$$

and so by Dieudonné's theorem (see [22, Thorem I-10] or [43, Proposition 1.]) the set $\overline{\operatorname{co}}^{w^*} \mathcal{G} + \{0\} \times \mathbb{R}_+$ is closed. Hence, (5.32) reads

$$\left[\operatorname{cl}^{w^*} \mathcal{E}\right]_\infty = \left[\overline{\operatorname{co}}^{w^*} \mathcal{G} + \{0\} \times \mathbb{R}_+\right]_\infty = \left[\overline{\operatorname{co}}^{w^*} \mathcal{G}\right]_\infty + \{0\} \times \mathbb{R}_+.$$

Now, we take $x^* \in X^*$ such that $(x^*, \langle x^*, x\rangle + \varepsilon) \in \left[\operatorname{cl}^{w^*} \mathcal{E}\right]_\infty$. Then, by the last relations, there exist $(y^*, \gamma) \in \left[\overline{\operatorname{co}}^{w^*} \mathcal{G}\right]_\infty$ and $\eta \geq 0$ such that $(x^*, \langle x^*, x\rangle + \varepsilon) = (y^*, \gamma + \eta)$; hence, $x^* = y^*$. Moreover, using Theorem 5.32, we have

$$\operatorname{dom} I_f \times \{-1\} \subseteq [(\operatorname{epi}(I_f)^*)_\infty]^\circ = \left[\left(\operatorname{cl}^{w^*} \mathcal{E}\right)_\infty\right]^\circ \subseteq \left[\left(\operatorname{cl}^{w^*} \mathcal{G}\right)_\infty\right]^\circ,$$

so that $\langle (x^*,\gamma),(x,-1)\rangle \leq 0$, and $\eta = \langle x^*,x\rangle - \gamma + \varepsilon \leq \varepsilon$; that is,

$$(x^*,\langle x^*,x\rangle) \in \left[\overline{\mathrm{co}}^{w^*}(\mathcal{G})\right]_\infty + \{0\}\times[0,\varepsilon]$$

□

Consequently, we obtain a complete explicit characterization of the ε-subdifferential of I_f :

Theorem 5.37 *Assume that f is a convex normal integrand. Then for every $x \in X$ and $\varepsilon \geq 0$ we have that*

$$\partial_\varepsilon I_f(x) = \bigcap_{L\in\mathcal{F}(x)} \bigcup_{\substack{\varepsilon_1,\varepsilon_2\geq 0\\ \varepsilon=\varepsilon_1+\varepsilon_2\\ \ell\in\mathcal{I}(\varepsilon_1)\\ \eta\in L^1}} \bigcap_{\eta\in L^1(T,(0,+\infty))} \mathrm{cl}\left\{\int_T \left(\partial_{\ell(t)+\eta(t)} f_t(x) + A_L^{\varepsilon_2}(x)\right) \mathrm{d}\mu(t)\right\}$$

$$= \bigcap_{L\in\mathcal{F}(x)} \bigcup_{\substack{\varepsilon_1,\varepsilon_2\geq 0\\ \varepsilon=\varepsilon_1+\varepsilon_2\\ \ell\in\mathcal{I}(\varepsilon_1)\\ \eta\in L^1}} \bigcap_{\eta\in L^1(T,(0,+\infty))} \mathrm{cl}\left\{\int_T \left(\partial_{\ell(t)+\eta(t)} f_t(x) + B_L^{\varepsilon_2}(x)\right) \mathrm{d}\mu(t)\right\},$$

where the closure is taken with respect to the strong topology $\beta(X^*,X)$,

$$A_L^{\varepsilon_2}(x) := \left\{x^*\in X^* : (x^*,\langle x^*,x\rangle+\varepsilon_2) \in \left[\mathrm{cl}^{w^*}\left(\mathcal{E} + L^\perp\times\mathbb{R}_+\right)\right]_\infty\right\}$$
$$B_L^{\varepsilon_2}(x) := \left\{x^*\in X^* : (x^*,\langle x^*,x\rangle+\varepsilon_2) \in \left[\mathrm{cl}^{w^*}\left(\mathrm{co}\left(\mathcal{G}\right) + L^\perp\times\mathbb{R}_+\right)+\right]_\infty + \{0\}\times[0,\varepsilon_2]\right\}.$$

and $\mathcal{E}$ and $\mathcal{G}$ are defined in (5.30) and (5.31) respectively.

Proof. According to Theorem 5.19, we only have to prove that for every $L \in \mathcal{F}(x)$ and $\varepsilon \geq 0$, $N^\varepsilon_{\mathrm{dom}\, I_f\cap L}(x) = A_L^\varepsilon(x) = B_L^\varepsilon(x)$. Indeed, it suffices to apply Proposition 5.36 with the measurable space $(\tilde{T},\tilde{\mathcal{A}},\tilde{\mu})$, where $\tilde{T} := T\cup\{\omega_0\}$ for an element $\omega_0 \notin T$, $\tilde{\mathcal{A}}$ is the σ-Algebra generated by $(\mathcal{A}\cup\{\omega_0\})$, and $\tilde{\mu}$ is defined by

$$\tilde{\mu}(G) := \begin{cases} \mu(G\backslash\{\omega_0\}) + 1 & \text{if } \omega_0 \in G\\ \mu(G) & \text{if } \omega_0 \notin G,\end{cases}$$

and the integrand function $g(t,x) := f(t,x)$ for $t\in T$ and $g(\omega_0,x) := \delta_L(x)$. □

5.7 Characterizations via (exact-) subdifferentials

In this section, we use the previous results and Bronsted-Rockafellar theorems to obtain sequential formulas for the subdifferential of integral functions. As in the previous section, we suppose that either X, X^* are Suslin or $(T,\mathcal{A}) = (\mathbb{N},\mathcal{P}(\mathbb{N}))$.

We recall that a net of weakly measurable functions $g_{\mathrm{i}} : T\to X$ is said to converge uniformly ae to $g : T\to X$ if for all continuous seminorm ρ in X, the net $\rho(g_{\mathrm{i}} - g)$ converges to 0 in $\mathrm{L}^\infty(T,\mathbb{R})$.

Theorem 5.38 *Suppose that one of the following conditions holds:*

(i) X is Banach.

(ii) f_t are epi-pointed ae t.

Then for every $x \in X$, we have that $x^* \in \partial I_f(x)$ if and only if there exist a net of finite-dimensional subspaces $(\mathrm{L_i})_{\mathrm{i}}$ and nets measurable selections (x_{i}), (x_{i}^*) and (y_{i}), (y_{i}^*) such that $x_{\mathrm{i}}^*(t) \in \partial f(t, x_{\mathrm{i}}(t))$, $y_{\mathrm{i}}^*(t) \in N_{\mathrm{dom}\, I_f \cap \mathrm{L_i}}(y(t))$ ae, and:

(a) $(x_{\mathrm{i}}^* + y_{\mathrm{i}}^*) \subset \mathrm{L}^1(T, X^*)$ and $x^* = w^*\text{-}\lim \int\limits_T (x_{\mathrm{i}}^*(t) + y_{\mathrm{i}}^*(t)) \mathrm{d}\mu(t)$.

(b) $x_{\mathrm{i}}, y_{\mathrm{i}} \to x$ uniformly ae.

(c) $f(\cdot, x_{\mathrm{i}}(\cdot)) \to f(\cdot, x)$ uniformly ae.

(d) $\langle x_{\mathrm{i}}^*(\cdot), x_{\mathrm{i}}(\cdot) - x\rangle$, $\langle y_{\mathrm{i}}^*(\cdot), y_{\mathrm{i}}(\cdot) - x\rangle \to 0$ uniformly ae.
In addition, if X is reflexive and separable, then we take sequences instead of nets, and the w^*-convergence is replaced by the norm.

Proof. W.l.o.g. we may assume that $\mu(T) < +\infty$. Take $x_0^* \in \partial I_f(x_0)$, $x_0 \in X$, and fix $L \in \mathcal{F}(x_0)$. By (5.4) in Theorem 5.9 we find a measurable function $z^*(\cdot)$ such that $z^*(t) \in \partial(f_t + \delta_{L \cap \mathrm{dom}\, I_f})(x_0)$ for all $t \in \tilde{T}$ (with $\mu(T \backslash \tilde{T}) = 0$), and $x_0^* = \int_T z^*(t) \mathrm{d}\mu(t)$. Next, given $n \in \mathbb{N}$, continuous seminorms ρ_X on X and a w^*-continuous seminorm $\rho_{X^*} := \sigma_C$ on X^* such that C is finite and $\operatorname{span} C \supset \operatorname{span}(L \cap \mathrm{dom}\, I_f)$, we define the multifunction $B : \tilde{T} \rightrightarrows X \times X^* \times X \times X^*$ by $(x, x^*, y, y^*) \in B(t)$ if and only if

A(i) $x^* \in \partial f(t, x)$, $y^* \in N_{\mathrm{dom}\, I_f \cap L}(y)$.

A(ii) $\rho_X(x - x_0) \leq 1/n$, $\rho_X(y - x_0) \leq 1/n$, $\rho_{X^*}(z^*(t) - x^* - y^*) \leq 1/n$.

A(iii) $|f(t, x) \to f(t, x_0)| \leq 1/n$, $|\langle x^*, x - x_0\rangle| \leq 1/n$ and $|\langle y^*, y - x_0\rangle| \leq 1/n$.

By [89, Theorem 2.3] (see, also, [113, Theorem 3]) (in case (i)) or by Theorem 3.14 (in case (ii)), $B(t)$ is non-empty for all $t \in \tilde{T}$. Hence, due to the measurability of the involved functions, B has a measurable graph, so that by Proposition 1.21 we conclude the existence of nets of measurable functions $x(\cdot)$, $y(\cdot)$, $x^*(\cdot)$, and $\tilde{y}^*(\cdot)$, which satisfy the properties A(i), A(ii) and A(iii) above. Now, we consider $y^*(t) := P_{\operatorname{span} C}^*(x^*(t)) - x^*(t) + P_{\operatorname{span} C}^*(\tilde{y}^*(t))$, where P^* is the adjoint of the a projection $P_{\operatorname{span} C}$ onto $\operatorname{span} C$. Then

$$\langle y_{\mathrm{i}}^*(t), u - y(t)\rangle = \langle \tilde{y}^*(t), u - y(t)\rangle \leq 0 \;\; \forall u \in \mathrm{dom}\, I_f \cap \mathrm{L_i}(\operatorname{span} C),$$

and so $y_{\mathrm{i}}^*(t) \in N_{\mathrm{dom}\, I_f \cap L}(y(t))$. Moreover, we have

$$\begin{aligned}
\rho_{X^*}(z^*(t) - x^*(t) - y^*(t)) &= \rho_{X^*}(z^*(t) - P_{\operatorname{span} C}^*(x^*(t)) - P_{\operatorname{span} C}^*(\tilde{y}^*(t))) \\
&= \sigma_C(z^*(t) - P_{\operatorname{span} C}^*(x^*(t)) - P_{\operatorname{span} C}^*(\tilde{y}^*(t))) \\
&= \sigma_C(z^*(t) - x^*(t) - \tilde{y}^*(t)) \leq 1/n,
\end{aligned}$$

$$
\begin{aligned}
|\langle y^*(t), y(t)-x_0\rangle| &= \left|\langle P^*_{\operatorname{span} C}(x^*(t)) - x^*(t) + P^*_{\operatorname{span} C}(\tilde{y}^*(t)), y(t)-x_0\rangle\right| \\
&= |\langle \tilde{y}^*(t), y(t)-x_0\rangle| \leq 1/n,
\end{aligned}
$$

and for all balanced bounded set $A \subset X$

$$
\begin{aligned}
\int_T \sigma_A(x^*(t)+y^*(t))\mathrm{d}\mu(t) &= \int_T \sigma_A(P^*_{\operatorname{span} C}(x^*(t)) + P^*_{\operatorname{span} C}(\tilde{y}^*(t)))\mathrm{d}\mu(t) \\
&= \int_T \sigma_{PA}(x^*(t)+\tilde{y}^*(t)-z^*(t))\mathrm{d}\mu(t) + \int_T \sigma_{PA}(z^*(t))\mathrm{d}\mu(t) < +\infty.
\end{aligned}
$$

The conclusion when X is reflexive and separable comes from the fact that we can take sequences instead of nets used above (using [89, Theorem 2.3]) and countable family of finite-dimensional subspaces (recall Remark 5.10).

It remains to verify the sufficiency implication. Take $x^* \in X^*$, $x \in X$, nets of finite-dimensional subspaces $\mathrm{L_i}$ and nets of measurable functions (x_{i}), (x^*_{i}) and (y_{i}), (y^*_{i}) as as in the statement of the theorem. Then for all $u \in X$ we obtain

$$
\begin{aligned}
\langle x^*, u-x\rangle &= \lim \int \langle x^*_{\mathrm{i}} + y^*_{\mathrm{i}}, u-x\rangle \mathrm{d}\mu \\
&= \int\limits_T \left(\langle x^*_{\mathrm{i}}(t), u-x_{\mathrm{i}}(t)\rangle + \langle x^*_{\mathrm{i}}(t), x_{\mathrm{i}}(t)-x\rangle + \langle y^*_{\mathrm{i}}(t), y_{\mathrm{i}}(t)-x\rangle\right)\mathrm{d}\mu(t) \\
&+ \int\limits_T \left(\langle y^*_{\mathrm{i}}(t), u-y_{\mathrm{i}}(t)\rangle\right)\mathrm{d}\mu(t) \\
&\leq \lim \int\limits_T \{f(t,u)-f(t,x_{\mathrm{i}}(t))\}\,\mathrm{d}\mu(t) + \lim \int\limits_T \{\langle x^*_{\mathrm{i}}, x_{\mathrm{i}}(t)-x\rangle\}\,\mathrm{d}\mu(t) \\
&+ \lim \int\limits_T \{\langle y^*_{\mathrm{i}}(t), y_{\mathrm{i}}(t)-x\rangle\}\,\mathrm{d}\mu(t) \\
&= I_f(u) - I_f(x);
\end{aligned}
$$

taht is, $x^* \in \partial I_f(x)$. $\square$

For the next result we need the following lemma.

Lemma 5.39 *Under the assumption of Proposition 5.32 suppose f_t are epi-pointed for almost every t and the integrable function satisfies $x^*(t) \in \operatorname{int}\operatorname{dom} f_t$ t-ae. Then*

$$
\partial I_f(x) = \bigcap_{\delta>0} \operatorname{cl}\left\{ \bigcup_{\ell\in\mathcal{I}(\delta)} (w) - \int\limits_T \partial_{\ell(t)} f(t,x) \cap \operatorname{int}(\operatorname{dom} f^*_t)\mathrm{d}\mu(t). \right\}
$$

Proof. Since the inclusion $\supseteq$ is trivial we focus on the opposite one. According to Theorem 5.32 it suffices to show that for every $\varepsilon > 0$, $\hat{\ell} \in \mathcal{I}(\varepsilon)$ and $z^* \in (w) - \int\limits_T \partial_{\hat{\ell}(t)} f(t,x)\mathrm{d}\mu(t)$ we have that $z^* \in \operatorname{cl}\left\{ \bigcup\limits_{\ell \in \mathcal{I}(2\varepsilon)} (w) - \int\limits_T \partial_{\ell(t)} f(t,x) \cap \operatorname{int}(\operatorname{dom} f_t^*) \mathrm{d}\mu(t) \right\}$. Since $f^*(\cdot, z^*(\cdot)) \in L^1(T, \mathbb{R})$, we have that $z^*(t) \in \operatorname{dom} f_t^*$ for ae $t \in T$, and so $z_\lambda^*(t) = (1-\lambda)z^*(t) + \lambda x_0^*(t) \in \operatorname{int}(\operatorname{dom} f_t^*) \cap \partial_{\ell_\lambda(t)} f_t(x)$ where $\lambda \in (0,1)$ and $\ell_\lambda(t) := \langle z_\lambda^*, x \rangle - f(t,x) - f^*(t, z_\lambda^*(t)) \geq 0$. By the Fenchel inequality and convexity of the f_t^*'s we get

$$\langle z_\lambda^*(t), x \rangle - f(t,x) - f^*(t, z^*(t)) \leq f^*(t, z_\lambda^*(t)) - f^*(t, z^*(t)) \leq \lambda(f^*(t, x_0^*(t)) - f^*(t, z^*(t))),$$

and so, since $f^*(t, z_\lambda^*) \to f^*(t, z^*(t))$ as $\lambda \downarrow 0$, we have $\ell_\lambda(t) \to \langle z^*, x \rangle - f(t,x) - f^*(t, z^*(t)) \leq \hat{\ell}(t)$, by the Lebesgue dominated convergence theorem we get

$$\lim_{\lambda \to 0} \int_T \ell_\lambda(t) \mathrm{d}\mu(t) \leq \int_T \hat{\ell}(t) \mathrm{d}\mu(t) \leq \varepsilon.$$

$\square$

Remark 5.40 The hypothesis $x^*(t) \in \operatorname{int} \operatorname{dom} f_t$ *t-ae* always holds when X is metrizable. Indeed, because X^* is separable and $\operatorname{int}(\operatorname{dom} f_t)$ is nonempty we can construct a measurable selection $w^*(t) \in \operatorname{int}(\operatorname{dom} f_t)$, moreover let $B := \{x \in X : \mathrm{d}(0,x) \leq 1\}$ the unity ball in X, then for every almost every $t \in T$ $|f^*(t, x^*(t))| + |\sigma_B(x^*(t) - w^*(t))| < +\infty$, then defining $z^*(t) := x^*(t) + \lambda(t)(w^*(t) - x^*(t)$ with $\lambda(t) := \big(1 + |f^*(t, w^*(t))| + |\sigma_B(x^*(t) - w^*(t))|\big)^{-1}$ we get that

$$\begin{aligned} |\alpha(t)| + 1 &\geq -(1-\lambda(t))\alpha(t) + \lambda(t) f^*(t, w^*(t)) \geq (1-\lambda(t)) f^*(t, x^*(t)) + \lambda(t) f^*(t, w^*(t)) \\ &\geq f^*(t, z^*(t)) \end{aligned}$$

Therefore $f(t,x) \geq \langle z^*(t), x \rangle - (|\alpha(t)| + 1)$ for every $t \in T$ and every $x \in X$.

Theorem 5.41 *Assume that the linear growthINTEGRAL condition (5.27) holds with $x_0^* \in L_{w^*}^\infty(T, X^*)$, and assume that either X is Banach, or the f_t's are epi-pointed and $x^*(t) \in \operatorname{int} \operatorname{dom} f_t$ ae. Then $x^* \in \partial I_f(x)$ if and only if there exist net of measurable functions $(x_{\mathrm{i}}) \subset X$, $(x_{\mathrm{i}}^*) \subset L_{w^*}^\infty(T, X^*)$ such that $x_{\mathrm{i}}^*(t) \in \partial f(t, x_{\mathrm{i}}(t))$ ae, and*

(a) $x^* = w^*\!-\lim \int\limits_T x_{\mathrm{i}}^*(t) \mathrm{d}\mu(t)$.

(b) $x_{\mathrm{i}} \to x$ *uniformly ae.*

(c) $\displaystyle\int\limits_T |f(t, x_{\mathrm{i}}(t)) - \langle x_{\mathrm{i}}^*(t), x_{\mathrm{i}}(t) - x_0 \rangle - f(t, x_0)| \mathrm{d}\mu(t) \to 0$

If X is reflexive, then the above nets are replaced by sequences, and the w^*-convergence by norm-convergence.

Proof. Let $u_0^* \in \partial I_f(x_0)$ for $x_0 \in X$, $n \geq 0$, ρ_X a continuous seminorm in X and ρ_{X^*} a w^*-continuous seminorm in X^*. We choose $\varepsilon \in (0, 1/2n)$ such that $\varepsilon \sup_{y^* \in \mathbb{B}_{\rho_X}^\circ(0,1)} \rho_{X^*}(y^*) \leq 1/(2n)$. Then, by Theorem 5.32 (or Lemma 5.39, when f_t are epi-pointed), we can choose $z^*(t) \in L_{w^+}^1(T, X^*)$ such that $\rho_{X^*}(x_0^* - \int z^*) \leq 1/2n$ and $z^*(t) \in \partial_{\ell(t)} f(t, x_0)$ t-ae ($z^*(t) \in \partial_{\ell(t)} f(t, x_0) \cap \operatorname{int}(\operatorname{dom} f_t)$ when f_t^* are epi-pointed) with $\int \ell(t) \leq \varepsilon^2$, then we define the measurable function $B : T \to X \times X^*$ by $(x^*, x^*) \in B(t)$ if and only if

(i) $x^* \in \partial f(t,x)$

(ii) $\rho_X(x - x_0) \leq \varepsilon$.

(iii) $x^* - z^*(t) \in \frac{\ell(t)}{\varepsilon} B^\circ_{\rho_X}(0,1)$.

(iv) $|f(t,x) - \langle x^*, x - x_0\rangle - f(t,x_0)| \leq 2\ell(t)$.

By Brønsted-Rockafellar's theorem in the case of X Banach (see [8, Theorem 1]), or by Theorem 3.8 in the case of epi-pointed functions the set $B(t)$ is nonempty ae, moreover by the measurability of the involved functions we can apply Theorem 1.21 and conclude that there exists $x^*(t) \in \partial f(t, x(t))$ (by 5.41 (i)) such that $x^*(t) \in L^1_{w^*}(T,X)$ (in $L^1(T,X)$ if $z^* \in z^* L^1(T,X)$), $\rho_X(x_0 - x_{(}\cdot))_\infty \leq 1/n$,

$$
\begin{aligned}
\rho_{X^*}(x_0^* - \int x^*) &\leq \rho_{X^*}(x_0^* - \int z^*) + \rho_{X^*}(x_0^* - \int z^*) && (5.33)\\
&\leq 1/2n + \sup_{y^* \in B^\circ_{\rho_X}(0,1)} \rho_{X^*}(y^*) \int \frac{\ell(t)}{\varepsilon} && (5.34)\\
&\leq 1/2n + 1/2n = 1/n, && (5.35)
\end{aligned}
$$

and $\int\limits_T |f(t, x_{\mathrm{i}}(t)) - \langle x^*_{\mathrm{i}}(t), x_{\mathrm{i}}(t) - x_0\rangle - f(t,x_0)| \mathrm{d}\mu(t) \leq 1/n$. To prove the sufficiency, let $x^* \in X^*$, $x \in X$, and nets $x_{\mathrm{i}}(\cdot)$, $x^*_{\mathrm{i}}(\cdot)$ that satisfies the conclusion of the Theorem, then for every $y \in X$

$$
\begin{aligned}
\langle x_0^*, y - x_0\rangle &= \langle x_0^* - \int x^*_{\mathrm{i}}, y - x_0\rangle + \int \langle x^*_{\mathrm{i}}(t), y - x_{\mathrm{i}}(t)\rangle + \int \langle x^*_{\mathrm{i}}(t), x_{\mathrm{i}}(t) - x_0\rangle\\
&\leq \langle x_0^* - \int x^*_{\mathrm{i}}, y - x_0\rangle + \int f(t,y) - \int f(t, x_{\mathrm{i}}(t)) + \int \langle x^*_{\mathrm{i}}(t), x_{\mathrm{i}}(t) - x_0\rangle\\
&\leq \langle x_0^* - \int x^*_{\mathrm{i}}, y - x_0\rangle + I_f(y) - I_f(x)\\
&+ \int\limits_T |f(t, x_{\mathrm{i}}(t)) - \langle x^*_{\mathrm{i}}(t), x_{\mathrm{i}}(t) - x_0\rangle - f(t,x_0)| \mathrm{d}\mu(t).
\end{aligned}
$$

So taking the limits we conclude. $\square$

Remark 5.42 The sequential formulas for subdiferential for convex integrand functional was first studied by Ioffe [64], in this paper the author present a characterization of the subdiferential of an integrand functional for the case where the measurable net belongs to $x_{\mathrm{i}}(\cdot) \in \mathrm{L}^p(T,X)$ and $x^*_{\mathrm{i}}(\cdot) \in \mathrm{L}^q_{w^*}(T,X^*)$ with $1/p + 1/q = 1$ and $p \in [1, +\infty)$, also the author left the open question of what happen when $p = \infty$. Later, Lopez-Thibault [80] using the representation of the dual of $\mathrm{L}^\infty(T,X)$ (recall that $\mathrm{L}^\infty(T,X)^* = \mathrm{L}^1_{w^*}(T,X^*) \oplus \mathrm{L}^{\mathrm{sing}}(T,X)$) and the sequential formulas for the subdiferential of the composition [116, Theorem 1] (see also [89,113,115]. Our Theorems 5.38 and 5.41 give another answer to the question presented by Ioffe, where the convergence of the measurable selection x_{i} converge uniformly almost everywhere.

Chapter 6

Sequential and exact formulae for the subdifferential of nonconvex integral functionals

In this Chapter $(X, \|\cdot\|)$ will be a Banach space. Unless stated otherwise in the document, in the remainder of this chapter we may assume f is an integrand function from $T \times X$ to $[0, +\infty]$ and one of the following settings: X is separable, or $(T, \mathcal{A}) = (\mathbb{N}, \mathcal{P}(\mathbb{N}))$, here $\mathcal{P}(\mathbb{N})$ denote the power set of $\mathbb{N}$, that is, the set of all subset of $\mathbb{N}$. Although the assumption about the range of the values of the integrand appear less general, many of the results in the literature can be obtained in our setting by modifying appropriately the integrand. We will talk more in depth about these techniques in Theorems 6.11 and 6.12. it is important to recall that under our framework the integral functional $\tilde{I}_f : (\mathrm{L}^p(T, X), \|\cdot\|_p) \to \overline{\mathbb{R}}$ is lsc.

6.1 Sequential formulae for subdifferential calculus of nonconvex integrand functionals

6.1.1 Subdifferential of I_f by means of the chain rule

First, we start with the analysis of the subdifferential of I_f under the hypothesis that the measure is non-atomic. To complete this task we establish some sequential approximation rules for the Fréchet subdifferential using the well-known chain rule for this subdifferential together with formulas for the Fréchet subdifferential of an integral functional defined on $\mathrm{L}^p(T, X)$, which are available in the literature.

We have the following theorem. Recall that by a modulus α, we mean a nondecreasing function such that $\alpha(0) = 0$ and $\alpha(u) \to 0$ as $u \to 0$.

Theorem 6.1 *Let $(T, \mathcal{A}, \mu)$ be a finite non-atomic measure space, let X be a separable*

Asplund space and let $f : T \times X \to [0, +\infty]$ be a normal integrand function. Consider $x^* \in \partial I_f(\bar{x})$. Then, for any w^*-continuous seminorm ρ and $p \in (1, \infty)$, there exist sequences $y_n \in X$, $x_n \in \mathrm{L}^p(T, X)$, $x_n^* \in \mathrm{L}^q(T, X^*)$ (with $1/p + 1/q = 1$) such that:

(a) *For some numbers $r_n > 0$, moduli $\alpha_n : [0, r_n] \to \mathbb{R}_+$, and families $(a_s^n)_{s \in (0, r_n)}$ of $\mathrm{L}^1(T, \mathbb{R}_+)$ and $b_s^n := \alpha_n(s) - \|a_s^n\|_1 \geq 0$ ($s \in (0, r_n]$) ($n \geq 1$), we have that for every $s \in (0, r_n]$ and $(t, x) \in T \times X$*

$$f(t, x_n(t) + sx) - f(t, x_n(t)) - \langle x^*(t), sx \rangle \geq -s(a_s^n(t) + b_s^n \|x\|^p).$$

(b) $\|\bar{x} - y_n\| \to 0$, $\int_T \|\bar{x} - x_n(t)\|^p \mathrm{d}\mu(t) \to 0$,

(c) $\|x_n^*(\cdot)\|_q \|x_n(\cdot) - y_n\|_p \to 0$,

(d) $\rho(x^* - \int_T x_n^*(t) \mathrm{d}\mu(t)) \to 0$,

(e) $\int_T |f(t, x_n(t)) - f(t, \bar{x})| \mathrm{d}\mu(t)$.

Proof. Consider $\varepsilon_n \to 0^+$, the space $Y := \mathrm{L}^p(T, X)$, the linear function $A : X \to Y$ given by $A(x) = x \mathbb{1}_T$, and the functional $\hat{I}_f : Y \to \overline{\mathbb{R}}$. So, by [14, Theorem 3.5.2] there exist $x_n \in B(\bar{x}, \varepsilon_n)$, $y_n \in \mathbb{B}(A(\bar{x}), \varepsilon_n)$, $y_n^* \in \partial_F \hat{I}_f(y_n(\cdot))$, $\|\lambda_n^* - y_n^*\| < \varepsilon_n$ and $z_n^* = A^* \circ \lambda_n^*$ such that $|\hat{I}_f(y_n) - \hat{I}_f(A(\bar{x}))| < \varepsilon_n$, and

$$\max(\|\lambda_n^*\|, \|y_n^*\|, \|z_n^*\|) \|y_n - F(x_n)\| < \varepsilon_n, \;\; \rho(x^* - z_n^*) < \varepsilon_n.$$

Thus, by [90, Theorem 22] there exist $r_n > 0$, a modulus $\alpha_n : [0, r] \to \mathbb{R}_+$ and a family $(a_s^n)_{s \in (0, r_n)}$ of $\mathrm{L}^1(T, \mathbb{R}_+)$ such that for every $s \in (0, r_n]$, $(t, x) \in T \times X$ one has $b_s := \alpha(s) - \|a_s\|_1 \geq 0$ and

$$f(t, x(t) + sx) - f(t, x(t)) - \langle x^*(t), sx \rangle \geq -s(a_s(t) + b_t \|x\|^p).$$

Moreover, translating the estimate given by the chain rule in terms of measurable selections, we get

1. $\|\bar{x} - y_n\| \to 0$, $\displaystyle\int_T \|\bar{x} - x_n(t)\|^p \mathrm{d}\mu(t) \to 0$.
2. $\|x_n^*\|_q \|x_n(\cdot) - y_n\|_p \to 0$.
3. $\rho(x^* - \int_T x_n^*(t) \mathrm{d}\mu(t)) \to 0$.
4. $\int_T f(t, x_n(t)) \mathrm{d}\mu(t) \to \int_T f(t, \bar{x}) \mathrm{d}\mu(t)$.

Finally, Lemma 6.5 implies $\displaystyle\int_T |f(t, x_n(t)) - f(t, x_0)| \mathrm{d}\mu(t) \to 0$. □

By the following example we show the impossibility of using the same technique to find an approach of the subdifferential of the integral functional for $p = 1$.

Example 6.2 Let $(T, \mathcal{A}, \mu) = (]0, 1], \mathcal{L}, \lambda)$ be the Lebesgue measure on $[0, 1]$, and consider the normal integrand $f(t, x) = \mathrm{e}^{-x^2}$. Then $I_f(x) = \int_0^1 f(t, x) \mathrm{d}\lambda(t) = \mathrm{e}^{-x^2}$ for all $x \in \mathbb{R}$. By [90, Theorem 12] we have that $\partial \hat{I}_f(w(\cdot)) = \emptyset$ for every $w \in \mathrm{L}^1(T, \mathbb{R})$. In other words, if we consider the operator I_f as the composition of $\hat{I}_f$ and the linear function $A : \mathbb{R} \to \mathrm{L}^1(T, \mathbb{R})$ given by $x \to x \mathbb{1}_T$, then it is impossible to approximate the subdifferential of I_f in terms of $A^*(\partial \hat{I}_f(w(\cdot))$.

6.1.2 Robusted local minima

In order to guarantee a general framework (without the atmoless restriction on the measure space) for the subdifferential of integral functions, we adopt here the notion of *Robusted local minima* (see for example [14, 65, 91]) to the case of integral functions. This definition allows us to understand better the approximate formulae for the subdifferential of such integrand functions. In what follows, we work with an arbitrary complete σ-finite nonnegative measure space $(T, \mathcal{A}, \mu)$.

Definition 6.3 *Consider a function $f : T \times X \to \overline{\mathbb{R}}$ and $p \in [1, +\infty]$. We define the p-stabilized infimum of I_f on $B \subseteq X$ by*

$$\wedge_{p,B} I_f := \sup_{\varepsilon>0} \inf\{ \int\limits_T f(t, x(t)) \mid x(\cdot) \in \mathrm{L}^p(T, X), \;\; y \in B \text{ and } \int\limits_T \|x(t) - y\|^p \mathrm{d}\mu(t) \leq \varepsilon\}.$$

The infimum of I_f is called p-robust if $\wedge_{p,B} I_f = \inf_B I_f$ and both quantities are finite, in which case a minimizer of I_f on B will be called a p-robust minimizer of I_f on B. A point x will be called a p-robust minimum of I_f provided the existence of some $\delta > 0$ such that x is a p-robust minimizer on $\mathbb{B}(x, \delta)$.

It is worth mentioning that one can easily prove (using *Hölder's inequality* [7, Corollary 2.11.5]) that when the measure is finite, a p-robust minimizer of I_f is also an r-robust minimizer for every $r \geq p$.

Example 6.4 (A p-robust minimum which is not an r-robust minimum for every $r < p$) Consider $(T, \mathcal{A}, \mu) = (]0, 1], \mathcal{L}, \lambda)$, the Lebesgue measure on $]0, 1]$ and $\ell \in \mathrm{L}^p(T, \mathbb{R}) \backslash \mathrm{L}^r(T, \mathbb{R})$. We use the integrand $f(t, x) = -\|x\|^p + \delta_{\mathbb{B}(0,1)}(x)$, so that for every $r < p$, $\wedge_{r,\mathbb{B}(0,1)} I_f = -\infty$. But, if we take $r \geq p$, then for every ε_n-minimizer (y_n, x_n) of $\wedge_{r,\mathbb{B}(0,1)} I_f$, with $\varepsilon_n \to 0^+$, we can take a convergent subsequence $y_{n_k} \to \bar{y}$. Then $\int\limits_T \|x_n(t) - \bar{y}\|^r \mathrm{d}\mu(t) \to 0$, and so $\wedge_{r,\mathbb{B}(0,1)} I_f = \int_T -\|x_n(t)\|^p \mathrm{d}\mu(t) \to \int_T -\|\bar{y}\|^p \geq \inf_{\mathbb{B}(0,1)} I_f$.

The following Lemma shows that the graphical convergence of an integral function gives convergence in $\mathrm{L}^1(T, \mathbb{R})$ of the values of the function.

Lemma 6.5 *Consider $x_n \in \mathrm{L}^p(T, X)$ such that $x_n \xrightarrow{\mathrm{L}^p} x$ and*

$$\lim \int\limits_T f(t, x_n(t)) \mathrm{d}\mu(t) = \int\limits_T f(t, x(t)) \mathrm{d}\mu(t) \in \mathbb{R}.$$

Then $\lim \int\limits_T |f(t, x_n(t)) - f(t, x(t))| \mathrm{d}\mu(t) = 0.$

Proof. Consider $\delta > 0$. By the lower semicontinuty of $\hat{I}_f$ in $\mathrm{L}^p(T, X)$ there exists $\varepsilon > 0$ such that $-\delta/4 + \hat{I}_f(x) \leq \hat{I}_f(y(\cdot))$ for every $y \in \mathbb{B}_{\mathrm{L}^p(T,X)}(x, \varepsilon)$. Since $x_n \to x$, there exists $n_1 \in \mathbb{N}$

such that $x_n \in B_{\mathrm{L}^p(T,X)}(x,\varepsilon)$ for every $n \geq n_1$. In particular, for every $A \in \mathcal{A}$ and every $n \geq n_1$ the function $y := x_n \mathbb{1}_A + x\mathbb{1}_{A^c} \in \mathbb{B}_{\mathrm{L}^p(T,X)}(x,\varepsilon)$, and then

$$-\delta/4 + \int_A f(t,x(t))\mathrm{d}\mu(t) \leq \int_A f(t,x_n(t))\mathrm{d}\mu(t)$$

for every $A \in \mathcal{A}$ and every $n \geq n_1$. This yields

$$\begin{aligned}-\delta/4 + \int\limits_A f(t,x(t))\mathrm{d}\mu(t) &\leq \int\limits_A f(t,x_n(t))\mathrm{d}\mu(t) = \int\limits_T f(t,x_n(t))\mathrm{d}\mu(t) - \int\limits_{A^c} f(t,x_n(t))\mathrm{d}\mu(t)\\ &\leq \int\limits_T f(t,x_n(t))\mathrm{d}\mu(t) - \int\limits_{A^c} f(t,x(t)) + \delta/4,\ \forall A \in \mathcal{A},\ \forall n \geq n_1.\end{aligned}$$

From the fact that $\lim \int_T f(t,x_n(t))\mathrm{d}\mu(t) = \int_T f(t,x(t))\mathrm{d}\mu(t)$ there must exist $n_2 \geq n_1$ such that $\int_T f(t,x_n(t))\mathrm{d}\mu(t) \leq \int_T f(t,x(t))\mathrm{d}\mu(t) + \delta/4$ for all $n \geq n_2$. Thus

$$\begin{aligned}-\delta/4 + \int\limits_A f(t,x(t))\mathrm{d}\mu(t) &\leq \int\limits_A f(t,x_n(t))\mathrm{d}\mu(t) \leq \int\limits_T f(t,x_n(t))\mathrm{d}\mu(t) - \int\limits_{A^c} f(t,x(t)) + \delta/4\\ &\leq \int\limits_T f(t,x(t))\mathrm{d}\mu(t) + \delta/4 - \int\limits_{A^c} f(t,x(t)) + \delta/4\\ &= \int\limits_A f(t,x(t))\mathrm{d}\mu(t) + \delta/2,\ \forall A \in \mathcal{A},\ \forall n \geq n_2.\end{aligned}$$

Then considering the measurable sets $A_n^+ := \{\, t \in T : f(t,x_n(t)) - f(t,x(t)) > 0\}$ and $A_n^- := \{\, t \in T : f(t,x_n(t)) - f(t,x(t)) < 0\}$, we get

$$\begin{aligned}\int\limits_T |f(t,x_n(t)) - f(t,x(t))|\mathrm{d}\mu(t) &= \int\limits_{A_n^+} f(t,x_n(t)) - f(t,x(t))\mathrm{d}\mu(t)\\ &\quad + \int\limits_{A_n^-} f(t,x(t)) - f(t,x_n(t))\mathrm{d}\mu(t)\\ &\leq \delta/2 + \delta/4 < \delta;\end{aligned}$$

that is, $\int\limits_T |f(t,x_n(t)) - f(t,x(t))|\mathrm{d}\mu(t) \to 0.$ □

The following Lemma is a simple application of classical rules concerning differentiation of integral functionals. For the sake of completeness we give a proof.

Lemma 6.6 *Let μ be a finite measure and let $f : T \times X \to \overline{\mathbb{R}}$ be a normal integrand Lipschitz on $\mathbb{B}(x_0,\delta)$ with some p-integrable constant, that is to say, there exists $K \in \mathrm{L}^p(T,\mathbb{R})$ such that $|f(t,x) - f(t,y)| \leq K(t)|x-y|$, for all $x, y \in \mathbb{B}(x_0,\delta)$ and all $t \in T$. Assume that the functions f_t are β-differentiable at x_0 ae. Then I_f is β-differentiable at x_0, $\nabla_\beta f(t,x_0)$ belongs to $\mathrm{L}^p_{w^*}(T,X^*)$ and $\nabla_\beta I_f(x_0) = (G) - \int_T \nabla f_t(x_0)\mathrm{d}\mu(t)$. Moreover, if f_t are β-smooth on $\operatorname{int}(\mathbb{B}(x_0,\delta))$, then I_f is β-smooth on $\operatorname{int}(\mathbb{B}(x_0,\delta))$. Finally, if X is a Hilbert space, the functions f_t are C^2 on $\operatorname{int}(\mathbb{B}(x_0,\delta))$ and the derivative $\nabla f : T \times X \to X^*$ is Lipschitz on $\mathbb{B}(x_0,\delta)$; hence, I_f is C^2 on $\operatorname{int}(\mathbb{B}(x_0,\delta))$.*

Proof. First, the w^*-integrability of the function $t \to \nabla_\beta f_t(x_0)$ follows from the fact that for every $h \in X$, $\langle \nabla_\beta f_t(x_0), h\rangle = \lim\limits_{s\to 0^+} \frac{f(t,x_0+sh)-f(t,x_0)}{s}$ and $\|\nabla_\beta f_t(x_0)\| \le K(t)$. Now take $U \in \beta$, and any sequence $s_n \to 0^+$. Since U is bounded we may assume that $x_0 + s_n h \in \mathbb{B}(x_0,\delta)$ for every $n \in \mathbb{N}$ and $h \in U$. So, when the space X is separable, the measurability of $t \to \sup\limits_{h\in U}\left|\frac{f_t(x_0+s_nh)-f_t(x_0)}{s_n} - \langle \nabla_\beta f_t(x_0), h\rangle\right|$ follows from the Lipschitz continuity of the integrand and the separability of U. We notice that this function is bounded from above by K; moreover, it converges to zero (ae) as $n \to \infty$. Then by Lebesgue's dominated convergence theorem we get $\lim\limits_{n\to\infty} \sup_{h\in U}\left|\frac{I_f(x_0+s_nh)-I_f(x_0)}{s_n} - \int_T \langle \nabla_\beta f_t(x_0), h\rangle \mathrm{d}\mu(t)\right| \overset{n\to\infty}{\to} 0$, which concludes the first part. To prove the continuity of the derivative $\nabla_\beta I_f : \mathrm{int}(\mathbb{B}(x_0,\delta)) \to (X^*, \beta^*)$ we fix $U \in \beta$, $x \in \mathrm{int}(\mathbb{B}(x_0,\delta))$ and $x_n \to x$ with $x_n \in \mathbb{B}(x_0,\delta)$. By the boundedness of U the number $M := \sup_{h\in U} \| h\|$ is finite. Then we notice that for almost all $t \in T$, $\lim\limits_{n\to\infty} \sup_{h\in U} |\langle \nabla f_t(x) - \nabla f_t(x_n), h\rangle| = 0$, and $g_n(t) := \sup_{h\in U} |\langle \nabla f_t(x) - \nabla f_t(x_n), h\rangle| \le 2MK(t)$ ae. Then again by Lebesgue's dominated convergence theorem, we get $\sup_{h\in U}\left|\langle \nabla_\beta I_f(x) - \nabla_\beta I_f(x_n), h\rangle\right| \overset{n\to\infty}{\longrightarrow} 0$. In the final case when X is a Hilbert space, the fact that functions f_t are C^2 uses similar arguments and so we omit the proof. □

We recall that in every β-smooth space X there exists a *Leduc function* ψ, which is a (globally) Lipschitz continuous function, β-smooth away from the origin, and satisfies the existence of some constant $a > 0$ such that $\|x\| \le \psi(x) \le a\|x\|$ and $\psi(tx) = t\psi(x)$ for all $x \in X$ and $t > 0$ (for more details see [69] and the references therein).

Lemma 6.7 *Let $p \in [1,+\infty)$ and $x \in X$ be a p-robust minimizer of I_f on $B \subseteq X$. Then for every ε_n-minimizer $(x_n(\cdot), y_n)$ (with $\varepsilon_n \to 0$) of $\varphi_n : \mathrm{L}^p(T,X) \times X \to \overline{\mathbb{R}}$,*

$$\varphi_n(w,u) := \int\limits_T f(t,w(t))\mathrm{d}\mu(t) + n\int\limits_T \psi^p(w(t)-u)\mathrm{d}\mu(t) + \psi^p(x_0-u) + I_B(u),$$

where ψ is a Leduc function, we have:

(a) $n(\|x_n(\cdot) - y_n\|_p)^p$, $\|x_n(\cdot) - x\|_p$, $\|y_n - x\| \to 0$, *and*

(b) $\int\limits_T |f(t,x_n(t)) - f(t,x)|\mathrm{d}\mu(t) \to 0.$

In particular

$$\sup_{n\in\mathbb{N}} \inf_{\substack{w\in \mathrm{L}^p(T,X)\\ u\in X}} \varphi_n(w,u) = I_f(x). \tag{6.1}$$

Proof. First, for $n \ge 1$ and $\delta > 0$ define

$$\nu_n := \inf\{\, \varphi_n(w,u) \mid w \in \mathrm{L}^p(T,X) \text{ and } u \in X\},$$
$$\xi_\delta := \inf\{\int_T f(t,w(t)) \mid \int_T \|w(t)-u\|\mathrm{d}\mu(t) \le \delta,\ w \in \mathrm{L}^p(T,X) \text{ and } u \in B\}.$$

Now we have, $n(\|x_n(\cdot) - y_n\|_r)^p \leq \int_T (f(t, x_n(t)) + n\psi^p(x_n(t) - y_n)) \,\mathrm{d}\mu(t) + \psi^p(y_n - x) \leq I_f(x) + \varepsilon_n$. The last inequality implies $\int_T \|x_n(t) - y_n\|^p \mathrm{d}\mu(t) \to 0$. Then, setting $\delta_n := \int_T \|x_n(t) - y_n\|^p \mathrm{d}\mu(t)$, we have

$$\xi_{\delta_n} - \varepsilon_n \leq \int_T f(t, x_n(t))\mathrm{d}\mu(t) - \varepsilon_n \leq \varphi_n^r(x_n, y_n) - \varepsilon_n \leq \nu_n \leq I_f(x_0).$$

Taking the limits we conclude that $\displaystyle\int_T f(t, x_n(t))\mathrm{d}\mu(t) \to \int_T f(t, x_0)\mathrm{d}\mu(t)$, and consequently (a) and equation (6.1) follow. Finally, using Lemma 6.5 we obtain (b). □

Lemma 6.8 *Let $z(\cdot)$ be a measurable function with values on X, and let $\varepsilon(\cdot)$ and $\lambda(\cdot)$ be two strictly positive measurable functions. Suppose that $z(t)$ is an $\varepsilon(t)$-minimum of f_t and one of the following condition holds:*

(i) $\partial = \partial_\beta$ and X is a β-smooth space.

(ii) $\partial = \partial_P$ and X is a Hilbert space.

Then there exist a constant $L = L(X, \beta)$, a measurable function $y(t)$ and a w^*-measurable function $y^*(t)$ such that $y^*(t) \in \partial f(t, y(t))$, $\|y(t) - z(t)\| \leq \lambda(t)$, $|f(t, y(t)) - f(t, z(t))| \leq \delta(t)$ and $\|y^*(t)\| \leq 2L\delta(t)/\lambda(t)$.

Proof. Consider a Leduc function ψ and take $L > 0$ such that ψ is L-Lipschitz and $\|x\| \leq \psi(x) \leq L\|x\|$, $\|x\| \leq \|\nabla_\beta \psi^2(x)\| \leq L\|x\|/2$ for all $x \in X$. Consider $\delta_{\mathrm{i}} > 0$ with $\delta_0 = 1$ such that $\sum_{\mathrm{i}=0}^\infty \delta_{\mathrm{i}} = 2$. Then define $\delta_{\mathrm{i}}(t) := \delta_{\mathrm{i}}\varepsilon(t)/\lambda^2(t)$, the space $S = X \times \prod_{\mathrm{i}=1}^\infty X$ with the product topology, and the function $(t, y, (x_{\mathrm{i}})) \to \varphi(y, (x_{\mathrm{i}})) = \sum_{\mathrm{i}=0}^\infty \delta_{\mathrm{i}}(t)\psi^2(y - x_{\mathrm{i}})$.

Now define the multifunction $(y, (x_{\mathrm{i}})) \in M(t)$ if and only if $\psi^2(y - z(t)) \leq \lambda^2(t)$, $\psi(x_{\mathrm{i}} - y)^2 \leq \lambda^2(t)/2^{\mathrm{i}}$ for all $\mathrm{i} = 1, 2, ...,$ $f(t, y) + \varphi(x, (y_{\mathrm{i}})) \leq f(z(t))$ and $f(t, w) + \varphi(w, (x_{\mathrm{i}})) \geq f(t, y) + \varphi(y, (x_{\mathrm{i}}))$ for all $w \in X$. It is not hard to prove that S is a Polish space (i.e. complete and separable), φ is measurable (with respect to $\mathcal{A} \otimes \mathcal{B}(S)$) and from the fact that every function involved in the multifunction M is $\mathcal{A} \otimes \mathcal{B}(S)$-measurable, we have $\operatorname{gph} M \in \mathcal{A} \otimes \mathcal{B}(S)$. Moreover, by Borwein-Preiss Variational Principle (see e.g. [14, Theorem 2.5.2]) $M(t)$ has non-empty values. Then by the measurable selection theorem (see e.g. [22, Theorem III.22]) there exist measurable functions $(y(t), x_{\mathrm{i}}(t)) \in M(t)$ ae. The last implies $\|y(t) - z(t)\| \leq \lambda(t)$, $\|x_{\mathrm{i}}(t) - y(t)\| \leq \lambda(t)/\sqrt{2}^{\mathrm{i}}$ for all $\mathrm{i} = 1, 2, ...,$ $|f(t, y(t)) - f(t, z(t))| \leq \varepsilon(t)$ and $f(t, w) + \varphi(w, (x_{\mathrm{i}}(t))) \geq f(t, y(t)) + \varphi(y(t), (x_{\mathrm{i}}(t)))$ for all $w \in X$ ae.

Finally, it is easy to see that $\phi(t, y) := \sum_{\mathrm{i}=0}^\infty \delta_{\mathrm{i}}(t)\psi^2(y - x_{\mathrm{i}}(t))$ is β-smooth (C^2 if X is a Hilbert space) with respect to the second argument (see Lemma 6.6), $\nabla_\beta \phi(t, y(t))$ is w^*-measurable and $\|\nabla_\beta \phi(t, y(t))\| \leq L\varepsilon(t)/\lambda(t)$. Hence, $f(t, \cdot) + \phi(t, \cdot)$ attains a minimum at $y(t)$ and so, $y^*(t) := -\nabla_\beta \phi(t, y(t)) \in \partial f(t, y)$. □

The following gives a sufficient condition for robusted local minima.

Proposition 6.9 *Consider $p, q \in [1, +\infty]$ with $1/p + 1/q = 1$, and $B \supseteq X$ such that* $\operatorname{dom} I_f \cap B \neq \emptyset$. *Suppose one of the following conditions is satisfied:*

(a) For almost every $t \in T$, $f(t,\cdot)$ is τ-lsc, B is τ-closed and there exists $A \in \mathcal{A}$ with $\mu(A) > 0$ such that for all $t \in A$, $f(t,\cdot)$ is sequentially τ-inf-compact, with τ some coarser topology than the norm topology (i.e. $\tau \subseteq \tau_{\|\cdot\|}$).

(b) For almost every $t \in T$, f is Lipschitz on X with some q-integrable constant.

Then

$$\wedge_{p,B} I_f = \inf_B I_f.$$

Proof. In the first case define $\nu_n := \inf_{\substack{w \in \mathrm{L}^p(T,X) \\ u \in B}} \{\int_T f(t, w(t)) \mid \int_T \|w(t) - u\|^p \leq 1/n\}$, take $\varepsilon_n \to 0^+$ and $(x_n, y_n) \in \mathrm{L}^p(T,X) \times X$ such that

$$-\varepsilon_n + \int_T f(t, x_n(t)) \leq \nu_n, \tag{6.2}$$

and $\int_T \|x_n(t) - y_n\|^p \leq 1/n$. We can suppose that for every $t_1 \in T$ and $t_2 \in A$, $\|x_n(t_1) - y_n\| \to 0$ and $f(t_2,\cdot)$ is sequentially τ-inf-compact. So, by Fatou's lemma we have that for every subsequence x_{n_k} of x_n we have

$$\int_T \liminf f(t, x_{n_k}(t)) \mathrm{d}\mu(t) \leq \liminf \int_T f(t, x_{n_k}(t)) \leq \inf_B I_f < +\infty. \tag{6.3}$$

Then in particular for some $t_0 \in A$, $\liminf f(t_0, x_n(t_0)) < +\infty$, and there exist a subsequence $x_{n_{k(t_0)}}(t_0)$ and a constant M_{t_0} such that $f(t_0, x_{n_{k(t_0)}}(t_0)) \leq M_{t_0}$. By the inf-compactness of $f(t_0,\cdot)$, there exists a subsequence z_n of $x_{n_{k(t_0)}}(t_0)$ such that $z_n \to w_0 \in X$. Because $\|x_n(t_0) - y_n\| \to 0$, we get the existence of a subsequence $y_{\phi(n)}$ of y_n such that $y_{\phi(n)} \xrightarrow{\tau} w_0 \in B$ (because B is τ-closed). Then from the fact that $\|x_n(t) - y_n\| \to 0$, we get $x_{\phi(n)}(t) \xrightarrow{\tau} w$ for all $t \in T$. Finally, taking into account (6.2) and using the lsc of the integrand in (6.3) we obtain

$$\inf_B I_f \leq I_f(w_0) \leq \int_T \liminf f(t, x_{\phi(n)}(t)) \mathrm{d}\mu(t) \leq \liminf \int_T f(t, x_{\phi(n)}(t)) \mathrm{d}\mu(t) \leq \wedge_{p,B} I_f \leq \inf_B I_f.$$

In the second case, let K be the q-integrable Lipschitz constant and consider $w \in \mathrm{L}^p(T,X)$ and $y \in B$. Then

$$\begin{aligned}
\int_T f(t, w(t)) \mathrm{d}\mu(t) &\geq -\int_T |f(t, w(t)) - f(t, y)| \mathrm{d}\mu(t) + \int_T f(t, y) \mathrm{d}\mu(t) \\
&\geq -\int_T K(t) \|w(t) - y\| \mathrm{d}\mu(t) + \inf_B I_f.
\end{aligned}$$

So, the result follows taking the appropriate limits. $\square$

6.1.3 β-smooth and proximal subdifferentials

Theorem 6.10 *Consider $p \in (1, +\infty)$. Assume that the measure μ is finite and $x_0 \in X$ is a p-robust local minimizer of I_f. Then if ∂, X and p satisfy one of the following conditions:*

(i) $\partial = \partial_\beta^-$, X is a β-smooth space, *(ii) $\partial = \partial_P$, X is a Hilbert space and $p \geq 2$,*

we have the existence of sequences $y_n \in X$, $x_n \in \mathrm{L}^p(T, X)$, $x_n^* \in \mathrm{L}^q_{w^*}(T, X^*)$ (with $1/p + 1/q = 1$) such that:

(a) $x_n^*(t) \in \partial f(t, x_n(t))$ ae,

(b) $\|x_0 - y_n\| \to 0, \int\limits_T \|x_0 - x(t)\|^p \mathrm{d}\mu(t) \to 0$,

(c) $\|x_n^*(\cdot)\|_q \|x_n(\cdot) - y_n\|_p \to 0$,

(d) $\|\int\limits_T x_n^*(t) \mathrm{d}\mu(t)\| \to 0$,

(e) $\int\limits_T |f(t, x_n(t)) - f(t, x_0)| \mathrm{d}\mu(t) \to 0$.

Proof. Consider the function $\ell(x) = \psi^p(x)$, where in case (i) ψ is a Leduc function and in case (ii) ψ is the norm. When X is a β-smooth space one can follow the construction given in [69, Lemma 2.5] and derive the existence of a constant L such that; $\|x\| \leq \psi(x) \leq L\|x\|$ and $\|x\|^{p-1} \leq \|\nabla_\beta \ell(x)\| \leq L\|x\|^{p-1}$ for all $x \in X$. Moreover, it is easy to see that ℓ is β-smooth everywhere. When X is a Hilbert space it is well-known that ℓ is C^2 on X and ℓ satisfies the same estimates.

Consider $\delta \in (0, 1)$ such that x_0 is a p-robust minimizer of I_f on $B := \mathbb{B}(x_0, \delta)$, and fix a family of $\delta_{\mathrm{i}} > 0$ such that $\delta_0 = \delta$ and $\sum_{\mathrm{i}=0}^{+\infty} \delta_{\mathrm{i}} = 1$. Now define $\varphi_n : \mathrm{L}^p(T, X) \times X \to \overline{\mathbb{R}}$ by

$$\varphi_n(x, y) = \int\limits_T f(t, x(t)) \mathrm{d}\mu(t) + n \int\limits_T \ell(x(t) - y) \mathrm{d}\mu(t) + \ell(y - x_0) + \delta_B(y).$$

Then Lemma 6.7 says that

$$\sup_{n \in \mathbb{N}} \inf_{\substack{w \in \mathrm{L}^p(T,X) \\ u \in X}} \varphi_n(w, u) = I_f(x_0),$$

and so there exists $\varepsilon_n \to 0^+$ (with $\varepsilon_n \in (0, \delta_0^2)$ for large enough n) such that (x_0, x_0) is an ε_n-minimum of φ_n. Then we can apply the Borwein-Preiss variational principle (see, e.g. [14, Theorem 2.5.3]) with the type-gauge function $\rho : (\mathrm{L}^p(T, X) \times X)^2 \to \mathbb{R}$ given by $\rho((w_1, u_1), (w_2, u_2)) := \int_T \ell(w_1(t) - w_2(t)) \mathrm{d}\mu(t) + \ell(u_1 - u_2)$ and find points $(x_{\mathrm{i}}^n, y_{\mathrm{i}}^n)_{\mathrm{i} \in \mathbb{N}}, (x_\infty^n, y_\infty^n) \in \mathrm{L}^p(T, X) \times X$ such that:

(BP.1) $\displaystyle\int\limits_T \ell(x_0 - x_\infty^n(t)) \mathrm{d}\mu(t) + \ell(x_0 - y_\infty^n) \leq \frac{\varepsilon_n}{\delta_0}, \int\limits_T \ell(x_{\mathrm{i}}^n(t) - x_\infty^n(t)) \mathrm{d}\mu(t) + \ell(y_{\mathrm{i}}^n - y_\infty^n) \leq \frac{\varepsilon_n}{2^{\mathrm{i}} \delta_0}$,

(BP.2) $\varphi_n(x_\infty^n, y_\infty^n) + \phi_n(x_\infty^n, y_\infty^n) \leq \varphi_n(x_0, x_0)$, and

(BP.3) $\varphi_n(w, u) + \phi_n(w, u) > \varphi_n(x_\infty^n, y_\infty^n) + \phi_n(x_\infty^n, y_\infty^n)$ for all $(w, u) \in \mathrm{L}^p(T, X) \times X \setminus \{(x_\infty^n, y_\infty^n)\}$, where

$$\phi_n(w,u) = \sum_{\mathrm{i}=1}^{\infty} \delta_{\mathrm{i}} \left(\int_T \ell(w(t) - x_{\mathrm{i}}^n(t)) \mathrm{d}\mu(t) \right) + \sum_{\mathrm{i}=1}^{\infty} \delta_{\mathrm{i}} \ell(u - y_{\mathrm{i}}^n)$$
$$= \int_T \left(\sum_{\mathrm{i}=1}^{\infty} \delta_{\mathrm{i}} \ell(w(t) - x_{\mathrm{i}}^n(t)) \right) \mathrm{d}\mu(t) + \sum_{\mathrm{i}=1}^{\infty} \delta_{\mathrm{i}} \ell(u - y_{\mathrm{i}}^n).$$

On the one hand, by (BP.2) $\int_T h_n(t, x_\infty^n(t))\mathrm{d}\mu(t)$ is finite, where $h_n(t,v) := f(t,v) + n\ell(v - y_\infty^n) + \sum_{\mathrm{i}=1}^{\infty} \delta_{\mathrm{i}} \ell(v - x_{\mathrm{i}}^n(t))$ is a normal integrand functional, and by (BP.3) (taking $u = y_\infty^n$)

$$\int_T h_n(t, x_\infty^n(t))\mathrm{d}\mu(t) = \inf_{w \in \mathrm{L}^p(T,X)} \int_T h_n(t, w(t))\mathrm{d}\mu(t)$$
$$\text{(by Proposition 5.2)} = \int_T \inf_{u \in X} h_n(t,u)\mathrm{d}\mu(t).$$

Then, by definition in case (i), and by the sum rule for case (ii), we have

$$0 \in \partial f(t, x_\infty^n(t)) + n u_n^*(t) + v_n^*(t) \; ae, \tag{6.4}$$

where $u_n^*(t) := \nabla_\beta \ell(x_\infty^n(t) - y_\infty^n)$ and $v^*(t) := \sum_{\mathrm{i}=1}^{\infty} \delta_{\mathrm{i}} \nabla_\beta \ell(x_\infty^n(t) - x_{\mathrm{i}}^n(t))$. The measurability and differentiability of these functions follow from Lemma 6.6 (notice that this infinite sum can also be seen as an integral functional). The estimate of the gradient of the function ℓ gives us $\|u_n^*(t)\|^q \leq \mathrm{L}^q \|x_\infty^n(t) - y_\infty^n\|^p$ and $\int_T \|v_n^*(t)\|^q \mathrm{d}\mu(t) \to 0$. On the other hand, by (BP.3) (taking $w = x_\infty^n$)

$$n \int_T \ell(x_\infty^n(t) - y_\infty^n)\mathrm{d}\mu(t) + \ell(y_\infty^n - x_0) + \sum_{\mathrm{i}=1}^{\infty} \delta_{\mathrm{i}} \ell(y_\infty^n - y_{\mathrm{i}}^n)$$
$$= \inf_{u \in X} \left(n \int_T \ell(x_\infty^n(t) - u)\mathrm{d}\mu(t) + \ell(u - x_0) + \sum_{\mathrm{i}=1}^{\infty} \delta_{\mathrm{i}} \ell(y_\infty^n - y_{\mathrm{i}}^n) \right).$$

Hence, again Lemma 6.6 gives us the differentiability of these three functions and simple calculus implies $0 = -n \int_T u_n^*(t)\mathrm{d}\mu(t) + w_n^*$ with $\|w_n^*\| \to 0$. Thus, there exists $x_n^* := -n u_n^*(t) - v_n^*(t) \in \mathrm{L}^q(T, X^*)$ such that $x_n^*(t) \in \partial f(t, x_\infty^n(t))$ (see equation (6.4)), and the previous computations give us

$$(\int_T \|x_n^*(t)\|^q)^{1/q} \leq n(\int_T \|u_n^*(t)\|^q)^{1/q} + (\int_T \|v_n^*(t)\|^q)^{1/q},$$

and $\| \int_T x_n^*(t)\mathrm{d}\mu(t)\| \leq \| \int_T v_n^*(t)\mathrm{d}\mu(t)\| + \|w_n^*\| \to 0$. By (BP.2) we have that (x_∞^n, y_∞^n) is an ε_n-minimizer of φ_n, and so by Lemma 6.7 we conclude that $n(\|x_\infty^n(t) - y_\infty^n\|_p)^p \to 0$,

$\int_T |f(t, x_\infty^n(t)) - f(t, x_0)| \mathrm{d}\mu(t)$. Finally,

$$\begin{aligned}
\|x_n^*(\cdot)\|_q \cdot \|x_n(\cdot) - y_n\|_p \leq & \left(n(\int_T \|u_n^*(t)\|^q \mathrm{d}\mu(t))^{1/q} + (\int_T \|v_n^*(t)\|^q \mathrm{d}\mu(t))^{1/q} \right) \|x_n(\cdot) - y_n\|_p \\
& \leq nL(\|x_n(\cdot) - y_n\|_p)^{p/q} (\|x_n(\cdot) - y_n\|_p) + \|v_n^*(\cdot)\|_q \cdot \|x_n(\cdot) - y_n\|_p \\
=& nL(\|x_n(\cdot) - y_n\|_p)^p + \|v_n^*(\cdot)\|_q \cdot \|x_n(\cdot) - y_n\|_p \to 0.
\end{aligned}$$

□

Now we establish the two main results of this section. In order to show how to adopt some of the settings available in the literature to our framework, we consider in the following theorems two normal integrands f, g from $T \times X$ to $\overline{\mathbb{R}}$ satisfying the following properties relying on the smoothness of X:

$\mathcal{P}_1$ For all $t \in T$ and all $x \in X$, $f(t, x) \geq g(t, x)$.

$\mathcal{P}_2$ If X is a β-smooth space, the functions g_t are assumed to be β-smooth on X for all $t \in T$, or if X is a Hilbert space, the functions g_t are assumed to be C^2 on X for all $t \in T$.

$\mathcal{P}_3$ If X is a β-smooth space, the integrand I_g is β-smooth on X, or if X is a Hilbert space the integrand I_g is C^2 on X.

Some new results about the study of the subdifferential of a convex integral function (i.e. when $f(t, \cdot)$ is convex ae) used the function $g(t, x) = \langle a^*(t), x \rangle + \alpha(t)$ with $a^* \in L^p_{w^*}(T, X^*)$ and $\alpha \in \mathrm{L}^1(T, \mathbb{R})$ (see [29, 64, 80]).

Theorem 6.11 *Let f, g be two normal integrands satisfying $\mathcal{P}_1$, $\mathcal{P}_2$ and $\mathcal{P}_3$. Consider $x^* \in \partial I_f(x)$ and $p, q \in (1, +\infty)$ with $1/p + 1/q = 1$. If μ is finite, $\sup_{x \in X} \|\nabla_\beta g(\cdot, x)\| \in \mathrm{L}^q(T, \mathbb{R})$ and ∂, X, p satisfy one of the following conditions:*

(i) $\partial = \partial_\beta^-$, *$X$ is a β-smooth space,*

(ii) $\partial = \partial_P$, *X is a Hilbert space and $p \geq 2$,*

then for every w^*-continuous seminorm ρ in X^*, there exist sequences $y_n \in X$, $x_n \in \mathrm{L}^p(T, X)$, $x_n^* \in \mathrm{L}^q_{w^*}(T, X^*)$ such that:

(a) $x_n^*(t) \in \partial f(t, x_n(t))$ *ae.*

(b) $\|x - y_n\| \to 0, \int_T \|x - x_n(t)\|^p \mathrm{d}\mu(t) \to 0,$

(c) $\|x_n^*(\cdot)\|_q \|x_n(\cdot) - y_n\|_p \to 0,$

(d) $\int_T \langle x_n^*(t), x_n(t) - x \rangle \mathrm{d}\mu(t) \to 0,$

(e) $\rho\left(\int_T x_n^*(t) \mathrm{d}\mu(t) - x^* \right) \to 0,$

(f) $\int_T |f(t, x_n(t)) - f(t, x)| \mathrm{d}\mu(t) \to 0.$

Moreover, if one of the conditions of Proposition 6.9 holds, then (e) can be changed to $\| \int_T x_n^*(t) \mathrm{d}\mu(t) - x^* \| \to 0.$

PROOF. First, assume that $g = 0$. Then consider $\varepsilon > 0$ and $\{\mathrm{e_i}\}_{\mathrm{i}=1,\dots,k}$ a finite family of points such that $\rho(\cdot) = \max\{\langle\cdot, \mathrm{e_i}\rangle : \mathrm{i} = 1,..,k\}$, and denote by $L := \mathrm{span}\{x, \mathrm{e_i}\}_{\mathrm{i}=1,\dots,k}$ and $K = L \cap \mathbb{B}(x, 1)$ (if one of the conditions of Proposition 6.9 holds we must proceed taking simply $K = X$). Then there are a ball $\mathbb{B}(x, \eta)$ and a Lipschitz, β-smooth function (a C^2 function if (ii) holds) $\phi : \mathbb{B}(x, \eta) \to \mathbb{R}$ such that $\nabla_\beta \phi(x) = x^*$ and $I_f - \phi$ attains a local minimum at x.

Now we consider the measure space $(\tilde{T}, \tilde{\mathcal{A}}, \tilde{\mu})$, where $\tilde{T} = T \cup \{\omega_1, \omega_2\}$, $\tilde{\mathcal{A}} = \sigma(\mathcal{A}, \{\omega_1\}, \{\omega_2\})$ and $\tilde{\mu}(A) = \mu(A \backslash \{\omega_1, \omega_2\}) + \mathbb{1}_A(\omega_1) + \mathbb{1}_A(\omega_2)$, together with the integrand functional $\tilde{f}(t, x) = f(t, x) + \mathbb{1}_{\{\omega_1\}}(t)\phi(x) + \mathbb{1}_{\{\omega_2\}}(t)\delta_K(x)$. Then condition (a) of Proposition 6.9 holds, and so x is a p-robust minimizer of $I_{\tilde{f}}$ on $\mathbb{B}(x, \eta)$. By Proposition 6.10 there exist sequences $\tilde{y}_n \in X$, $\tilde{x} \in \mathrm{L}^p(\tilde{T}, X)$, $\tilde{x}_n^* \in \mathrm{L}^q_{w^*}(\tilde{T}, X^*)$ (with $1/p + 1/q = 1$) such that:

1. $\tilde{x}_n^*(t) \in \partial \tilde{f}(t, \tilde{x}_n(t))$ ae,
2. $\|x - \tilde{y}_n\| \to 0$, $\displaystyle\int\limits_T \|x - \tilde{x}_n(t)\|^p \mathrm{d}\tilde{\mu}(t) \to 0$,
3. $\|\tilde{x}_n^*(\cdot)\|_q \|\tilde{x}_n(\cdot) - \tilde{y}_n\|_p \to 0$,
4. $\|\displaystyle\int\limits_T \tilde{x}_n^*(t)\mathrm{d}\tilde{\mu}(t)\| \to 0$,
5. $\displaystyle\int\limits_T |f(t, \tilde{x}_n(t)) - f(t, x)|\mathrm{d}\tilde{\mu}(t) \to 0$.

In particular, $\int_{\tilde{T}} \tilde{x}_n^*(t)\mathrm{d}\tilde{\mu}(t)$ is bounded, and so $\langle \int_{\tilde{T}} \tilde{x}_n^*(t)\mathrm{d}\tilde{\mu}(t), \tilde{y}_n - x\rangle \to 0$; hence,

$$\left| \int_{\tilde{T}} \langle \tilde{x}_n^*(t), \tilde{x}_n(t) - x\rangle \mathrm{d}\tilde{\mu}(t) \right| \leq \left| \int\limits_{\tilde{T}} \langle \tilde{x}_n^*(t), \tilde{y}_n - x\rangle \mathrm{d}\tilde{\mu}(t) \right| + \int\limits_{\tilde{T}} \|\tilde{x}_n^*(t)\| \|\tilde{x}_n(t) - \tilde{y}_n\| \mathrm{d}\tilde{\mu}(t) \to 0.$$

Next, define $x_n(t) := \tilde{x}_n(t)$, $x_n^*(t) := \tilde{x}_n^*(t)$ with $t \in T$ and $y_n = \tilde{y}_n$. So, $x_n \in \mathrm{L}^p(T, X)$, $x_n^* \in \mathrm{L}^q_{w^*}(T, X^*)$, $x_n^*(t) \in \partial f(t, x_n(t))$ ae, $\|x - y_n\| \to 0$, $\int_T \|x - x(t)\|^p \mathrm{d}\mu(t) \to 0$, $\|x_n^*(\cdot)\|_q \|x_n(\cdot) - y_n\|_p \to 0$ and $\int_T |f(t, x_n(t)) - f(t, x)|\mathrm{d}\mu(t) \to 0$.

Now $\tilde{x}_n^*(\omega_1) = -\nabla_\beta \phi(x_n(\omega_1)) \xrightarrow{w^*} -\nabla_\beta \phi(x) = -x^*$. By the convexity of K we have that for large enough n, $\tilde{x}_n^*(\omega_2) \in N_K(\tilde{x}_n(\omega_2)) = L^\perp$. Therefore,

$$\begin{aligned}
\rho\left(\int\limits_T x_n^*(t)\mathrm{d}\mu(t) - x^* \right) &\leq \rho\left(\int\limits_T x_n^*(t)\mathrm{d}\mu(t) + \tilde{x^*}_n(\omega_1) + \tilde{x^*}_n(\omega_2) \right) \\
&\quad + \rho(-\tilde{x}_n^*(\omega_1) - x^*) + \rho(\tilde{x}_n^*(\omega_2)) \\
&\leq \| \int\limits_T \tilde{x}_n^*(t)\mathrm{d}\tilde{\mu}(t)\| + \rho(-\tilde{x}_n^*(\omega_1) - x^*) + \rho(\tilde{x}_n^*(\omega_2)) \to 0.
\end{aligned}$$

On the one hand, since $\tilde{x}_n^*(\omega_1)$ is bounded and $x_n(\omega_1) \to x$, we have $\langle \tilde{x}_n^*(\omega_1), x_n(\omega_1) - x\rangle \to 0$. On the other hand, since for large enough n, $\langle \tilde{x}_n^*(\omega_2), x_n(\omega_2) - x\rangle = 0$, we get

$$\begin{aligned}
\left| \int_T \langle x_n^*(t), x_n(t) - x\rangle \mathrm{d}\mu(t) \right| \leq& \left| \int_{\tilde{T}} \langle \tilde{x}_n^*(t), \tilde{x}_n(t) - x\rangle \mathrm{d}\tilde{\mu}(t) \right| + \left| \langle \tilde{x}_n^*(\omega_1), x_n(\omega_1) - x\rangle \right| \\
&+ \left| \langle \tilde{x}_n^*(\omega_2), x_n(\omega_2) - x\rangle \right| \to 0.
\end{aligned}$$

At this step, if g is different than zero, we know by Lemma 6.6 that the gradient of I_g is given by $\int_T \nabla_\beta g_t(x) \mathrm{d}\mu(t)$. Then we must apply the result to the integrand function $h := f - g$, with the gradient $y^* := x^* - \int_T \nabla_\beta g_t(x) \mathrm{d}\mu(t) \in \partial I_h(x)$, and then by making some computations we easily get the result. $\square$

The next Theorem corresponds to the uniform convergence of the measurable functions x_n given in the previous theorem, which can be seen as the corresponding result when case $p = +\infty$.

Theorem 6.12 *Let f, g be two normal integrands satisfying $\mathcal{P}_1$, $\mathcal{P}_2$ and $\mathcal{P}_3$. Consider $x^* \in \partial I_f(x)$. We assume that $\sup_{x \in X} \|\nabla_\beta g(\cdot, x)\| \in \mathrm{L}^1(T, \mathbb{R})$ and ∂, X satisfy one of the following conditions:*

(i) $\partial = \partial_\beta^-$, *$X$ is a β-smooth space,*

(ii) $\partial = \partial_P$, *X is a Hilbert space.*

Then for every w^*-continuous seminorm ρ in X^*, there exist sequences $x_n \in \mathrm{L}^\infty(T, X)$, $x_n^* \in \mathrm{L}^1_{w^*}(T, X^*)$ such that

(a) $x_n^*(t) \in \partial f(t, x_n(t))$ *ae.*

(b) $\|x - y_n\| \to 0, \|x - x_n(\cdot)\|_\infty \to 0.$

(c) $\displaystyle\int_T \|x_n^*(t)\| \|x_n(t) - y_n\| \mathrm{d}\mu(t) \to 0.$

(d) $\displaystyle\int_T \langle x_n^*(t), x_n(t) - x \rangle \mathrm{d}\mu(t) \to 0.$

(e) $\displaystyle\rho\Big(\int_T x_n^*(t) \mathrm{d}\mu(t) - x^*\Big) \to 0.$

(f) $\displaystyle\int_T |f(t, x_n(t)) - f(t, x)| \mathrm{d}\mu(t) \to 0.$

Moreover, if one of the conditions of Proposition 6.9 holds, then (e) can be changed by $\displaystyle\|\int_T x_n^*(t) \mathrm{d}\mu(t) - x^*\| \to 0.$

Proof. Consider ρ and $x^* \in \partial I_f(x)$ as in the statement. First we assume that μ is finite and $g = 0$, and so we have that $f(t, x) \geq 0$ for all $t \in T$ and all $x \in X$. Let $\varepsilon \in (0, 1)$ and define $\tilde{f}(t, x') := f(t, x') + \delta_{\mathbb{B}(x, \varepsilon)}(x')$. It follows that $x^* \in \partial I_{\tilde{f}}(x)$. Then by Theorem 6.11 there exist measurable functions $\tilde{x}_n \in \mathrm{L}^2(T, X)$, $\tilde{x}_n^*(t) \in \mathrm{L}^2_{w^*}(T, X^*)$ such that:

1. $\tilde{x}_n^*(t) \in \partial \tilde{f}(t, \tilde{x}_n(t))$ ae,
2. $\|x - \tilde{y}_n\| \to 0, \displaystyle\int_T \|x - \tilde{x}_n(t)\|^2 \mathrm{d}\mu(t) \to 0,$
3. $\|\tilde{x}_n^*(\cdot)\|_2 \|\tilde{x}_n(\cdot) - \tilde{y}_n\|_2,$
4. $\rho(\displaystyle\int_T \tilde{x}_n^*(t) \mathrm{d}\mu(t) - x^*) \to 0,$
5. $\displaystyle\int_T |f(t, \tilde{x}_n(t)) - f(t, x)| \mathrm{d}\mu(t) \to 0.$

It is easy to see that if $\|\tilde{x}_n(t) - x\| < \varepsilon$, then $\tilde{x}_n^*(t) \in \partial f(t, \tilde{x}_n(t))$. Let $L = L(X, \beta)$ be as in Lemma 6.8, and define the measurable set $A_n := \{ t \in T : \| \tilde{x}_n(t) - x\| = \varepsilon\}$.

The convergence in $\mathrm{L}^2(T,X)$ implies that $\mu(A_n)\to 0$. Now take $n\in\mathbb{N}$ such that $\|x-\tilde{y}_n\|\leq\varepsilon/2$, $\rho(\int\limits_T\tilde{x}_n^*(t)\mathrm{d}\mu(t)-x^*)\leq\varepsilon/3$, $\int_T\|\tilde{x}_n^*(t)\|\|\ \tilde{y}_n-\tilde{x}_n(t)\|\mathrm{d}\mu(t)\leq\varepsilon^2/6$, $\int_T|f(t,\tilde{x}_n(t))-f(t,x)|\mathrm{d}\tilde{\mu}(t)\leq\varepsilon/2$, $\displaystyle\int_T\langle x_n^*(t),x_n(t)-x\rangle\mathrm{d}\mu(t)\leq\varepsilon/3$ and $\int_{A_n}f(t,x)\mathrm{d}\mu(t)\leq\varepsilon^2/6(L+1)$. It follows that

$$\begin{aligned}\frac{\varepsilon^2}{6}&\geq\int_T\|\tilde{x}_n^*(t)\|\|\ \tilde{y}_n-\tilde{x}_n(t)\|\mathrm{d}\mu(t)\geq\int_{A_n}\|\tilde{x}_n^*(t)\|\|\ \tilde{y}_n-\tilde{x}_n(t)\|\mathrm{d}\mu(t)\\&\geq\int_{A_n}\|\tilde{x}_n^*(t)\|\big\{\|\ x-\tilde{x}_n(t)\|-\|\ x-\tilde{y}_n\|\big\}\mathrm{d}\mu(t)\geq\int_{A_n}\|\tilde{x}_n^*(t)\|\{\varepsilon-\frac{\varepsilon}{2}\}\mathrm{d}\mu(t)\geq\frac{\varepsilon}{2}\int_{A_n}\|\tilde{x}_n^*(t)\|\mathrm{d}\mu(t).\end{aligned}$$

Therefore, $\int_{A_n}\|\tilde{x}_n^*(t)\|\mathrm{d}\mu(t)\leq\frac{\varepsilon}{3}$. Now define $\varepsilon(t):=f(t,x)$. By the nonnegativity of the integrand, we have that x is a $\varepsilon(t)$-minimum of $f(t,\cdot)$ for almost all $t\in A_n$. Then by Lemma 6.8 there exist measurable functions $(y(t),y^*(t))\in X\times X^*$ such that for almost all $t\in A_n$, $y^*(t)\in\partial f(t,y(t))$, $\|y(t)-x\|\leq\varepsilon/2$, $|f(t,y(t))-f(t,x)|\leq\varepsilon(t)$ and $\|\ y^*(t)\|\leq 2L\varepsilon(t)/\varepsilon$.

Then we define $x(t):=\tilde{x}_n(t)\mathbb{1}_{A_n^c}(t)+y(t)\mathbb{1}_{A_n}$ and $x^*(t):=\tilde{x}_n^*(t)\mathbb{1}_{A_n^c}(t)+y^*(t)\mathbb{1}_{A_n}$. Hence, $x^*(t)\in\partial f(t,x(t))$ ae, $\|x-x(\cdot)\|_\infty\leq\varepsilon$,

$$\begin{aligned}\int_T\|y^*(t)\|\mathrm{d}\mu(t)&=\int_{A_n^c}\|\tilde{x}_n^*(t)\|\mathrm{d}\mu(t)+\int_{A_n}\|y^*(t)\|\mathrm{d}\mu(t)\\&\leq\mu(T)^{1/2}\|\tilde{x}_n^*(\cdot)\|_2+\varepsilon/2,\int_T|f(t,x(t))-f(t,x)|\mathrm{d}\mu(t)\\&=\int_{A_n^c}|f(t,\tilde{x}_n(t))-f(t,x)|\mathrm{d}\mu(t)+\int_{A_n}|f(t,y(t))-f(t,x)|\mathrm{d}\mu(t)\leq\varepsilon.\end{aligned}$$

Furthermore, $\rho(\int_T x^*(t)\mathrm{d}\mu(t)-x^*)\leq\rho(\int_T\tilde{x}_n^*(t)\mathrm{d}\mu(t)-x^*)+\int_{A_n}\|\tilde{x}_n^*(t)\|\mathrm{d}\mu(t))+\int_{A_n}\|y^*(t)\|\mathrm{d}\mu(t)\leq\varepsilon$, so that

$$\begin{aligned}\int\limits_T\|x^*(t)\|\|\tilde{y}_n-x(t)\|\mathrm{d}\mu(t)&=\int\limits_{A_n^c}\|\tilde{x}^*(t)\ \|\ \|\tilde{y}_n-\tilde{x}_n(t)\|\mathrm{d}\mu(t)+\int_{A_n}\|y^*(t)\ \|\big(\|\tilde{y}_n-x\|+\|y(t)-x\|\big)\mathrm{d}\mu(t)\\&\leq\varepsilon^2/6+\varepsilon^2/6+\varepsilon^2/6\leq\varepsilon.\end{aligned}$$

Finally,

$$\begin{aligned}\Big|\int_T\langle x^*(t),x(t)-x\rangle\mathrm{d}\mu(t)\Big|&\leq\Big|\int_{A_n^c}\langle\tilde{x}_n^*(t),\tilde{x}_n(t)-x\rangle\mathrm{d}\mu(t)+\int_{A_n}\langle y^*(t),y(t)-x\rangle\mathrm{d}\mu(t)\Big|\\&\leq\Big|\int_T\langle\tilde{x}_n^*(t),\tilde{x}_n(t)-x\rangle\mathrm{d}\mu(t)\Big|+\int_{A_n}\|\tilde{x}_n^*(t)\|\cdot\|\tilde{x}_n(t)-x\|\mathrm{d}\mu(t)\\&\quad+\int_{A_n}\|y^*(t)\|\cdot\|y(t)-x\|\mathrm{d}\mu(t)\leq\varepsilon.\end{aligned}$$

Now if μ is σ-finite, consider $\nu(\cdot)=\int_\cdot k(t)\mathrm{d}\mu(t)$, where $k>0$ is integrable and consider the integrand $\tilde{f}(t,x)=f(t,x)/k(t)$. So, $I_{\tilde{f}}^\nu=I_f^\mu$, and then by applying the previous Theorem we easily get the result. The general case, when g is not zero, follows the same arguments given in the proof of Theorem 6.11. □

Remark 6.13 The reader can easily notice that one can modify the measurable selection $\{x_n, x_n^*\}$ in Theorem 6.11 for the case X is a Hilbert space, using a similar technique as in the Theorem above. Indeed it is enough to repeat the proof of the Theorem above and use Lemma 6.8 to derive the existence of measurable functions $(y(t), y^*(t)) \in X \times X^*$ such that for almost all $t \in A_n$, $y^*(t) \in \partial f(t, y(t))$, $\|y(t) - x\| \leq \varepsilon(t)^{1/p}$, $|f(t, y(t)) - f(t, x)| \leq \varepsilon(t)$ and $\| y^*(t)\| \leq L\varepsilon^{1-1/p}(t)$. This gives us a function $x(t) := \tilde{x}_n(t)\mathbb{1}_{A_n^c}(t) + y(t)\mathbb{1}_{A_n}$ satisfying the conclusion of Theorem 6.12 with an arbitrary $p \in (1, 2)$.

In the next result we apply techniques of *separable reduction* for the case of Fréchet subdifferential of infinite series. For this propose we introduce the concept of a *rich family*. The symbol $S(X \times X^*)$ denotes the family of set $U \times Y$ where U and Y are (norm-) separable closed linear subspaces of X and X^*. A set $\mathcal{R} \subseteq S(X \times X^*)$ is called a *rich family* if for every $U \times Y \in S(X \times X^*)$, there exists $V \times Z \in \mathcal{R}$ such that $U \subseteq V$, $Y \subseteq Z$ and $\overline{\bigcup_{n\in\mathbb{N}} U_n} \times \overline{\bigcup_{n\in\mathbb{N}} Y_n} \in \mathcal{R}$ whenever the sequence $(U_n \times Y_n)_{n\in\mathbb{N}} \subseteq \mathcal{R}$ satisfy $U_n \subseteq U_{n+1}$ and $Y_n \subseteq Y_{n+1}$ (for more details see [36, 37, 47] and the reference therein). In [36, Theorem 3.1] the authors showed that there exist a rich family in (rather non-separable) Asplund space.

Proposition 6.14 *Let $(X, \|\cdot\|)$ be an Asplund space and let $f : X \to \mathbb{R} \cup \{+\infty\}$ be any proper function. Then there exists a rich family $\mathcal{R} \subseteq S(X \times X^*)$ such that $Y_1 \subseteq Y_2$ whenever $V_1 \times Y_1, V_2 \times Y_2 \in \mathcal{R}$ and $V_1 \subseteq V_2$, with further properties that for every $V \times Y \in \mathcal{R}$, the assignment $Y \ni x^* \to x^*_{|V} \in V^*$ is a surjective isometry from Y to V^*, and for every $v \in V$ we have that*

$$(\partial_F f(v) \cap Y)_{|V} = (\partial_F f(v))_{|V} = \partial_F f_{|V}(v); \tag{6.5}$$

that is, in more detail, if $v^* \in \partial_F f_{|V}(v)$, there exists a unique $x^* \in \partial_F f(v) \cap Y$ such that $x^*_{|V} = v^*$ and $\|x^*\| = \|v^*\|$.

Besides, it has been proved that *intersection of countably many rich families of a given space is (not only non-empty but even) rich* (see [10, Proposition 1.1] or [47, Proposition 1.2]). Then for the case $(T, \mathcal{A}) = (\mathbb{N}, \mathcal{P}(\mathbb{N}))$ there must exists a rich family $\mathcal{R}$ for the integrand function (f_n), satisfying the properties of Proposition 6.14 and with (6.5) uniformly for every $n \in \mathbb{N}$, as well as for the integral functional I_f. Using this family, we can extend all the previous statements to arbitrary Asplund spaces in the case when $(T, \mathcal{A}) = (\mathbb{N}, \mathcal{P}(\mathbb{N}))$.

Corollary 6.15 *The statement of Theorem 6.11 and 6.12 hold if we assume that X is an Asplund space and $(T, \mathcal{A}) = (\mathbb{N}, \mathcal{P}(\mathbb{N}))$.*

Proof. Assume that the assumptions on f, g in Theorem 6.11 hold in an Asplund space X (respectively the assumptions in Theorem 6.12). Let $x^* \in \partial_F I_f(x)$ and ρ be a w^*-continuous seminorm on X^*; for instance, $\rho = \max_{\mathrm{i}=1,\ldots,p}\langle \cdot, \mathrm{e_i}\rangle$ with some $\mathrm{e_i} \in X$. Then consider $V \times Y \in \mathcal{R}$ such that $x, \mathrm{e_i} \in V$, $\mathrm{i} = 1, \ldots, p$ and $x^* \in Y$. Then $x^*_{|V} =: y^* \in \partial(I_f)|_V(x)$. Because the space V is separable and V has separable dual V^* (because X is and Asplund space), V has a Fréchet smooth renorm. Then $\partial_F^- = \partial_F$ and consequently, applying Theorem 6.11, there exist sequences $y_n \in U$, $x_n \in \mathrm{L}^p(T, V)$, $z_n^* \in \mathrm{L}^q_{w^*}(T, V^*)$ such that:

(a) $z_n^*(t) \in \partial_F f_{|V}(t, x_n(t))$ ae.

(b) $\|x - y_n\| \to 0, \displaystyle\int_T \|x - x_n(t)\|^p \mathrm{d}\mu(t) \to 0,$

(c) $\|z_n^*(\cdot)\|_q \|x_n(\cdot) - y_n\|_p \to 0$, ($\displaystyle\int_T \|x_n^*(t)\| \|x_n(t) - y_n\| \mathrm{d}\mu(t) \to 0$ resp.)

(d) $\displaystyle\int_T \langle z_n^*(t), x_n(t) - x\rangle \mathrm{d}\mu(t) \to 0,$

(e) $\rho\left(\displaystyle\int_T z_n^*(t)\mathrm{d}\mu(t) - x^*\right) \to 0,$

(f) $\displaystyle\int_T |f(t, x_n(t)) - f(t, x)| \mathrm{d}\mu(t) \to 0.$

Then define $x_n^*(t)$ as the unique element in $\partial f(t, x_n(t)) \cap Y$ such that $\|x_n^*(t)\| = \|z_n^*(t)\|$ and $(x_n^*(t))_{|V} = z_n^*(t)$. Now $\|y_n^*(\cdot)\|_q = \|x_n^*(\cdot)\|_q$, which implies $\|x_n^*(\cdot)\|_q \|x_n(\cdot) - y_n\|_p \to 0$ (respectively $\displaystyle\int_T \|x_n^*(t)\| \|x_n(t) - y_n\| \mathrm{d}\mu(t) \to 0$) and $x_n^* \in \mathrm{L}^q(T, X^*)$. From the fact that $x_n(t), y_n, \mathrm{e}_i, x \in V$ and $(x_n^*(t))_{|V} = z_n^*(t)$ we conclude that $\int_T \langle x_n^*(t), x_n(t) - x\rangle \mathrm{d}\mu(t) \to 0$ and $\rho\left(\int_T x_n^*(t)\mathrm{d}\mu(t) - x^*\right) \to 0$. Then, the sequences y_n, $x_n(\cdot)$ and $x_n^*(\cdot)$ satisfy the required properties.

Finally, if one of the conditions of Proposition 6.9 holds we notice that for each n $\int_T x_n^*(t)\mathrm{d}\mu(t) = \sum_{k\in\mathbb{N}} x_n^*(k)\mathrm{d}\mu(k) \in Y^*$ (because Y^* is linear closed space). Then the norm of $\int_T x_n^*(t)\mathrm{d}\mu(t)$ must be equal to $(\int_T x_n^*(t)\mathrm{d}\mu(t))_{|V} = \int_T y_n^*(t)\mathrm{d}\mu(t)$ and it goes to zero if one of the conditions of Proposition 6.9 holds. $\square$

The final result corresponds to an extension of [80, Corollary 1.2.1] to the case $p = +\infty$.

Corollary 6.16 *In the setting of Theorem 6.12 assume that f is a convex normal integrand (i.e. f_t is convex for all $t \in T$). Then one has $x^* \in \partial f(x)$ if and only if there are nets $x_\nu \in \mathrm{L}^\infty(T, X)$ and $x_\nu^* \in \mathrm{L}^1(T, X^*)$ such that $x_\nu^*(t) \in \partial f(t, x_\nu(t))$ ae, $\|x - x_\nu(\cdot)\|_\infty \to 0$, $\int_T x_\nu^*(t)\mathrm{d}\mu(t) \xrightarrow{w^*} x^*$, $\int_T \langle x_\nu^*(t), x_\nu(t) - x\rangle \mathrm{d}\mu(t) \to 0$ and $\int_T |f(t, x_\nu(t)) - f(t, x)| \mathrm{d}\mu(t) \to 0$. If the space X is reflexive we can take sequences instead of nets and the w^*-convergence of $\int_T x_\nu^*(t)\mathrm{d}\mu(t)$ is in norm topology.*

Proof. When X is reflexive, without loss of generality, we can assume that criterion (a) of Proposition 6.9 is satisfied; otherwise, we take $\tilde{f}_t := f_t + I_{\mathbb{B}(x,1)}$. So the construction of the sequence follows similar and classical arguments which we are going to give in the next paragraphs. First assume $x^* \in \partial I_f(x)$. Then take $\mathcal{N}_0$ the neighborhood system of zero for the w^*-topology and consider the ordered set $\mathbb{A} := \mathbb{N} \times \mathcal{N}_0$ as $(n_1, U_1) \leq (n_2, U_2)$ if and only if $n_1 \leq n_2$ and $U_2 \subseteq U_1$. Then by Theorem 6.12 for every $\nu = (n, U)$ there must be $x_\nu \in \mathrm{L}^\infty(T, X)$ and $x_\nu^* \in \mathrm{L}^1_{w^*}(T, X^*)$ such that

1. $x_\nu^*(t) \in \partial f(t, x_\nu(t))$ ae,
2. $\|x - x_\nu(\cdot)\|_\infty \leq 1/n$,
3. $\displaystyle\int_T \langle x_\nu^*(t), x_\nu(t) - x\rangle \mathrm{d}\mu(t) \to 1/n$,
4. $\int x_\nu^*(t)\mathrm{d}\mu(t) - x^* \in U$,
5. $\int_T |f(t, x_\nu(t)) - f(t, x)| \mathrm{d}\mu(t) \to 0$.

Hence the net (x_ν, x_ν^*) satisfies the required properties. Conversely, assume that the net

(x_ν, x_ν^*) satisfies the above properties. Then for all $y \in X$

$$\begin{aligned}\langle x^*, y - x\rangle \leq & \langle x^* - \int_T x_\nu^*(t)\mathrm{d}\mu(t), y - x\rangle + \int_T \langle x_\nu^*(t), y - x_\nu(t)\rangle \mathrm{d}\mu(t) \\ & + \int_T \langle x_\nu^*(t), x_\nu(t) - x\rangle \mathrm{d}\mu(t) \\ \leq & \langle x^* - \int_T x_\nu^*(t)\mathrm{d}\mu(t), y - x\rangle + I_f(y) - \int_T f(t, x_\nu(t)) \\ & + \int_T \langle x_\nu^*(t), x_\nu(t) - x\rangle \mathrm{d}\mu(t).\end{aligned}$$

So, taking the limits we conclude $\langle x^*, y - x\rangle \leq I_f(y) - I_f(x)$, and from the arbitrarily of y we get the result. $\square$

Remark 6.17 It important to notice that all the result above does not required any smoothness if we assume that f is a convex normal integral and $(T, \mathcal{A}) = (\mathbb{N}, \mathcal{P}(\mathbb{N}))$. Moreover, in this case one can simplify the proof of Lemma 6.8 and Theorem 6.10 applying Ekeland's variational principle (see, e.g., [45, Theorem 1]) instead of Borwein-Preiss variational principle. Also one can show the result using [47, Theorem 4.1], where the authors claim that for any proper function $g : X \to \mathbb{R}$ there exists a rich family $\mathcal{R}$ with the property that for every $U \in \mathcal{R}$ and $\bar{x} \in U$ one has $x^* \in \partial_F g(\bar{x})$ and $\|x^*\| \leq c$ whenever there exists $y^* \in \partial(f_{|_Y})(\bar{x})$ and $\|y^*\| \leq c$. Then, the proof follow similar arguments than Corollary 6.15. However, we refused the uses of this family in Corollary 6.15, because it does not give us an comparison between the norm of $\int_T z_n^*(t)\mathrm{d}\mu(t)$ and $\int_T x_n^*(t)\mathrm{d}\mu(t)$, which is necessary to prove the convergence in norm topology of $\int_T x_n^*(t)\mathrm{d}\mu(t)$ when one of the conditions of Proposition 6.9 holds.

Remark 6.18 It is worth comparing the results given in [80, Theorem 1.4.2] with Corollary 6.16, where we have proved similar results to the convex case. The advantage of our approach is that we do not need to use *singular elements* in the dual of $L^\infty(T, X)$ to characterize the subdifferential of I_f.

6.2 Limiting/Mordukhovich, G- and Clarke-Rockafellar subdifferentials

The aim of this section is to establish an upper-estimation for the Mordukhovich, G- and Clarke-Rockafellar subdifferentials at a point $x \in \operatorname{dom} I_f$, in terms of the corresponding subdifferential of the data function f_t at the same point. For simplicity, we will focus on the case when X is a separable Banach space. For this reason, in the sequel we adopt the following notation: If X is an F-smooth space (i.e. when $(X^*, \|\cdot\|)$ is also separable) $\partial = \partial_F^-$, $\partial_L = \partial_M$, $\partial_L^\infty = \partial_M^\infty$, otherwise $\partial = \partial_H^-$, $\partial_L = \tilde{\partial}_G^\infty$, $\partial_L^\infty = \tilde{\partial}_G^\infty$.

Lemma 6.19 *Consider $x^* \in \partial_L I_f(x)$, $y^* \in \partial_{\mathrm{L}}^\infty I_f(x_0)$, a finite family of linearly independent points $\{\mathrm{e_i}\}_{\mathrm{i}=1}^p$, $W := \operatorname{span}\{\mathrm{e_i}\}$ and $\rho(\cdot) := \max\{|\langle \cdot, \mathrm{e_i}\rangle|\}$. Then there exist sequences $x_n, y_n \in \mathrm{L}^\infty(T, X)$, $x_n^*, y_n^* \in \mathrm{L}^1_{w^*}(T, X^*)$ and $\lambda_n \to 0^+$ such that:*

(a) $x_n^*(t) \in \partial f(t, x_n(t))$ *ae,*

(b) $\|x - x_n(\cdot)\|_\infty \to 0$,

(c) $\rho(\int_T x_n^*(t) \mathrm{d}\mu(t) - x^*) \to 0$,

(d) $\lim \int_T |f(t, x_n(t)) - f(t, x)| \mathrm{d}\mu(t) \to 0$,

(a^∞) $y_n^*(t) \in \partial f(t, y_n(t))$ *ae,*

(b^∞) $\|x - y_n(\cdot)\|_\infty \to 0$,

(c^∞) $\rho(\lambda_n \cdot \int_T y_n^*(t) \mathrm{d}\mu(t) - y^*) \to 0$,

(d^∞) $\lim \int_T |f(t, y_n(t)) - f(t, x)| \mathrm{d}\mu(t) \to 0$,

(e^∞) $y^* \in (B(y_n, y_n^*) \cap W)^-$,

where

$$B(y_n, y_n^*) := \{ \ \xi \in X : \liminf \int_T \langle y_n^*, \xi \rangle^+ \mathrm{d}\mu < +\infty \}.$$

Moreover, if there exists a bounded sequence x_n^* *(in* $\mathrm{L}^1_{w^*}(T, X^*)$*) (or* $\lambda_n y_n^*$ *respectively) satisfying the above properties, then*

$$x^* \in \int_T \partial_L f(t, x) \mathrm{d}\mu(t) + C((x_n, x_n^*))^- + F^\perp,$$

$$\textit{(respectively) } y^* \in \int_T \partial_{\mathrm{L}}^\infty f(t, x) \mathrm{d}\mu(t) + C((y_n, \lambda_n y_n^*))^- + F^\perp,$$

where $C(x_n, x_n^*) := \{ \ \xi \in X : (\langle x_n^*(\cdot), \xi \rangle^+)_{n \in \mathbb{N}}$ *is uniformly integrable*$\}$.

Proof. By definition of $\partial_L I_f(x)$ and $\partial_L^\infty I_f(x)$ there exist sequences $z_n^* \in \partial I_f(z_n)$, $s_n^* \in \partial I_f(s_n)$ and $\lambda_n \to 0^+$ such that $z_n, s_n \xrightarrow{I_f} x$, $z_n^* \xrightarrow{w^*} x^*$ and $\lambda_n \cdot s_n^* \xrightarrow{w^*} y^*$. By Theorem 6.12 (and using a diagonal argument) there exist sequences $x_n, y_n \in \mathrm{L}^\infty(T, X)$, $x_n^*, y_n^* \in \mathrm{L}^1_{w^*}(T, X^*)$ such that

(i) $x_n^*(t) \in \partial f(t, x_n(t))$ ae,

(ii) $\|x - x_n(\cdot)\|_\infty \to 0$,

(iii) $\rho(\int_T x_n^*(t) \mathrm{d}\mu(t) - x^*) \to 0$,

(iv) $\int_T f(t, x_n(t)) \mathrm{d}\mu(t) \to \int_T f(t, x) \mathrm{d}\mu(t)$,

(i^∞) $y_n^*(t) \in \partial f(t, y_n(t))$ ae,

(ii^∞) $\|x - y_n(\cdot)\|_\infty \to 0$,

(iii^∞) $\rho(\lambda_n \cdot \int_T y_n^*(t) \mathrm{d}\mu(t) - x^*) \to 0$,

(iv^∞) $\int_T f(t, y_n(t)) \mathrm{d}\mu(t) \to \int_T f(t, x) \mathrm{d}\mu(t)$.

So, by Lemma 6.5 we conclude that $\lim \int_T |f(t, x_n(t)) - f(t, x)| \mathrm{d}\mu(t) = 0$ and $\lim \int_T |f(t, y_n(t)) - f(t, x)| \mathrm{d}\mu(t) = 0$. Moreover if we take $\xi \in B(y_n, y_n^*) \cap W$, then $\langle y^*, \xi \rangle = \lim \lambda_n \int \langle y_n^*, \xi \rangle \mathrm{d}\mu(t) \le \liminf \lambda_n \int \langle y_n^*, \xi \rangle^+ = 0$; this proves the first part. To prove the second part, consider a continuous projection $P_W : X \to W$, and assume that $\sup_n \int_T \|x_n^*(t)\| \mathrm{d}\mu(t) < \infty$. So, by [4, Corollary

4.1] we have that

$$\mathrm{Ls}^{w^*}\{\int_T x_n^*(t)\mathrm{d}\mu(t)\} \subseteq \int_T \mathrm{Ls}^{w^*}\{x_n^*(t)\}\mathrm{d}\mu(t) + C(x_n, x_n^*)^- + F^\perp,$$

where $\mathrm{Ls}^{w^*}\{x_n^*(t)\}$ represents the sequential upper limit of the sequence $(x_n^*(t))$. Moreover, if $\sup_n \int_T \|x_n^*(t)\|\mathrm{d}\mu(t)$ is finite, then (under subsequence) $w_n^* := \int_T x_n^*(t)\mathrm{d}\mu(t) \to w_0^*$. But $w_n^* = P_W^*(w_n^*) + w_n^* - P_W^*(w_n^*)$ and by (iii) we get that $P_W^*(w_n^*) \to P_W^*(x^*)$. Therefore, $w_0^* - P_W^*(x^*) \in W^\perp$, and then we conclude that $x^* = w_0^* + P_W^*(x^*) - w_0^* + x^* - P_W^*(x^*) \in \int_T \mathrm{Ls}^{w^*}\{x_n^*(t)\}\mathrm{d}\mu(t) + C((x_n, x_n^*)^- + W^\perp$. Finally, we have to prove that almost everywhere $\mathrm{Ls}^{w^*}\{x_n^*(t)\} \subseteq \partial_L f(t,x)$. Indeed, from the previous convergence (and under a subsequence, see e.g. [1, Theorem 13.6]) we can take a measurable set $\tilde{T}$ such that $\mu(T\backslash\tilde{T}) = 0$ and for every $t \in \tilde{T}$, $x_n(t) \to x_0$ and $f(t, x_n(t)) \to f(t,x)$. Then take an integrable selection $a^*(t) \in \mathrm{Ls}^{w^*}\{x_n^*(t)\}$ and fix $t_0 \in \tilde{T}$. So, there exists a subsequence $x^*_{n_{k(t_0)}}(t_0) \to a^*(t_0)$, and then $x_{n_{k(t_0)}}(t_0) \to x_0$ and $f(t_0, x_{n_{k(t_0)}}) \to f(t,x)$. Hence, $a^*(t_0) \in \partial_L f(t_0, x)$. The case for the point y^* is similar, and so we omit the proof.

□

Considering the above result we need to ensure the boundedness of approximate sequences in the subdifferential to establish an exact upper estimate. So, the next part concerns criteria to guarantee this property. For this reason, we introduce the following definitions that allow us to extend the classical results, which consider locally Lipschitz continuity of the integral functional (see for instance [23, Theorem 2.7.2] or [86]).

We adopt the definition of w^*-compact sole (see [24, Proposition 2.1]) in the integral sense as follows.

Definition 6.20 (Integrable compact sole) *Consider a measurable multifunction $C : T \rightrightarrows X^*$ with non-empty closed values. We say that C has an integrable compact sole if and only if there exist $\mathrm{e} \in X$ and $\delta > 0$ such that for every measurable selection c^* of M*

$$\delta\langle c^*(t), \mathrm{e}\rangle \geq \|c^*(t)\| \text{ ae.}$$

Moreover, we denote $UI(C) := \{\xi \in X : \sigma_{C(\cdot)}(\xi)^+ \in \mathrm{L}^1\}$.

Theorem 6.21 *Let $x \in \operatorname{dom} I_f$ and suppose there exist $\delta > 0$, a measurable multifunction $C : T \rightrightarrows X^*$ which has an integrable compact sole, and an integrable function $K(\cdot) > 0$ such that*

$$\partial f(t, x') \subseteq K(t)\mathbb{B}(0,1) + C(t), \forall x' \in \mathbb{B}(x,\delta),\ \forall t \in T \tag{6.6}$$

Then

$$\partial_L I_f(x) \subseteq \bigcap \Big\{ \int_T \partial_L f(t,x)\mathrm{d}\mu(t) + UI(C)^- + W^\perp \Big\}, \tag{6.7}$$

$$\partial_L^\infty I_f(x) \subseteq \bigcap \Big\{ \int_T \partial_{\mathrm{L}}^\infty f(t,x)\mathrm{d}\mu(t) + UI(C)^- + W^\perp \Big\}, \tag{6.8}$$

where the intersection is over all finite dimensional subspaces $W \subseteq X$. Consequently,

$$\partial^C I_f(x) \subseteq \overline{\text{co}}^{w^*} \left\{ \int\limits_T \partial_L f(t,x)\mathrm{d}\mu + \int_T \partial_{\mathrm{L}}^{\infty} f(t,x)\mathrm{d}\mu(t) + UI(C)^- \right\}.$$

Proof. Let $x^* \in \partial_L I_f(x)$ and $y^* \in \partial_{\mathrm{L}}^{\infty} I_f(x)$. Consider a finite family of linearly independent points $\{\mathrm{e_i}\}_{\mathrm{i}=1}^p$, $W := \text{span}\{\mathrm{e_i}\}$ and $\rho(\cdot) := \max\{|\langle \cdot, \mathrm{e_i}\rangle|\}$. Then by Lemma 6.19 there exist sequences $x_n, y_n \in \mathrm{L}^{\infty}(T,X)$, $x_n^*, y_n^* \in \mathrm{L}^1_{w^*}(T,X^*)$ and $\lambda_n \to 0^+$ such that:

(i) $x_n^*(t) \in \partial f(t, x_n(t))$ ae,

(ii) $\|x - x_n(\cdot)\|_{\infty}\mathrm{d}\mu(t) \to 0$,

(iii) $\rho(\int_T x_n^*(t)\mathrm{d}\mu(t) - x^*) \to 0$,

(iv) $\lim \int_T |f(t,x_n(t)) - f(t,x)|\mathrm{d}\mu(t) \to 0$,

(i$^\infty$) $y_n^*(t) \in \partial f(t, y_n(t))$ ae,

(ii$^\infty$) $\|x_0 - y_n(\cdot)\|_{\infty} \to 0$,

(iii$^\infty$) $\rho(\lambda_n \cdot \int_T y_n^*(t)\mathrm{d}\mu(t) - y^*) \to 0$,

(iv$^\infty$) $\lim \int_T |f(t,y_n(t)) - f(t,x_0)|\mathrm{d}\mu(t) \to 0$.

Hence, (for large enough n) relation (6.6) implies that $x_n^*(t) \in K(t)\mathbb{B}(0,1) + C(t)$ and $y_n^*(t) \in K(t)\mathbb{B}(0,1)+C(t)$ ae. Therefore, by the measurability of the involved functions in the last inclusion there are measurable selections $h_n^1(t), h_n^2(t) \in B(0,K(t))$ and $c_n^1(t), c_n^2(t) \in C(t)$ such that $x_n^*(t) = h_n^1(t) + c_n^1(t)$ and $y_n^*(t) = h_n^2(t) + c_n^2(t)$ (see [22, Theorem III.22]). From the fact that C has an integrable compact sole there exist $\mathrm{e} \in X$ and $\delta > 0$ such that $\|c_n^{\mathrm{i}}\| \leq \delta\langle c_n^{\mathrm{i}}, \mathrm{e}\rangle$ for $\mathrm{i} = 1,2$. Then

$$\int_T \| x_n^*\|\mathrm{d}\mu \leq \int_T K\mathrm{d}\mu + \delta \int_T \langle c_n^1, \mathrm{e}\rangle\mathrm{d}\mu = \int_T K\mathrm{d}\mu + \delta\Big(\int_T \langle x_n^*, \mathrm{e}\rangle\mathrm{d}\mu - \int_T \langle h_n^1, \mathrm{e}\rangle\mathrm{d}\mu\Big) \tag{6.9}$$

$$\leq (1 + \| \mathrm{e}\|) \int_T K\mathrm{d}\mu + \delta \int_T \langle x_n^*, \mathrm{e}\rangle\mathrm{d}\mu. \tag{6.10}$$

So, assuming that $\mathrm{e} \in W$, we have that the sequence x_n^* is bounded in $\mathrm{L}^1_{w^*}(T,X^*)$ and obviously the same holds for the sequence y_n^*. Then, by Lemma 6.19, $x^* \in \int\limits_T \partial_L f(t,x)\mathrm{d}\mu(t)+ UI(C)^- + W^{\perp}$ and $y^* \in \int\limits_T \partial_L^{\infty} f(t,x)\mathrm{d}\mu(t) + UI(C)^- + W^{\perp}$. The final formula follows from [84, Theorem 3.57] or Proposition 1.5 depending on the smoothness of X. □

The following lemma allows us to understand the definition of integrable w^*-compact sole in terms of the primal space instead of the dual space, using the polar cone.

Lemma 6.22 *Let $C : T \rightrightarrows X$ be a measurable multifunction with non-empty w-closed convex values, let $\mathrm{e} \in X$ and let $\delta > 0$. Then the following statement are equivalent:*

(a) For every measurable selection c^* of $C^-(t)(:= (C(t))^-)$ one has $\delta^{-1}\langle c^*(t), -\mathrm{e}\rangle \geq \|c^*(t)\|$ ae.

(b) The ball in $L^{\infty}(T,X)$ around e with radius δ is contained in $\{x(\cdot) \in \mathrm{L}^{\infty}(T,X) : x(t) \in C(t)$ ae$\}$.

Proof. First by Castaing's representation there exist sequences of measurable selections c_n and c_n^* of C and C^- respectively such that $C(t) = \overline{\{c_n(t)\}}^{\|\cdot\|}$ and $C^-(t) = \overline{\{c_n^*(t)\}}^{w^*}$ (the measurability of C^- follows from the fact that $C^-(t) = \bigcap_{n\in\mathbb{N}}\{ x^* \in X^* : \langle c_n(t), x^*\rangle \leq 0\}$). Assume (a) and consider $h \in \mathrm{L}^\infty(T, X)$ and with $\|h\|_\infty \leq 1$. Then $\langle \mathrm{e} + \delta h(t), c_n^*(t)\rangle \leq \langle \mathrm{e}, c_n^*(t)\rangle + \delta\|c_n^*(t)\| \leq 0$ ae. Since the last inequality holds for all n and the measurable selection $c_n^*(t)$ is dense in $C^-(t)$, we conclude that $\mathrm{e} + \delta h(t) \in (C^-(t))^-$ ae. Then using the Bipolar theorem (see e.g. [46, Theorem 3.38]), we have $\mathrm{e} + \delta h(t) \in C(t)$ ae. Now assume (b) and consider a measurable selection c^* of C^- and $\varepsilon_n > 0 \to 0$ and take a measurable selection $h_n(t) \in B(0, 1)$ such that $\langle h_n(t), c^*(t)\rangle \geq \|c^*(t)\| - \varepsilon_n$ ae. So, $\mathrm{e} + \delta h(t) \in C(t)$ ae, and hence $\langle c^*(t), \mathrm{e} + \delta h(t)\rangle \leq 0$ ae. Therefore, $\delta^{-1}\langle c^*(t), -\mathrm{e}\rangle \geq \|c^*(t)\| - \varepsilon_n$ ae. From the stability of null sets under countable intersections we get $\delta^{-1}\langle c^*(t), -\mathrm{e}\rangle \geq \|c^*(t)\|$. □

Remark 6.23 When the measurable function C has cone values, it is easy to see that relation (6.6) implies that for all $t \in T$, $\partial_L^\infty f(t, x) \subseteq C(t)$ and $UI(C) = \{\xi \in X : \xi \in C^-(t) \text{ ae}\}$. In addition, if the values of C are also w^*-closed and convex, then the integrable compact sole property can be understood in terms of the negative polar set $C^-(t)$ (see Lemma 6.22). Finally, the most simple case is when C is a fixed w^*-closed convex cone; in this case Lemma 6.22 characterizes the compact sole property as the non-empty interior of the polar cone $C^-(\subseteq X)$.

Remark 6.24 It is worth commenting that in Theorem 6.21 one can also admit a dependence on x in the measurable multifunction K. Indeed, one can define the *Integrable compact sole at x* property of a multifunction as: Consider a measurable multifunction $K : T \times X \rightrightarrows X^*$ with non-empty closed values, we say that K has an integrable compact sole at x if and only if for every measurable sequence x_n such that $\|x_n - x\|_\infty \to 0$ and $\int_T |f(t, x_n(t)) - f(t, x)| \mathrm{d}\mu(t) \to 0$, there exist $\mathrm{e} \in X$ and $\delta > 0$ such that $t \to K(t, x_n(t))$ is measurable and for any measurable selection c_n^* of $K(\cdot, x_n(\cdot))$ one has

$$\delta\langle c^*(t), \mathrm{e}\rangle \geq \|c^*(t)\| \; ae.$$

Moreover, we denote $UI(K) := \bigcap\{\xi \in X : \sigma_{K(\cdot, x_n(\cdot))}(\xi)^+ \in \mathrm{L}^1\}$, where the intersection is over all sequences $x_n \in \mathrm{L}^\infty(T, X)$ such that $\|x_n - x\|_\infty \to 0$, $\int_T |f(t, x_n(t)) - f(t, x)| \mathrm{d}\mu(t) \to 0$.

Finally, one can construct cones satisfying inclusion (6.6) with dependence on x as follows. Consider $\eta > 0$ and a positive integrable function ℓ. Then, define $C_\varepsilon^\ell(t, z) = \{h \in X : \mathrm{d}^- f_t(z; h) \leq \ell(t)\|h\|\}$, where $\mathrm{d}^- f_t(z; h)$ refers to the *(lower) Dini-Hadamard subderivative*. Then it is not difficult to check that for every $t \in T$ and $z \in X$

$$\partial f(t, z) \subseteq \ell(t)\mathbb{B}(0, 1) + (C_\varepsilon^\ell(t, z))^-.$$

Remark 6.25 Following the same proof as the above Theorem, one can easily remove the set $UI(C)$ in Equations (6.7) and (6.8) imposing the additional assumption that the point e selected in (6.9) satisfies that the sequence $\{\langle x_n^*(t), \mathrm{e}\rangle^+\}$ is uniformly integrable.

Corollary 6.26 *In the setting of Theorem 6.21 assume that for any sequence of measurable functions $x_n^*(t) \in \partial f(t, x_n(t))$ with $\|x_n(\cdot) - x\|_\infty \to 0$, $\int_T f(t, x_n(t)) \mathrm{d}\mu(t) \to I_f(x)$ and $\|x_n^*\|$ is bounded in $\mathrm{L}^1(T, X^*)$, there exists a constant function* e *in the interior of* $\{x(\cdot) \in \mathrm{L}^\infty(T, X) :$

$x(t) \in C(t)$ *ae*} *such that the sequence* $\{ \langle x_n^*(t), \mathrm{e}\rangle^+\}$ *is uniformly integrable. Then*

$$\partial_L I_f(x) \subseteq \bigcap_{\substack{F \subseteq X \\ \dim(F) < +\infty}} \left\{ \int_T \partial_L f(t,x) \mathrm{d}\mu(t) + F^\perp \right\}, \tag{6.11}$$

$$\partial_L^\infty I_f(x) \subseteq \bigcap_{\substack{F \subseteq X \\ \dim(F) < +\infty}} \left\{ \int_T \partial_{\mathrm{L}}^\infty f(t,x) \mathrm{d}\mu(t) + F^\perp \right\}, \tag{6.12}$$

and consequently,

$$\partial^C I_f(x) \subseteq \overline{\mathrm{co}}^{w^*} \left\{ \int_T \partial_L f(t,x) \mathrm{d}\mu + \int_T \partial_{\mathrm{L}}^\infty f(t,x) \mathrm{d}\mu(t) \right\}.$$

Corollary 6.27 *In the setting of Theorem 6.21 assume that the multifunction C is a constant cone. Then*

$$\partial_L I_f(x_0) \subseteq \bigcap \{ \int_T \partial_L f(t,x) \mathrm{d}\mu(t) + C + W^\perp \}; \quad \textit{and } \partial_L^\infty I_f(x_0) \subseteq C,$$

where the intersection is over all finite dimensional subspaces $W \subseteq X$. Consequently,

$$\partial_C I_f(x_0) \subseteq \overline{\mathrm{co}}^{w^*} \left\{ \int_T \partial_L f(t,x) \mathrm{d}\mu + C \right\}.$$

Remark 6.28 The motivation for using boundedness condition (6.6) comes from an application to stochastic programming; more precisely, applications to probability constraints (see [122], [123]), where the authors impose boundedness conditions over the gradients of the involved functions to guarantee the interchange between the sign of the integral and the subdifferential.

The following examples show the importance of C in Theorem 6.21 and Corollary 6.27.

Example 6.29 Consider the integrand $f :]0,1] \times \mathbb{R} \to [0,+\infty)$ given by

$$f(t,x) = \begin{cases} x^{3/2} t^{-1+x} & if \quad x > 0, \\ 0 & if \quad not. \end{cases}$$

It is easy to check that f is continuously differentiable with respect to x and

$$I_f(x) = \begin{cases} \sqrt{x} & if \quad x > 0, \\ 0 & if \quad not. \end{cases}$$

Then we easily get $\partial_M I_f(0) = [0,+\infty)$, $\partial_F f(t,x) = \begin{cases} \frac{3}{2} x^{1/2} t^{-1+x} + x^{3/2} \ln(t) t^{-1+x} & if \quad x > 0, \\ 0 & if \quad not, \end{cases}$ and $\partial_M f(t,x) = \{0\}$. Then we can consider $C = [0,+\infty)$, so that $\partial I_f(0) = \int_{]0,1]} \partial_M f_t(0) \mathrm{d}\mu(t) + C = \{0\} + [0,+\infty)$. The same example can be modified as

$$f(t,x) = \begin{cases} x^2 t^{-1+x} & if \quad x > 0, \\ 0 & if \quad not. \end{cases}$$

Then one has

$$I_f(x) = \begin{cases} x & \text{if} \quad x > 0, \\ 0 & \text{if} \quad \textit{not}. \end{cases}$$

So, the integral functional I_f is Lipschitz continuous, but it is not true that $\partial_M I_f(0) = \{0, 1\}$ is included in $\int_{]0,1]} \partial_M f(t, 0) \mathrm{d}\mu(t) = \{0\}$, as in classical results (see [85, Lemma 6.18] and also [86] for an extension of this result). However, Corollary 6.27 guarantees the inclusion $\partial_M I_f(0) \subseteq \int_{]0,1]} \partial_M f(t, 0) \mathrm{d}\mu(t) + [0, +\infty)$.

Remark 6.30 As a final comment we recall that in the finite dimensional setting two lsc functions f_1, f_2 satisfy the sum rule inclusion $\partial_L (f_1 + f_2)(x) \subseteq \partial^L f_1(x) + \partial^L f(x)$ at a point x provided that the asymptotic qualification condition $\partial_M^\infty f_1(x) \cap \partial_M^\infty f_2(x) = \{0\}$ holds. However, the reader can notice that in the above example the integrand is continuously differentiable, then the singular subdifferential $\partial_M^\infty f_t(0) = \{0\}$ for all $t \in T$. In other words, it is not possible to recover similar criteria, as in the finite sum, in terms of the singular subdifferentials, to get an inclusion of the form $\partial I_f(x) = \int_T \partial_M f_t(x) \mathrm{d}\mu(t)$.

Corollary 6.31 *In the setting of Theorem 6.27 assume that the multifunction $C = \{0\}$. Then I_f is locally Lipschitz around x. In addition, if X is finite dimensional and $\partial_L f(t, x')$ is single valued ae for all x' in a neighborhood of x, then I_f is continuous differentiable differentiable at x.*

Proof. By Theorem 6.27, the Clarke subdifferential $\partial_C I_f$ is bounded by $M := \int K \mathrm{d}\mu$ in a neighborhood of x. Then a straightforward application of Zagrodny's Mean Value Theorem (see [130, Theorem 4.3]) shows that I_f is Locally Lipschitz around x. Furthermore, if X is finite dimensional and $\partial_l f(t, x')$ is single valued ae for all x' in a neighborhood of x, then $\partial_C I_f$ single valued for all x' in a neighborhood of x, and so [23, Proposition 2.2.4 and its Corollary] implies the result. $\square$

Chapter 7

Subdifferential characterization of probability functions under Gaussian distribution

7.1 Introduction

The aim of this paper is to investigate subdifferential properties of Gaussian probability functions induced by nonnecessarily smooth initial data. This topic combines aspects of stochastic programming with arguments from variational analysis, two areas which have been crucially influenced by the fundamental work of Prof. Roger J-B Wets (see, e.g., [106], [129] and many other references). The motivation to study analytical properties of probability functions comes from their importance in the context of engineering problems affected by random parameters. They are at the core of probabilistic programming (i.e., optimization problems subject to probabilistic constraints) (e.g., [95], [110]) or of reliability maximization (e.g., [44]).

A probability function assigns to a control or decision variable the probability that a certain random inequality system induced by this decision variable be satisfied (see (7.1) below). Since such functions are typical constituents of optimization problems under uncertainty, it is natural to ask for their analytical properties, first of all differentiability. Roughly speaking, this can be guaranteed under three assumptions: the differentiability of the input data, an appropriate constraint qualification for the given random inequality system and the compactness of the set of realizations of the random vector for the fixed decision vector (e.g., [74], [93], [120]). While the first two assumptions are quite natural, the last one appears to be restrictive in problems involving random vectors with unbounded support. Failure of the compactness condition, however, may result in general in nonsmoothness of the probability function despite the fact that all input data are smooth and a standard constraint qualification is satisfied (see [55, Prop. 2.2]). In order to keep the differentiability while doing without the compactness assumption, one may restrict to special distributions such as Gaussian or Gaussian-like as in [55], [123]. The working horse for deriving differentiability and gradient formulae in these cases is the so-called *spheric-radial decomposition* of Gaussian random vectors [49, p. 29]. The resulting formulae for the gradient of the probability function are represented - similar to the formulae for the probability values themselves - as integrals over the unit sphere with respect to the uniform measure. The latter can be efficiently approximated by QMC methods tailored to this specific measure (e.g., [18]). Such approach, by exploiting special properties of the distribution, promises more efficiency in the solution of probabilistic programs than general gradient formulae in terms of possibly complicated surface or volume integrals. Successful applications of this methodology in the context of probabilistic programming in gas network optimization is demonstrated in [52, 53].

The aim of this paper is to substantially extend the earlier results in [55], [123] in two directions: first, decisions will be allowed to be infinite-dimensional and second, the random inequality may be just locally Lipschitzian rather than smooth. As the resulting probability function can be expected to be continuous only (rather than locally Lipschitzian or even smooth), appropriate tools (subdifferentials) from variational analysis will be employed for an analytic characterization.

We consider a probability function $\varphi : X \to \mathbb{R}$ defined by

$$\varphi(x) := \mathbb{P}\left(g\left(x, \xi\right) \leq 0\right), \tag{7.1}$$

where X is a Banach space, $g : X \times \mathbb{R}^m \to \mathbb{R}$ is a function depending on the realizations of an m-dimensional random vector ξ. Such probability functions are important in many optimization problems dealing with reliability maximization or probabilistic constraints. The latter one refers to an inequality $\varphi(x) \geq p$ constraining the set of feasible decisions in an optimization problem, in order to guarantee that the underlying random inequality $g\left(x, \xi\right) \leq 0$ is satisfied under decision x with probability at least $p \in (0, 1]$, referred to as a a probability level (or safety level). Since we allow in our paper the function g to be locally Lipschitzian, there is no loss of generality in considering a single random inequality only because in a finite system of such inequalities one could pass to the maximum of components.

Throughout the paper, we shall make the following basic assumptions on the data of (7.1):

(H)

1. X is a reflexive and separable Banach space.
2. Function g is locally Lipschitzian as a function of both arguments simultaneously, and convex as a function of the second argument.
3. The random vector ξ is Gaussian of type $\xi \sim \mathcal{N}\left(0, A\right)$, where A is a correlation matrix.

A brief discussion of these assumptions is in order here: reflexivity of X is imposed in order to work with the limiting (Mordukhovich) subdifferential (actually, one could consider the more general case of Asplund spaces, or simply separable Banach space considering the G-subdifferential). The separability of X is needed in order to make use of an interchange formula for the limiting subdifferential and integration sign (see Chapter 6). For the same reason, g is required to be locally Lipschitzian. As already mentioned above, considering just one inequality rather than a system is no more restriction then. In particular, the single inequality $g\left(x, z\right) \leq 0$ could represent a finite or (compactly indexed) infinite system of smooth inequalities. Considering a Gaussian random vector ξ allows one to pass to a whole class of Gaussian-like multivariate distributions (e.g., Student, Log-normal, truncated Gaussian, χ^2 etc.) upon shifting their nonlinear transformations to a Gaussian one into a modified function $\tilde{g}$ satisfying the same assumptions as required for g here (e.g. [55, Section 4.3]). Moreover, assuming a centered Gaussian distribution with unit variances isn't a restriction either, because in the general case $\xi \sim \mathcal{N}(\mu, \Sigma)$, we may pass to the standardized vector $\tilde{\xi} := D(\xi - \mu)$, where D is the diagonal matrix with elements $D_{\mathrm{ii}} := 1/\sqrt{\Sigma_{\mathrm{ii}}}$. Then, as required above, $\tilde{\xi} \sim \mathcal{N}(0, A)$, with A being the correlation matrix associated with Σ and so

$$\varphi(x) = \mathbb{P}(g(x,\xi) \leq 0) = \mathbb{P}\left(\tilde{g}\left(x, \tilde{\xi}\right) \leq 0\right); \quad \tilde{g}(x,z) := g\left(x, D^{-1}z + \mu\right).$$

Clearly, $\tilde{g}$ is locally Lipschitzian and is convex in the second argument if g is so. Hence, there is no loss of generality in assuming that $\xi \sim \mathcal{N}(0, A)$ from the very beginning.

Our first observation is that our basic assumptions above do not guarantee the continuity of φ even if g is continuously differentiable. A simple two-dimensional example is given by $g(r,s) := r \cdot s$ (which is convex in the second argument) and $\xi \sim \mathcal{N}(0,1)$. Then, $\varphi(r) = 0.5$ for $r \neq 0$ and $\varphi(0) = 1$. Since we want to have the continuity as a minimum initial property of φ in our analysis, we will add the additional assumption that $g(\bar{x}, 0) < 0$ holds true at a point of interest $\bar{x}$ (at which a subdifferential of φ is computed). In other words, given the convexity of g in the second argument, zero is a Slater point for the inequality $g(x,z) \leq 0$, $z \in \mathbb{R}^m$. As shown in [55, Proposition 3.11], the opposite case would entail that $\varphi(\bar{x}) \leq 0.5$. Since one deals in typical applications like probabilistic programming or reliability maximization with probabilities close to one, it follows that the assumption $g(\bar{x}, 0) < 0$ can be made without any practical loss of generality.

The paper is organized as follows: In Section 3 and 4, we provide all the auxiliary results (continuity and partial subdifferential of the radial probability function) which are needed for the derivation of the main subdifferential formula presented in Section 5. This main result which is valid for general continuous probability functions will be specified then by adding additional hypotheses to the locally Lipschitzian and differentiable case. An application to probability functions induced by a finite system of smooth inequalities is given in Subsection 7.5.4.

7.2 Spheric-radial decomposition of Gaussian random vectors

We recall the fact that any Gaussian random vector $\xi \sim \mathcal{N}(0, A)$ has a so-called spheric-radial decomposition, which means that the probability of ξ taking values in an arbitrary Borel subset M of $\mathbb{R}^m$ can be represented as (e.g., [38, p. 105])

$$\mathbb{P}(\xi \in M) = \int\limits_{v \in \mathbb{S}^{m-1}} \mu_\eta(\{r \geq 0 \mid rLv \in M\})\, \mathrm{d}\mu_\zeta(v),$$

where $\mathbb{S}^{m-1} := \left\{v \in \mathbb{R}^m \mid \|v\|^2 = 1\right\}$ denotes the unit sphere in $\mathbb{R}^m$, μ_η is the one-dimensional Chi-distribution with m degrees of freedom, and μ_ζ refers to the uniform distribution on $\mathbb{S}^{m-1}$.

Moreover, the (non-singular) matrix L is supposed to be a factor in a decomposition $A = LL^T$ of the positive definite correlation matrix A (e.g. Cholesky decomposition).

The spheric-radial decomposition allows us to rewrite the probability function (7.1) in the form

$$\varphi(x) = \int_{\mathbb{S}^{m-1}} \mathrm{e}(v,x)\mathrm{d}\mu_\zeta(v) \quad \forall x \in X, \tag{7.2}$$

where $\mathrm{e} : \mathbb{S}^{m-1} \times X \to \mathbb{R}$ refers to the *radial probability function* defined by

$$\mathrm{e}(v,x) := \mu_\eta\left(\{r \geq 0 \mid g(x, rLv) \leq 0\}\right). \tag{7.3}$$

With any $x \in X$ satisfying $g(x,0) < 0$, we will associate the finite and infinite directions defined respectively as

$$\begin{aligned} F(x) \quad &:= \{v \in \mathbb{S}^{m-1} \mid \exists r \geq 0 : g(x, rLv) = 0\}, \\ I(x) \quad &:= \{v \in \mathbb{S}^{m-1} \mid \forall r \geq 0 : g(x, rLv) < 0\}. \end{aligned}$$

It is easily observed that $F(x) \cap I(x) = \emptyset$ and that $F(x) \cup I(x) = \mathbb{S}^{m-1}$ by continuity of g. Moreover, the number $r \geq 0$ satisfying $g(x, rLv) = 0$ in the case of $v \in F(x)$ is uniquely defined, due to the convexity of g in the second argument. This leads us to define the following radius function for any x with $g(x,0) < 0$ and any $v \in \mathbb{S}^{m-1}$:

$$\rho(v,x) := \begin{cases} r \text{ such that } g(x, rLv) = 0 & \text{if } v \in F(x) \\ +\infty & \text{if } v \in I(x). \end{cases} \tag{7.4}$$

This definition allows us to rewrite the radial probability function e from (7.3) in the form

$$\mathrm{e}(v,x) = \mu_\eta\left([0, \rho(v,x)]\right) = F_\eta\left(\rho(v,x)\right) \tag{7.5}$$

whenever $g(x,0) < 0$. Here, F_η refers to the distribution function of the Chi-distribution with m degrees of freedom, so that $F'_\eta(t) = \chi(t)$, where χ is the corresponding density:

$$\chi(t) := Kt^{m-1}\mathrm{e}^{-t^2/2} \quad \forall t \geq 0. \tag{7.6}$$

The second equation in (7.5) follows from $F_\eta(0) = 0$. We formally put $F_\eta(\infty) := 1$ which translates the limiting property $F_\eta(t) \to_{t\to+\infty} 1$ of cumulative distribution functions.

To avoid misunderstandings for the function $g(x,z)$ of two variables, we will refer to its partial subdifferentials at a point $(\bar{x}, \bar{z})$ with a superindex as following

$$\partial_x^{F/M/C} g(\bar{x}, \bar{z}) := \partial_{F/M/C}\, g(\cdot, \bar{z})(\bar{x}); \quad \partial_z^{F/M/C} g(\bar{x}, \bar{z}) := \partial_{F/M/C}\, g(\bar{x}, \cdot)(\bar{z}),$$

for the Fréchet, Limiting/Mordukhovich and Clarke subdifferential respectively. For the functions e and ρ, the notations $\partial_{F/M/C}\mathrm{e}(v,x)$ and $\partial_{F/M/C}\rho(v,x)$ will be consider as in the previous two chapters, that is, the subdifferential is taken with respect to the variable in $x \in X$ and not with respect to the integration variable $v \in \mathbb{S}^{m-1}$.

7.3 Continuity properties

In this section, we investigate continuous properties of the radial probability and the radius functions, defined respectively in (7.3) and (7.4), which are the basis for deriving in Section 7.5 subdifferential formulae for probability function (7.1).

For all the following results, the basic assumption (H) formulated in the Introduction is tacitly required to hold; namely, function g is locally Lipschitzian as a function of both arguments simultaneously, and convex as a function of the second argument.

Lemma 7.1 *Define* $U := \{x \in X \mid g(x,0) < 0\}$.

1. *The radius function* ρ *is continuous at* (v,x) *for any* $x \in U$ *and any* $v \in F(x)$.

2. *For* $x \in U$ *and* $v \in I(x)$ *it holds that* $\lim_{k\to\infty} \rho(v_k, x_k) = \infty$ *for any sequence* $(x_k, v_k) \to (v,x)$ *such that* $v_k \in F(x_k)$.

Proof. Observe first, that ρ is defined (possibly extended-valued) on $U \times \mathbb{S}^{m-1}$. To verify 1., consider any sequence $(x_k, v_k) \to_k (v,x)$ with $v_k \in \mathbb{S}^{m-1}$. We show first that the sequence $\rho(x_k, v_k)$ is bounded. Indeed, otherwise there would exist a subsequence with $\rho(v_{k_l}, x_{k_l}) \to_l \infty$. Clearly $g(x_{k_l}, 0) < 0$ for l large enough, because of $g(x,0) < 0$. Fix an arbitrary $r \geq 0$. Then $\rho(x_{k_l}, v_{k_l}) > r$. We claim that $g(x_{k_l}, rLv_{k_l}) < 0$ for these l's. This is obvious in case that $v_{k_l} \in I(x_{k_l})$. If $v_{k_l} \in F(x_{k_l})$, then the relations

$$g(x_{k_l}, 0) < 0, \quad g(x_{k_l}, \rho(v_{k_l}, x_{k_l}) Lv_{k_l}) = 0, \quad \rho(v_{k_l}, x_{k_l}) > r,$$

and

$$g(x_{k_l}, rLv_{k_l}) \geq 0,$$

would contradict the convexity of g in the second argument. Hence, for l sufficiently large,

$$g(x_{k_l}, rLv_{k_l}) < 0,$$

and passing to the limit yields that $g(x, rLv) \leq 0$, which holds true for all $r \geq 0$ because the latter was chosen arbitrary. But then, $g(x, rLv) < 0$ for all $r \geq 0$, because otherwise once more a contradiction with convexity of g in the second argument would arise from $g(x,0) < 0$. This, however, amounts to $v \in I(x)$ contradicting our assumption $v \in F(x)$. Summarizing, we have shown that $\rho(v_k, x_k)$ is bounded and, in particular, $v_k \in F(x_k)$ for all k. Let $\rho(v_{k_l}, x_{k_l}) \to_l r_0$ be an arbitrary convergent subsequence. Then, we may pass to the limit in the relation $g(x_{k_l}, \rho(v_{k_l}, x_{k_l}) Lv_{k_l}) = 0$ in order to derive that $g(x, r_0 Lv) = 0$, which in turn implies that $r_0 = \rho(v,x)$. Hence, all convergent subsequences of $\rho(v_k, x_k)$ have the same limit $\rho(v,x)$. This implies that $\rho(v_k, x_k) \to_k \rho(v,x)$ and altogether that ρ is continuous at (x,v).

As for 2., observe that if $\rho(v_k, x_k)$ would not tend to infinity, then there would exist a converging subsequence $\rho(x_{k_l}, v_{k_l}) \to_l r_1$ for some $r_1 \geq 0$. Since $\rho(v_{k_l}, x_{k_l}) < \infty$ and $g(x_{k_l}, 0) < 0$ for l large enough, we infer that $v_{k_l} \in F(x_{k_l})$ and, hence, $g(x_{k_l}, \rho(v_{k_l}, x_{k_l}) Lv_{k_l}) = 0$ for all these l's. Now, passing to the limit yields that $g(x, r_1 Lv) = 0$, whence $v \in F(x)$, a contradiction. □

Lemma 7.2 *If $g(x,0) < 0$ and $v \in F(x)$, then there exist neighborhoods U and V of x and v, respectively, such that $v' \in F(x')$ for all $x' \in U$ and $v' \in V \cap \mathbb{S}^{m-1}$.*

Proof. If the statement wasn't true, then there existed a sequence $(v_k, x_k) \to (v, x)$ with $g(x_k, 0) < 0$, $v_k \in \mathbb{S}^{m-1}$ and $v_k \in I(x_k)$. Hence, $\rho(v_k, x_k) = \infty$ and so $\rho(v, x) = \infty$ by 1. in Lemma 7.1. This yields the contradiction $v \in I(x)$. □

Lemma 7.3 *Let $x \in X$ and $r \geq 0$ be such that $g(x,0) < 0$ and $g(x, rLv) = 0$. Then*

$$\langle z^*, Lv \rangle \geq -\frac{g(x,0)}{r} > 0 \quad \forall z^* \in \partial_z g(x, rLv).$$

Proof. By convexity of g in the second variable and by definition of the convex subdifferential, one has that

$$\begin{aligned} -\frac{r}{2}\langle z^*, Lv \rangle &= \left\langle z^*, \frac{r}{2}Lv - rLv \right\rangle \leq g\left(x, \frac{r}{2}Lv\right) - g(x, rLv) \\ &= g\left(x, \frac{r}{2}Lv\right) \leq \frac{1}{2}g(x,0) + \frac{1}{2}g(x, rLv) = \frac{1}{2}g(x,0). \end{aligned}$$

Since our assumptions imply that $r > 0$, the assertion follows. □

We get in the following proposition the desired continuity of the radial probability function e defined in (7.3).

Proposition 7.4 *The radial probability function is continuous at any $(v,x) \in \mathbb{S}^{m-1} \times X$ with $g(x,0) < 0$.*

Proof. Fix a point $(v,x) \in \mathbb{S}^{m-1} \times X$ with $g(x,0) < 0$. Consider any sequence $(v_k, x_k) \to (v, x)$ with $v_k \in \mathbb{S}^{m-1}$ and assume first that $v \in F(x)$. Then, $\rho(v_k, x_k) \to_k \rho(v, x)$ by 1. in Lemma 7.1, and $v_k \in F(x_k)$ for k large, by Lemma 7.2. Hence, by (7.5) it follows that

$$\mathrm{e}(v_k, x_k) = F_\eta(\rho(v_k, x_k)) \to_k F_\eta(\rho(v, x)) = \mathrm{e}(v, x),$$

where the convergence follows from the continuity of the Chi-distribution function F_η.

If in contrast $v \in I(x)$, then, by (7.3), $\mathrm{e}(v,x) = \mu_\eta(\mathbb{R}_+) = 1$. We'll be done if we can show that $\mathrm{e}(v_k, x_k) \to_k 1$. If this did not hold true, then there would exist a subsequence and some $\varepsilon > 0$ such that

$$|\mathrm{e}(v_{k_l}, x_{k_l}) - 1| > \varepsilon \quad \forall\, l. \tag{7.7}$$

Since $v_{k_l} \in I(x_{k_l})$ would imply as above that $\mathrm{e}(v_{k_l}, x_{k_l}) = \mu_\eta(\mathbb{R}_+) = 1$, a contradiction, we conclude that $v_{k_l} \in F(x_{k_l})$ for all l. Now, 2. in Lemma 7.1 guarantees that $\rho(v_{k_l}, x_{k_l}) \to_l \infty$. Then, by (7.5), we arrive at the convergence

$$\mathrm{e}(v_{k_l}, x_{k_l}) = F_\eta(\rho(v_{k_l}, x_{k_l})) \to_l 1,$$

where we exploited the property $\lim_{t\to\infty} F_\eta(t) = 1$, following from F_η being a cumulative distribution function. This is a contradiction with (7.7), and the desired conclusion follows. □

Consequently, we obtain the continuity of the probability function φ, defined in (7.1).

Theorem 7.5 *The probability function is continuous at any point $x \in X$ with $g(x,0) < 0$.*

PROOF. For any sequence $x_n \to x$ one has by Proposition 7.4 that

$$\mathrm{e}\,(v, x_n) \to_n \mathrm{e}\,(v, x) \leq 1 \quad \forall v \in \mathbb{S}^{m-1},$$

where the inequality follows from e being a probability. Since the constant function 1 is integrable on $\mathbb{S}^{m-1}$, the assertion follows from Lebesgue's dominated convergence theorem. □

7.4 Subdifferential of the radial probability function

In this section, we provide characterizations of the Fréchet subdifferential of the radial probability function $\mathrm{e}\,(\cdot, v)$, defined in (7.3), for arbitrarily fixed directions $v \in \mathbb{S}^{m-1}$. As before, we also consider in this section our standard assumption (H).

We need first to estimate the set $\partial_F \rho(v,x)$:

Proposition 7.6 *Let $x \in X$ with $g(x,0) < 0$ and $v \in F(x)$ be arbitrary. Then, for every $y^* \in \partial_F \rho(v,x)$ and every $w \in X$, there exist $x^* \in \partial_C g(x, \rho(v,x)Lv)$ and $z^* \in \partial_z g\,(x, \rho(v,x)Lv)$ such that $\langle z^*, Lv\rangle > 0$ and*

$$\langle y^*, w\rangle \leq \frac{-1}{\langle z^*, Lv\rangle} \langle x^*, w\rangle\,.$$

PROOF. Fix $y^* \in \partial_F \rho(v,x)$ and $w \in X$; hence, $\rho(v,x) < \infty$ (because by assumption $v \in F(x)$). Let $M > 0$ be a Lipschitz constant of g at $(x, \rho(v,x)Lv)$. Then, there exists a neighborhood U of x such that the function $g(\cdot, \rho(v,x)Lv)$ is locally Lipschitzian with Lipschitz constant M at each $x' \in U$, and such that the functions $g(x', \cdot)$, $x' \in U$, are locally Lipschitzian with the same Lipschitz constant M at $\rho(v,x)Lv$. As a consequence of [23, Proposition 2.1.2], for all $x' \in U$ one has that

$$\|x^*\|\,, \|z^*\| \leq M \quad \forall x^* \in \partial_x^C g(x', \rho(v,x)Lv),\ \forall z^* \in \partial_z g(x', \rho(v,x)Lv). \tag{7.8}$$

Consider an arbitrary sequence $t_n \downarrow 0$ so that, by Lemma 7.2, we may assume $v \in F(x + t_n w)$ for all n. By convexity and continuity of the function g with respect to the second variable, the set $\partial g\,(x + t_n w, \cdot)\,(\rho(v,x)Lv)$ is nonempty for all n, and so we may select a sequence

$$z_n^* \in \partial_z g\,(x + t_n w, \cdot)\,(\rho(v,x)Lv); \tag{7.9}$$

hence, taking into account, from the definition of function ρ, that $g(x + t_n w, \rho(x + t_n w, v)Lv) =$

0 and $g(x, \rho(v,x)Lv) = 0$,

$$\begin{aligned}(\rho(x+t_n w, v) - \rho(v,x)) \langle z_n^*, Lv \rangle &= \langle z_n^*, \rho(x+t_n w, v)Lv - \rho(v,x)Lv \rangle \\ &\leq g(x+t_n w, \rho(x+t_n w, v)Lv) \\ &\qquad\qquad - g(x+t_n w, \rho(v,x)Lv) \\ &= -g(x+t_n w, \rho(v,x)Lv) \\ &= g(x, \rho(v,x)Lv) - g(x+t_n w, \rho(v,x)Lv). \end{aligned} \tag{7.10}$$

Next, Lebourg's mean value Theorem for Clarke's subdifferential [23, Theorem 2.3.7] yields some $\tau_n \in [0,1]$ and

$$x_n^* \in \partial_x^C g(x + \tau_n t_n w, \rho(v,x)Lv) \tag{7.11}$$

such that

$$g(x, \rho(v,x)Lv) - g(x+t_n w, \rho(v,x)Lv) \leq -t_n \langle x_n^*, w \rangle, \tag{7.12}$$

and, consequently, from (7.10),

$$(\rho(x+t_n w, v) - \rho(v,x)) \langle z_n^*, Lv \rangle \leq -t_n \langle x_n^*, w \rangle. \tag{7.13}$$

Since X is reflexive and $\|z_n^*\|, \ \|x_n^*\| \leq M$, by (7.8), there exists a subsequence $\left(x_{n_k}^*, z_{n_k}^*\right)$ and some $(x^*, z^*) \in X \times \mathbb{R}^m$ such that $x_{n_k}^* \rightharpoonup x^*$ and $z_{n_k}^* \to z^*$. The weak*-closedness of the graph of Clarke's subdifferential [23, Proposition 2.1.5] along with (7.11) and (7.9) implies that

$$x^* \in \partial_x^C g(x, \rho(v,x)Lv), \ \ z^* \in \partial_z g\left(x, \rho(v,x)Lv\right). \tag{7.14}$$

Now, Lemma 7.3 implies that

$$\langle z^*, Lv \rangle \geq \frac{-g(x,0)}{\rho(v,x)} > 0,$$

and, so, by passing to the (inferior) limit in (7.13), we arrive at

$$\langle z^*, Lv \rangle \liminf_{n\to\infty} t_n^{-1} \left(\rho(x+t_n w, v) - \rho(v,x)\right) \leq -\langle x^*, w \rangle. \tag{7.15}$$

Therefore, since $y^* \in \partial_F \rho(v,x)$,

$$\langle y^*, w \rangle \leq \liminf_{n\to\infty} t_n^{-1} \left(\rho(x+t_n w, v) - \rho(v,x)\right) \leq \frac{-1}{\langle z^*, Lv \rangle} \langle x^*, w \rangle,$$

as we wanted to prove. $\square$

Next, we give the desired estimate of the set $\partial_F \mathrm{e}(v,x)$. Recall that χ is the density of the one-dimensional Chi-distribution with m degrees of freedom (see (7.6)).

Theorem 7.7 *Let $x \in X$ with $g(x,0) < 0$ and $v \in F(x)$ be arbitrary. Then, for every $y^* \in \partial_F \mathrm{e}(v,x)$ and every $w \in X$, there exist $x^* \in \partial_x^C g(x, \rho(v,x)Lv)$ and $z^* \in \partial_z g\left(x, \rho(v,x)Lv\right)$ such that*

$$\langle y^*, w \rangle \leq \frac{-\chi\left(\rho(v,x)\right)}{\langle z^*, Lv \rangle} \langle x^*, w \rangle.$$

Consequently, if $M_{x,v}$ denotes a Lipschitz constant of $g(\cdot, \rho(v,x)Lv)$ at x, then

$$\|y^*\| \leq \frac{\rho(v,x) \cdot \chi\left(\rho(v,x)\right)}{|g(x,0)|} M_{x,v} \quad \forall y^* \in \partial_F \mathrm{e}(v,x).$$

Proof. By (7.5), for all y close to x we may write $\mathrm{e}(v,y) = F_\eta(\rho(v,y))$, with $\rho(v,y) < \infty$, as a consequence of Lemma 7.2. Since F_η is continuously differentiable and nondecreasing (as a distribution function), $F'_\eta(t) \geq 0$ for all $t \in \mathbb{R}$ and, from the calculus of Fréchet subdifferentials (e.g., [76, Corollary 1.14.1 and Proposition 1.11]), we obtain that

$$\begin{aligned}\partial_F \mathrm{e}(v,x) &= \partial_F\left(F'_\eta(\rho(v,x))\,\rho(\cdot,v)\right)(x)\\ &= F'_\eta(\rho(v,x))\partial_F\rho(\cdot,v)(x) = \chi(\rho(v,x))\,\partial_F\rho(v,x).\end{aligned}$$

Combination with Proposition 7.6 yields the first assertion.

To prove the second assertion, from the first part of the proposition we choose elements $x^* \in \partial_x^C g(x,\rho(v,x)Lv)$ and $z^* \in \partial_z g(x,\rho(v,x)Lv)$ such that

$$\langle y^*, w\rangle \leq \left|\frac{-\chi(\rho(v,x))}{\langle z^*, Lv\rangle}\right| \|x^*\|\,\|w\|,$$

and so, since $\langle z^*, Lv\rangle \geq \frac{-g(x,0)}{\rho(v,x)} > 0$ by Lemma 7.3,

$$\langle y^*, w\rangle \leq \frac{\rho(v,x)\cdot\chi(\rho(v,x))}{|g(x,0)|} M_{x,v}\,\|w\|,$$

yielding the desired conclusion. □

We shall also need the following result.

Corollary 7.8 (i) *For every $x_0 \in X$ with $g(x_0,0) < 0$ and every $v_0 \in F(x_0)$ there exist neighborhoods $\tilde{U}$ of x_0 and $\tilde{V}$ of v_0 as well as some $\alpha > 0$ such that*

$$\partial_F \mathrm{e}(v,x) \subseteq \mathbb{B}(0,\alpha) \quad \forall (v,x) \in \left(\tilde{V}\cap\mathbb{S}^{m-1}\right)\times\tilde{U}. \tag{7.16}$$

(ii) *For all $x \in X$ with $g(x,0) < 0$ and for all $v \in I(x)$ one has that $\partial_F \mathrm{e}(v,x) \subseteq \{0\}$.*

Proof. (i) Let $M > 0$ and define open neighborhoods $\tilde{U}$ of x_0 and $\tilde{V}$ of v_0 such that M is a Lipschitz constant of g on $\tilde{U}\times\tilde{V}$ and, for all $(v,x) \in \left(\tilde{V}\cap\mathbb{S}^{m-1}\right)\times\tilde{U}$ (recall Lemma 7.2),

$$g(x,0) < 0,\ \ \rho(v,x) < \infty.$$

Hence, by Theorem 7.7,

$$\partial_F \mathrm{e}(v,x) \subseteq \mathbb{B}(0,\alpha(v,x)),$$

where

$$\alpha(v,x) := \frac{\rho(v,x)\cdot\chi(\rho(v,x))}{|g(x,0)|} M_{x,v}.$$

Taking into account the continuity of ρ (see Lemma 7.1), we may suppose for all $(v,x) \in \left(\tilde{V}\cap\mathbb{S}^{m-1}\right)\times\tilde{U}$ that M is a Lipschitz constant for $g(\cdot,\rho(v,x)Lv)$ at the point x ($\in \tilde{U}$). Thus, we can replace $M_{x,v}$ by M in the definition of α above. Moreover, since g is continuous (also by

Lemma 7.1), as well as the Chi-density χ, we deduce that α is continuous on $\left(\tilde{V} \cap \mathbb{S}^{m-1}\right) \times \tilde{U}$. Then, after shrinking $\tilde{V} \times \tilde{U}$ if necessary, we may assume that for some $\alpha > 0$

$$\alpha(v,x) \leq \alpha \qquad \forall\,(v,x) \in \left(\tilde{V} \cap \mathbb{S}^{m-1}\right) \times \tilde{U}.$$

This proves (7.16).

(ii) As already observed in the proof of Proposition 7.4, $v \in I(x)$ implies that $\mathrm{e}(v,x) = 1$. Consequently, the function $\mathrm{e}(v,\cdot)$ (as the value of a probability) reaches a global maximum at x. Let $x^* \in \partial_F \mathrm{e}(v,x)$ and $u \in X \backslash \{0\}$ be arbitrary. Then,

$$\begin{aligned} -\left\langle x^*, \frac{u}{\|u\|} \right\rangle &= \liminf_{n \to \infty} -\frac{\langle x^*, n^{-1}u \rangle}{\|n^{-1}u\|} \\ &\geq \liminf_{n \to \infty} \frac{\mathrm{e}(x + n^{-1}u, v) - \mathrm{e}(v,x) - \langle x^*, n^{-1}u \rangle}{\|n^{-1}u\|} \\ &\geq \liminf_{h \to 0} \frac{\mathrm{e}(x+h,v) - \mathrm{e}(v,x) - \langle x^*, h \rangle}{\|h\|} \geq 0. \end{aligned}$$

Hence $\langle x^*, u \rangle \leq 0$ for all $u \in X$, and so $x^* = 0$ as desired. □

Definition 7.9 *For $x \in X$ and $l > 0$, we call*

$$C_l(x) := \{h \in X \mid g^\circ(\cdot,z)(y;h) \leq l \, \|z\|^{-m} \, \mathrm{e}^{\frac{\|z\|^2}{2\|L\|^2}} \, \|h\| \;\; \forall y \in \mathbb{B}\,(x, 1/l)\,,\, \|z\| \geq l\}$$

the l-cone of nice directions at $x \in X$. We denote the polar cone to $C_l(x)$ as $C_l^-(x)$.

Note that, by positive homogeneity of Clarke's directional derivative, $\{C_l\}_{l \in \mathbb{N}}$ defines a non-decreasing sequence of closed cones.

We give in the following theorem another estimate for $\partial_F \mathrm{e}(v,x)$, which will be useful in the sequel.

Theorem 7.10 *Fix $x_0 \in X$ such that $g(x_0, 0) < 0$. Then, for every $l > 0$, there exists some neighborhood U of x_0 and some $R > 0$ such that*

$$\partial_F \mathrm{e}(v,x) \subseteq \mathbb{B}\,(0,R) - C_l^-(x_0) \quad \forall x \in U,\, v \in \mathbb{S}^{m-1}.$$

Proof. Let $l > 0$ be arbitrarily fixed. It will be sufficient to show that for every $v_0 \in \mathbb{S}^{m-1}$ there are neighborhoods $\bar{U}$ of x_0 and $\bar{V}$ of v_0 and some $R > 0$ such that

$$\partial_F \mathrm{e}(v,x) \subseteq \mathbb{B}\,(0,R) - C_l^-(x_0) \quad \forall\,(v,x) \in (\bar{V} \cap \mathbb{S}^{m-1}) \times \bar{U}. \tag{7.17}$$

If this holds true, then the global inclusion in the statement of this proposition will follow from the local ones above by a standard compactness argument with respect to $\mathbb{S}^{m-1}$.

In order to prove (7.17), fix an arbitrary $v_0 \in \mathbb{S}^{m-1}$. Assume first that $v_0 \in I(x_0)$. Then, define open neighborhoods U^* of x_0 and V^* of v_0 such that $U^* \subseteq \mathbb{B}\,(x_0, 1/l)$ (with $l > 0$ as fixed above) and, for all $x \in U^*$ and $v \in V^* \cap F(x)$,

$$g(x,0) \leq \frac{1}{2} g(x_0, 0) < 0, \;\; \rho(v,x)\|Lv\| \geq l.$$

Note, that the last inequality is possible by virtue of 2. in Lemma 7.1 and by L being nonsingular and $\mathbb{S}^{m-1}$ being compact (therefore $\|Lv\| \geq \delta$ for all $v \in \mathbb{S}^{m-1}$ and some $\delta > 0$). From Corollary 7.8(ii) we derive that

$$\partial_F \mathrm{e}(v,x) \subseteq \{0\} \quad \forall x \in U^*,\ v \in I(x). \tag{7.18}$$

Now, consider an arbitrary $(v,x) \in V^* \times U^*$ such that $v \in F(x)$. Let also $y^* \in \partial_F \mathrm{e}(v,x)$ and $h \in -C_l(x_0)$ be arbitrarily given. Then, by Theorem 7.7, there exist $x^* \in \partial_x^C g(x, \rho(v,x)Lv)$ and $z^* \in \partial_z g\left(x, \rho(v,x)Lv\right)$ such that

$$\langle y^*, h\rangle \leq \frac{\chi\left(\rho(v,x)\right)}{\langle z^*, Lv\rangle} \langle x^*, -h\rangle \leq \frac{\chi\left(\rho(v,x)\right)}{\langle z^*, Lv\rangle} g^\circ(\cdot, \rho(v,x)Lv)(x;-h), \tag{7.19}$$

where the last inequality relies on (1.1) and on the fact that both the density function χ and $\langle z^*, Lv\rangle$ are positive (see Lemma 7.3). Since $-h \in C_l(x_0)$, our conditions on the neighborhoods U^* and V^* stated above guarantee that

$$\begin{aligned} g^\circ(\cdot, \rho(v,x)Lv)(x;-h) &\leq l \left\|\rho(v,x)Lv\right\|^{-m} \mathrm{e}^{\frac{\|\rho(v,x)Lv\|^2}{2\|L\|^2}} \|h\| \\ &\leq l \left\|\rho(v,x)Lv\right\|^{-m} \mathrm{e}^{\frac{\rho(v,x)^2}{2}} \|h\|, \end{aligned}$$

where we used the triangle inequality. This allows us to continue (7.19) as

$$\begin{aligned} \langle y^*, h\rangle &\leq \frac{\chi\left(\rho(v,x)\right)\rho(v,x) l}{|g(x,0)|} \left\|\rho(v,x)Lv\right\|^{-m} \mathrm{e}^{\frac{\rho(v,x)^2}{2}} \|h\| \\ &= \frac{lK}{|g(x,0)|} \|Lv\|^{-m} \|h\|, \end{aligned}$$

where we used Lemma 7.3 and the definition of the Chi-density with m degrees of freedom (see (7.6)). Owing to $g(x,0) \leq \frac{1}{2} g(x_0,0) < 0$, we may continue as

$$\langle y^*, h\rangle \leq \frac{2lKK^*}{|g(x_0,0)|} \|h\|, \tag{7.20}$$

where (recall that L is nonsingular)

$$K^* := \max_{v \in \mathbb{S}^{m-1}} \|Lv\|^{-m} \in \mathbb{R}_+.$$

Consequently, we have shown that for some $\tilde{K} > 0$, which is independent of x and v,

$$\langle y^*, h\rangle \leq \tilde{K} \|h\| \quad \forall y^* \in \partial_F \mathrm{e}(v,x),\ h \in -C_l(x_0).$$

Using indicator and support functions, respectively, this relation is rewritten as, for all $h \in X$,

$$\begin{aligned} \langle y^*, h\rangle &\leq \tilde{K} \|h\| + \delta_{-\overline{\mathrm{co}}\, C_l(x_0)}(h) \\ &= \sigma_{\mathbb{B}_{\tilde{K}}(0)}(h) + \sigma_{-C_l^-(x_0)}(h) \\ &= \sigma_{\left(\mathbb{B}_{\tilde{K}}(0) - C_l^-(x_0)\right)}(h). \end{aligned}$$

Consequently, we get

$$\sigma_{\partial_F \mathrm{e}(v,x)}(h) \leq \sigma_{\left(\mathbb{B}^*_{\tilde{K}}(0) - C_l^-(x_0)\right)}(h) \quad \forall h \in X,$$

which entails the inclusion

$$\partial_F \mathrm{e}(v,x) \subseteq \mathbb{B}_{\tilde{K}}(0) - C_l^-(x_0).$$

Since $(v,x) \in V^* \times U^*$ with $v \in F(x)$ were chosen arbitrarily, we may combine this with (7.18) to derive that

$$\partial_F \mathrm{e}(v,x) \subseteq \mathbb{B}_{\tilde{K}}(0) - C_l^-(x_0) \quad \forall\, (v,x) \in \left(V^* \cap \mathbb{S}^{m-1}\right) \times U^*.$$

Now, we suppose that $v_0 \in F(x_0)$. Then Corollary 7.8(i) guarantees the existence of neighborhoods $\tilde{U}$ of x_0 and $\tilde{V}$ of v_0 as well as some $\alpha > 0$ such that relation (7.16) holds true. □

Corollary 7.11 *Fix $x_0 \in X$ such that $g(x_0,0) < 0$, and assume one of the following alternative conditions:*

$$\{z \in \mathbb{R}^m \mid g(x_0,z) \le 0\} \text{ is a bounded set,} \tag{7.21}$$

or

$$\exists\, l > 0 \text{ such that } C_l(x_0) = X. \tag{7.22}$$

Then the partial radial probability functions $\mathrm{e}(v,\cdot)$, $v \in \mathbb{S}^{m-1}$, are uniformly locally Lipschitzian around x_0 with some common Lipschitz constant independent of v.

Proof. In the case of (7.21), one has that $I(x_0) = \emptyset$, whence $F(x_0) = \mathbb{S}^{m-1}$. Then, by Corollary 7.8(i), for every $v_0 \in \mathbb{S}^{m-1}$ there exist neighborhoods $\tilde{U}_{v_0}$ of x_0 and $\tilde{V}_{v_0}$ of v_0 as well as some $\alpha_{v_0} > 0$ such that

$$\partial_F \mathrm{e}(v,x) \subseteq \mathbb{B}(0,\alpha_{v_0}) \quad \forall\, (v,x) \in \left(\tilde{V}_{v_0} \cap \mathbb{S}^{m-1}\right) \times \tilde{U}_{v_0}.$$

Then, by the evident compactness argument with respect to the sphere $\mathbb{S}^{m-1}$ already alluded to in the beginning of the proof of Theorem 7.10, we derive the existence of a neighborhood $\tilde{U}$ of x_0 and of some $\alpha > 0$ such that

$$\partial_F \mathrm{e}(v,x) \subseteq \mathbb{B}(0,\alpha) \quad \forall\, (v,x) \in \mathbb{S}^{m-1} \times \tilde{U}.$$

In the case of (7.22), the same relation (with $\alpha := R$) is a direct consequence of Theorem 7.10 upon taking into account that $C_l(x_0) = X$ entails that $-C_l^-(x_0) = \{0\}$. Now, the claimed statement on uniform Lipschitz continuity follows from [84, Theorem 3.5.2]. □

7.5 Subdifferential of the Gaussian probability function φ

In this section, we provide the required formulae for the Fréchet, the Mordukhovich, and the Clarke subdifferentials of the Gaussian probability function φ, defined in (7.1). These results are next illustrated in Example 7.15, and in Subsection 7.5.3 to discuss the Lipschitz continuity and differentiability of φ. Finally, we study in this section, Subsection 7.5.4, the special and interesting setting of probability functions given by means of finite systems of smooth inequalities. In this case, formulae of the subdifferentials of φ are expressed in terms of the initial data in (7.1), i.e., in terms of the function g. All this is done under our standard assumption (H).

7.5.1 Main Result

Now, we are in a position to prove the main result of our paper.

Theorem 7.12 *Let $x_0 \in X$ be such that $g(x_0, 0) < 0$. Assume that the cone $C_l(x_0)$ has a non-empty interior for some $l > 0$. Then,*

(i) $\partial_M \varphi(x_0) \subseteq \mathrm{cl}^* \left\{ \int\limits_{\mathbb{S}^{m-1}} \partial_M \, \mathrm{e}(v, x_0) \mathrm{d}\mu_\zeta(v) - C_l^-(x_0) \right\}$.

(ii) *Provided that X is finite-dimensional,*

$$\partial_M \varphi(x_0) \subseteq \int\limits_{\mathbb{S}^{m-1}} \partial_M \, \mathrm{e}(v, x_0) \mathrm{d}\mu_\zeta(v) - C_l^-(x_0).$$

(iii) $\partial^\infty \varphi(x_0) \subseteq -C_l^-(x_0)$.

(vi) $\partial_C \varphi(x_0) \subseteq \overline{\mathrm{co}} \left\{ \int\limits_{\mathbb{S}^{m-1}} \partial_M \, \mathrm{e}(v, x_0) \mathrm{d}\mu_\zeta(v) - C_l^-(x_0) \right\}$.

Proof. We apply Proposition 6.27 by putting

$$f(\omega, x) := \mathrm{e}(\omega, x), \; C := -C_l^-(x_0),$$

and using the measurable space $(\mathbb{S}^{m-1}, \mathcal{A}, \mu_\zeta)$, with $\mathcal{A}$ being the σ-Algebra of measurable sets with respect to μ_ζ. It is known that μ_ζ is σ-finite and complete. The measurability property of f and the lower semicontinuity of $f(\omega, \cdot)$ are consequences of the continuity of e (see Proposition 7.4). The cone $C^- = \overline{\mathrm{co}}\, C_l(x_0)$ has a non-empty interior, by the current assumption. Condition (6.6) is a consequence of Theorem 7.10 upon defining $\mathcal{K}(\omega) := R$ for all $\omega \in \Omega = \mathbb{S}^{m-1}$, and observing that $\mathcal{K} \in \mathrm{L}^1(\mathbb{S}^{m-1}, \mathbb{R})$, due to $\mathbb{S}^{m-1}$ having finite (μ_ζ-) measure. Now, the claimed result follows from Proposition 6.27 by taking into account that $I_f = \varphi$ thanks to (7.2). □

Our main result motivates some investigation about the impact of the parameter $l > 0$ in the definition of the cones $C_l^-(x_0)$, $x_0 \in X$. From Definition 7.9, it follows immediately that $(C_l(x_0))_{l \geq 0}$ forms a non-decreasing family of closed cones, and hence

$$C_k(x_0) \subseteq C_{k+1}(x_0); \quad C_k^*(x_0) \supseteqq C_{k+1}^*(x_0) \quad \forall k \in \mathbb{N}. \tag{7.23}$$

Moreover, $C_k(x_0)$ having a non-empty interior as required in Theorem 7.12, implies that $C_{k+1}(x_0)$ does so too. This means that the upper estimates in the results of Theorem 7.12 become increasingly precise for $k \to \infty$. This immediately raises the question if we may pass to the limit in this result. Let us then introduce the *limiting cone of nice directions*

$$C_\infty(x_0) := \bigcup_{k \in \mathbb{N}} C_k(x_0) =$$
$$\{h \in X \mid \exists k \in \mathbb{N} : g^\circ(\cdot, z)(y; h) \leq k \|z\|^{-m} \exp(\tfrac{\|z\|^2}{2\|L\|^2}) \|h\|, \forall y \in \mathbb{B}\left(x, \tfrac{1}{k}\right), \|z\| \geq k\}.$$

The reader can simply notice (through Baire's Theorem) the non-emptiness of the interior of $C_\infty(x_0)$ is equivalent to the non-emptiness of the interior of $C_l(x_0)$ for some $l>0$. As far as the singular subdifferential is concerned, we may immediately pass to the limit:

Proposition 7.13 *Fix $x_0 \in X$ with $g(x_0,0)<0$, and assume that $C_l(x_0)$ has a non-empty interior for some $l>0$. Then $\partial^\infty\varphi(x_0) \subseteq -C_\infty^*(x_0)$.*

Proof. By Theorem 7.12(iii) we have that $\partial^\infty\varphi(x_0) \subseteq -C_l^-(x_0)$. Since along with $C_l(x_0)$ the larger cones $C_k(x_0)$ for $k \in \mathbb{N}, k \geq l$, have non-empty interiors too, it follows that

$$\partial^\infty\varphi(x_0) \subseteq \bigcap_{k\in\mathbb{N},k\geq l} -C_k^*(x_0) = -\left(\bigcup_{k\in\mathbb{N}} C_k(x_0)\right)^* = -C_\infty^*(x_0),$$

where the first equality relies on (7.23). □

In order to formulate a corresponding result for the Mordukhovich and Clarke subdifferentials, we need an additional boundedness assumption:

Proposition 7.14 *Fix $x_0 \in X$ with $g(x_0,0)<0$, and assume that $C_l(x_0)$ has a non-empty interior for some $l>0$. Moreover, suppose that $\partial_M \mathrm{e}(v,x_0)$ is integrably bounded; i.e., there exists some function $\varrho : \mathbb{S}^{m-1} \to \mathbb{R}_+$ with $\int\limits_{\mathbb{S}^{m-1}} \varrho(v)\mathrm{d}\mu_\zeta(v) < \infty$ such that*

$$\partial^M \mathrm{e}(v,x_0) \subseteq \mathbb{B}(0,\varrho(v)) \quad \mu_\zeta - a.\mathrm{e.}\ v \in \mathbb{S}^{m-1}.$$

Then

$$\partial_M\varphi(x_0) \subseteq \partial_C\varphi(x_0) \subseteq \mathrm{cl}\left\{\int\limits_{\mathbb{S}^{m-1}} \partial_M \mathrm{e}(v,x_0)\mathrm{d}\mu_\zeta(v)\right\} - C_\infty^*(x_0).$$

Proof. For the purpose of abbreviation, put

$$\mathcal{I} := \int\limits_{\mathbb{S}^{m-1}} \partial_M \mathrm{e}(v,x_0)\mathrm{d}\mu_\zeta(v).$$

From our assumption on $\mathbb{B}\mathrm{e}(v,x_0)$, being integrably bounded, it follows that $\mathcal{I}$ is bounded too. Consequently, $\mathrm{cl}^*\mathcal{I}$ is w^*-compact. With $C_l(x_0)$ having a non-empty interior, for all $k \in \mathbb{N}$ with $k \geq l$, from Theorem 7.12(i) it follows that

$$\partial_M\varphi(x_0) \quad \subseteq \quad \mathrm{cl}^*\left\{\mathcal{I} - C_k^*(x_0)\right\} = \mathrm{cl}^*\mathcal{I} - C_k^*(x_0) \quad \forall k \geq l.$$

Due to (7.23), we may continue as

$$\partial_M\varphi(x_0) \subseteq \bigcap_{k\in\mathbb{N}} \left\{\mathrm{cl}^*\mathcal{I} - C_k^*(x_0)\right\}, \tag{7.24}$$

which in turn, using again the w^*-compacity of $\mathrm{cl}^*\mathcal{I}$, gives us

$$\partial_M \varphi(x_0) \subseteq \mathrm{cl}^*\mathcal{I} - \bigcap_{k\in\mathbb{N}} C_k^*(x_0) = \mathrm{cl}^*\mathcal{I} - \left(\bigcup_{k\in\mathbb{N}} C_k(x_0)\right)^* = \mathrm{cl}^*\mathcal{I} - C_\infty^*(x_0).$$

Now, by [84, Theorem 3.57], by Proposition 7.13, and by convexity of $C_\infty^*(x_0)$, we arrive at

$$\begin{aligned}\partial_C \varphi(x_0) &= \overline{\mathrm{co}}\,\{\partial_M \varphi(x_0) + \partial_M^\infty \varphi(x_0)\}\\ &\subseteq \overline{\mathrm{co}}\,\{\mathrm{cl}^*\mathcal{I} - C_\infty^*(x_0) - C_\infty^*(x_0)\}\\ &= \overline{\mathrm{co}}\,\{\mathrm{cl}^*\mathcal{I} - C_\infty^*(x_0)\}\,.\end{aligned}$$

Now, as a consequence of [96, Theorem 3.1], the strong closure $\mathrm{cl}\,\mathcal{I}$ is convex (the measure μ_ζ being nonatomic), so that $\mathrm{cl}^*\mathcal{I} = \mathrm{cl}\,\mathcal{I}$ is convex, and the last inclusion above reads

$$\partial_C \varphi(x_0) \subseteq \mathrm{cl}\mathcal{I} - C_\infty^*(x_0).$$

This finishes the proof of our proposition. □

7.5.2 Two illustrating examples

In the following, we provide two example which, on the one hand, serves as an illustration of our main result Theorem 7.12 and, on the other hand, shows that even for a continuously differentiable inequality $g(x,\xi) \le 0$, satisfying a basic constraint qualification, the associated probability function φ may fail to be differentiable, actually even to be locally Lipschitzian (though it is continuous due to the constraint qualification; see Theorem 7.5).

Example 7.15 Define the function $g : \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}$ by

$$g(x, z_1, z_2) := \alpha(x)\mathrm{e}^{h(z_1)} + z_2 - 1,$$

where

$$\alpha(x) := \begin{cases} x^2 & x \ge 0 \\ 0 & x < 0, \end{cases}$$

$$h(t) := -1 - 4\log(1 - \Phi(t))\,; \quad \Phi(t) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} \mathrm{e}^{-\tau^2/2} \mathrm{d}\tau,$$

i.e., Φ is the distribution function of the one-dimensional standard normal distribution. Moreover, let ξ have a bivariate standard normal distribution, i.e.,

$$\xi = (\xi_1, \xi_2) \sim \mathcal{N}\left((0,0)\,, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right).$$

The following properties are shown in the Appendix:

1. g is continuously differentiable.

2. g is convex in (z_1, z_2).

3. $g\,(0,0,0) < 0$.

4. $C_1(0) = (-\infty, 0]$.

5. $\int\limits_{\mathbb{S}^1} \partial_M \, \mathrm{e}(v,0) \mathrm{d}\mu_\zeta(v) \subseteq (-\infty, 0]$.

6. φ fails to be locally Lipschitzian in 0.

Observe that, by 1. and 2., g satisfies our basic data assumptions, (H), and that 3. forces the probability function φ to be continuous. On the other hand, by 6., φ is not locally Lipschitzian -much less differentiable - in 0 despite the continuous differentiability of g and the satisfaction of Slater's condition. Now, Theorem 7.12(ii), along with 4. and 5. provides that

$$\partial_M \varphi(0) \subseteq (-\infty, 0] - [0, \infty) = (-\infty, 0]\,, \quad \partial^\infty \varphi(0) \subseteq (-\infty, 0]\,.$$

On the other hand, analytical verification along with the formula for φ provided in the Appendix (or alternatively visual inspection of the graph of φ) yields that $\partial_M \varphi(0) = \{0\}$ and $\partial^\infty \varphi(0) = (-\infty, 0]$, so that the upper estimate for the singular subdifferential is strict, while the one for the basic subdifferential is not (nevertheless this upper estimate is nontrivial due to being smaller than the whole space).

The second example shows a probability function which not satisfies the exponential growth condition at x_0 imposed in [122, 123], which is,

$$\exists l > 0 : \|\nabla_x g(x,z)\| \leq l \mathrm{e}^{\|z\|} \quad \forall x \in \mathbb{B}\,(x_0, 1/l)\,, \; \|z\| \geq l. \tag{7.25}$$

However using our results one can prove that the probability function is continuously differentiable.

Example 7.16 Define the function $g : \mathbb{R} \times \mathbb{R}^2 \to \mathbb{R}$ by

$$g\,(x, z_1, z_2) := \alpha(x) \frac{\exp(z_1^2/2)}{z_1^2 + 4} + z_2 - 1,$$

where

$$\alpha(x) := \begin{cases} x^2 & x \geq 0 \\ 0 & x < 0, \end{cases}$$

Moreover, let ξ have a bivariate standard normal distribution, i.e.,

$$\xi = (\xi_1, \xi_2) \sim \mathcal{N}\left((0,0)\,, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right).$$

The following properties are shown in the Appendix:

1. g is continuously differentiable.

2. g is convex in (z_1, z_2).

3. $g\,(0,0,0) < 0$.

4. $C_1(0) = \mathbb{R}$.

5. g does not satisfy the exponential growth condition at $x_0 = 0$.

6. φ is continuously differentiable at 0.

7.5.3 Lipschitz continuity and differentiability of φ

The following result on Lipschitz continuity of the probability function φ is an immediate consequence of Clarke's Theorem on interchanging subdifferentiation and integration [23, Theorem 2.7.2] and of Corollary 7.11, but also it can be obtained from Corollary 6.27 considering $C = \{0\}$:

Theorem 7.17 *Fix $x \in X$ such that $g(x, 0) < 0$. Under one of the alternative conditions* (7.21) *or* (7.22)*, the probability function φ is locally Lipschitz near x and the following estimate holds true*:

$$\partial_C \varphi(x) \subseteq \int_{\mathbb{S}^{m-1}} \partial_M \mathrm{e}(v, x) \mathrm{d}\mu_\zeta(v). \tag{7.26}$$

The next result provides conditions for differentiability of the probability function φ; recall that $\#A$ denotes the cardinal of a set A.

Proposition 7.18 *In addition to the assumptions of Theorem* 7.17*, assume that*

$$\#\partial_M \mathrm{e}(v, x) = 1 \quad \mu_\zeta\text{-}a.e.\, v \in \mathbb{S}^{m-1}. \tag{7.27}$$

Then φ is strictly differentiable at x and

$$\nabla\varphi(x) = \int_{v \in \mathbb{S}^{m-1}} \nabla \mathrm{e}(v, x) \mathrm{d}\mu_\zeta(v).$$

Consequently, if X is finite-dimensional and (7.27) *holds true in some neighborhood of x, then φ is even continuously differentiable at x.*

Proof. From the assumptions one can apply directly Corollary 6.31 and thus one gets the conclusions. □

7.5.4 Application to a finite system of smooth inequalities

In order to benefit from Theorem 7.12, one has to be able to express the integrand $\mathbb{B}\mathrm{e}(v, x_0)$ in terms of the initial data in (7.1), i.e., in terms of the function g. We will illustrate this for the case of a probability function defined over a finite system of continuously differentiable inequalities which are convex in their second argument:

$$\varphi(x) := \mathbb{P}\left(g_{\mathrm{i}}\left(x, \xi\right) \leq 0,\ \mathrm{i} = 1, \ldots, p\right),\ x \in X. \tag{7.28}$$

Clearly, this can be recast in the form of (7.1) upon defining

$$g := \max_{\mathrm{i}=1,\ldots,p} g_{\mathrm{i}}, \tag{7.29}$$

where g is locally Lipschitz as required and convex in the second argument because the g_{i}'s are supposed to be so. Since $g\left(x, 0\right) < 0$ implies that $g_{\mathrm{i}}\left(x, 0\right) < 0$ for all $\mathrm{i} = 1, \ldots, p$, we may associate with each component a function ρ_{i} satisfying the relation $g_{\mathrm{i}}\left(x, \rho_{\mathrm{i}}\left(v, x\right) Lv\right) = 0$, as we did in (7.4). The relation between ρ associated via (7.4) with g in (7.29) is, clearly,

$$\rho\left(v, x\right) = \min_{\mathrm{i}=1,\ldots,p} \rho_{\mathrm{i}}\left(v, x\right) \quad \forall x : g\left(x, 0\right) < 0,\ \forall v \in F(x). \tag{7.30}$$

Note, however, that unlike ρ, the functions ρ_{i} are continuously differentiable because the g_{i}'s are so. This is a consequence of the Implicit Function Theorem (see [55, Lemma 3.2]), which moreover yields the gradient formulae, for all x with $g\left(x, 0\right) < 0$ and all $v \in F(x)$,

$$\nabla_x \rho_{\mathrm{i}}\left(v, x\right) = -\frac{1}{\left\langle \nabla_z g_{\mathrm{i}}\left(x, \rho\left(v, x\right) Lv\right), Lv\right\rangle} \nabla_x g_{\mathrm{i}}\left(x, \rho\left(v, x\right) Lv\right),\ \mathrm{i} = 1, \ldots, p.$$

In the following proposition, we provide an explicit upper estimate of the subdifferential set $\mathbb{B}\mathrm{e}(v, x_0)$ in terms of the initial data, which can be used in the formula of Theorem 7.12 to get an upper estimate for the subdifferential of the probability function (7.28):

Proposition 7.19 *Fix $x \in X$ such that $g_{\mathrm{i}}\left(x, 0\right) < 0$ for $\mathrm{i} = 1, \ldots, p$. Then, for every $l > 0$, there exists some $R > 0$ such that the radial probability function associated with g in* (7.29) *via* (7.3) *satisfies*

$$\partial_M \mathrm{e}(v, x) \subseteq \begin{cases} -\displaystyle\bigcup_{\mathrm{i} \in T(v)} \left\{ \tfrac{\chi(\rho(x,v))}{\langle \nabla_z g_{\mathrm{i}}(x, \rho(x,v) Lv), Lv\rangle} \nabla_x g_{\mathrm{i}}\left(x, \rho\left(v, x\right) Lv\right)\right\} & v \in F\left(x\right) \\ \mathbb{B}\left(0, R\right) - C_l^-\left(x\right) & v \in I\left(x\right). \end{cases}$$

Here, $T(v) := \left\{\mathrm{i} \in \{1, \ldots, p\} \mid \rho_{\mathrm{i}}\left(x, v\right) = \rho\left(v, x\right)\right\}$.

Proof. Fix an arbitrary $v \in \mathbb{S}^{m-1}$. Given the continuity of e, we exploit the following representation [84, Theorem 2.34] of the Mordukhovich subdifferential in terms of the Fréchet subdifferential, which holds true in Asplund spaces (hence, in particular for reflexive Banach spaces)

$$x^* \in \mathbb{B}\mathrm{e}(v, x) \Longleftrightarrow \exists\, x_n \to_n x \text{ and } \exists\, x_n^* \rightharpoonup_n x^* :\ x_n^* \in \partial_F \mathrm{e}(v, x_n).$$

Then, the inclusion $\mathbb{B}\mathrm{e}(v, x) \subseteq \mathbb{B}\left(0, R\right) - C_l^-(x)$ follows from Theorem 7.10, since $\mathbb{B}\left(0, R\right)$ is weak*-compact and $C_l^-(x)$ is weak*-closed, entailing that $\mathbb{B}\left(0, R\right) - C_l^-(x)$ is weak*-closed. This yields the desired estimate of $\mathbb{B}\mathrm{e}(v, x)$ when $v \in I\left(x\right)$.

Suppose now in addition that $v \in F(x)$, and, according to Lemma 7.2, let U be a neighborhood of x such that, for all $y \in U$,

$$g(y, 0) < 0, \; v \in F(y).$$

From the proof of Theorem 7.7 we have seen that

$$\partial_F \mathrm{e}(v, y) = \chi(\rho(v, y))\, \partial_x^F \rho(v, y), \quad \forall y \in U,$$

which, by continuity of χ and by 1. in Lemma 7.1, immediately entails that

$$\partial_M \mathrm{e}(v, x) = \chi(\rho(v, x))\, \partial_x^M \rho(v, x).$$

From (7.30) and the calculus rule for minimum functions [84, Proposition 1.113] we conclude that

$$\partial_M \rho(v, x) \subseteq \bigcup_{\mathrm{i} \in T(v)} \nabla_x \rho_{\mathrm{i}}(v, x).$$

with $T(v)$ being defined as in the statement of the Proposition. Now, the assertion follows from (7.31). $\square$

We provide next a concrete characterization for the local Lipschitz continuity/differentiability of the probability function φ, defined in (7.28), along with an explicit subdifferential/gradient formula:

Theorem 7.20 *Fix $x_0 \in X$ with $g(x_0, 0) < 0$, and assume that for some $l > 0$ it holds, for $\mathrm{i} = 1, \ldots, p$,*

$$\|\nabla_x g_{\mathrm{i}}(x, z)\| \leq l \, \|z\|^{-m} \, \mathrm{e}^{\frac{\|z\|^2}{2\|L\|^2}} \quad \forall x \in \mathbb{B}(x_0, 1/l), \; \|z\| \geq l. \tag{7.31}$$

Then the probability function (7.28) is locally Lipschitz near x_0 and there exists a nonnegative number $R \leq \sup\{\|x^*\| \mid x^* \in \partial_x^M \mathrm{e}(x_0, v)$ and $v \in I(x_0)\}$ such that

$$\begin{aligned} \partial^C \varphi(x_0) \;\; & \subseteq \;\; - \int_{v \in F(x_0)} \mathrm{co} \left\{ \bigcup_{\mathrm{i} \in T(v)} \frac{\chi(\rho(x_0, v))\, \nabla_x g_{\mathrm{i}}(x_0, \rho(x_0, v) Lv)}{\langle \nabla_z g_{\mathrm{i}}(x_0, \rho(x_0, v) Lv), Lv \rangle} \right\} \mathrm{d}\mu_\zeta(v) \\ & \quad + \mu_\zeta(I(x_0)) \mathbb{B}^*(0, R). \end{aligned}$$

Proof. As a maximum of finitely many smooth functions, g is Clarke-regular, so that Clarke's directional derivative of g coincides with its usual directional derivative. Hence, by Danskin's Theorem and by (7.31), we get the following estimate, for all $h \in X$, $x \in \mathbb{B}_{1/l}(x_0)$ and $\|z\| \geq l$,

$$\begin{aligned} g^\circ(\cdot, z)(x; h) &= \langle \nabla_x g(x, z), h \rangle \\ &= \max \{ \langle \nabla_x g_{\mathrm{i}}(x, z), h \rangle \mid g_{\mathrm{i}}(x, z) = g(x, z) \} \\ &\leq \max_{\mathrm{i}=1,\ldots,p} \langle \nabla_x g_{\mathrm{i}}(x, z), h \rangle \leq l \, \|z\|^{-m} \, \mathrm{e}^{\frac{\|z\|^2}{2\|L\|^2}} \, \|h\| . \end{aligned}$$

Hence, $C_l(x_0) = X$ and, so, Theorem 7.17 guarantees that φ in (7.28) is locally Lipschitz near x_0 and that

$$\partial_C \varphi(x_0) \subseteq \int\limits_{F(x_0)} \partial_C \mathrm{e}(v, x_0) \mathrm{d}\mu_\zeta(v) + \int\limits_{I(x_0)} \partial_C \mathrm{e}(v, x_0) \mathrm{d}\mu_\zeta(v). \tag{7.32}$$

Since $\mathrm{e}(\cdot, v)$ is locally Lipschitzian for all $v \in \mathbb{S}^{m-1}$, it follows from [84, Theorem 3.57] and from Proposition 7.19 that

$$\begin{aligned} \partial_C \mathrm{e}(v, x_0) &= \overline{\mathrm{co}} \left\{ \partial_M \mathrm{e}(v, x_0) \right\} \\ &= -\mathrm{co} \left\{ \bigcup_{\mathrm{i} \in T(v)} \frac{\chi\left(\rho\left(x_0, v\right)\right) \nabla_x g_{\mathrm{i}}\left(x_0, \rho\left(x_0, v\right) L v\right)}{\left\langle \nabla_z g_{\mathrm{i}}\left(x_0, \rho\left(x_0, v\right) L v\right), L v\right\rangle} \right\}. \end{aligned}$$

Hence, the first term on the right-hand side of (7.32) coincides with the integral term in the asserted formula above. As for the second term, observe that $\partial_x^C \mathrm{e}(v, x_0) \subseteq \mathbb{B}(0, R)$ for some $R > 0$ by Theorem 7.10, which yields the second term in the upper estimate of this theorem. □

From Theorem 7.20 and Proposition 7.18, we immediately derive the following:

Corollary 7.21 *If in the setting of Theorem* 7.20 *one has that* $\mu_\zeta(I(x_0)) = 0$ (*in particular, under assumption* (7.21)), *or the constant* R *in Theorem 7.20 is zero, then*

$$\partial_C \varphi(x_0) \subseteq -\int\limits_{\mathbb{S}^{m-1}} \mathrm{co} \left\{ \bigcup_{\mathrm{i} \in T(v)} \frac{\chi\left(\rho\left(x_0, v\right)\right) \nabla_x g_{\mathrm{i}}\left(x_0, \rho\left(x_0, v\right) L v\right)}{\left\langle \nabla_z g_{\mathrm{i}}\left(x_0, \rho\left(x_0, v\right) L v\right), L v\right\rangle} \right\} \mathrm{d}\mu_\zeta(v).$$

If, in addition, for μ_ζ*-a.e.* $v \in \mathbb{S}^{m-1}$ *we have that* $\#T(v) = 1$ *(say:* $T(v) = \{\mathrm{i}^*(v)\}$*), then the probability function* (7.28) *is strictly differentiable with gradient*

$$\nabla \varphi(x_0) = -\int\limits_{v \in \mathbb{S}^{m-1}} \frac{\chi\left(\rho\left(x_0, v\right)\right) \nabla_x g_{\mathrm{i}^*(v)}\left(x_0, \rho\left(x_0, v\right) L v\right)}{\left\langle \nabla_z g_{\mathrm{i}^*(v)}\left(x_0, \rho\left(x_0, v\right) L v\right), L v\right\rangle} \mathrm{d}\mu_\zeta(v).$$

Consequently, if X *is finite-dimensional and* $\#T(v) = 1$ *holds true in some neighborhood of* x*, then* φ *is even continuously differentiable at* x*.*

Remark 7.22 It is worth mentioning that under the strengthened (compared with (7.31)) growth condition

$$\exists l > 0 : \|\nabla_x g_{\mathrm{i}}(x, z)\| \leq l \mathrm{e}^{\|z\|} \quad \forall x \in \mathbb{B}\left(x_0, 1/l\right), \ \|z\| \geq l, \ \mathrm{i} = 1, \ldots, p$$

the constant R in Theorem 7.20 and Corollary above is zero, as it can be seen in (7.20) (see also [123, Theorem 3.6 and Theorem 4.1]).

7.6 Appendix

7.6.1 Example 7.15

We verify in this Appendix properties 1.-6. in Example 7.15.

The continuous differentiability of g stated in 1. is obvious from the corresponding property of α and h. For h, this relies on the smoothness of the distribution function of the one-dimensional standard normal distribution Φ and on the fact that the argument $1-\Phi(t)$ of the logarithm is always strictly positive.

By nonnegativity of α it is sufficient to check that $\mathrm{e}^{h(t)}$ is convex in order to verify 2. To do so, it is sufficient to show that h itself is convex, which by definition would follow from the concavity of $\log(1-\Phi(t))$. This, however, is a consequence of $\log\Phi$ being concave, which in turn implies that $\log(1-\Phi)$ is concave (see [95, Theorem 4.2.4]).

Statement 3. follows immediately from the definition of the functions.

As for 4., observe first that, by continuous differentiability of g,

$$g^{\circ}(\cdot,z)(x;-1)=\nabla_x g(x,z_1,z_2)\cdot(-1)=-\alpha'(x)\mathrm{e}^{h(z_1)}\leq 0 \quad \forall x,z_1,z_2\in\mathbb{R},$$

whence $-1\in C_1(0)$ by Definition 7.9. On the other hand, putting $x:=1$ and $z:=(1,0)$, we have that $x\in\mathbb{B}_1(0)$, $\|z\|=1$ and

$$g^{\circ}(\cdot,z)(x;1)=\nabla_x g(1,1,0)\cdot 1=\alpha'(1)\mathrm{e}^{h(1)}=2\mathrm{e}^{h(1)}\approx 1161,$$

whereas, due to $m=2$ in this example,

$$\|z\|^{-m}\,\mathrm{e}^{\frac{\|z\|^2}{2\|L\|^2}}=\sqrt{\mathrm{e}}\approx 1.65.$$

Therefore, by Definition 7.9, $1\notin C_1(0)$. Since $C_1(0)$ is a closed cone, this together with $-1\in C_1(0)$ yields $C_1(0)=(-\infty,0]$.

For proving 5., it is sufficient to show that

$$\mathbb{B}\mathrm{e}(v,0)\subseteq(-\infty,0] \quad \forall v\in\mathbb{S}^1. \tag{7.33}$$

In order to calculate $\mathbb{B}\mathrm{e}(v,0)$ for an arbitrarily fixed $v\in\mathbb{S}^1$, we have to compute first the partial Fréchet subdifferentials $\partial_F\mathrm{e}(v,x)$ for x in a neighborhood U of 0. Define U such that $g(x,0,0)<0$ for all $x\in U$ (as a consequence of the already shown relation $g(0,0,0)<0$). If $x<0$, then, by definition of e and g,

$$\mathrm{e}(v,x)=\mu_\eta\left(\{r\geq 0\mid g(x,rLv)\leq 0\}\right)=\mu_\eta\left(\{r\geq 0\mid rLv_2\leq 1\}\right).$$

Hence, for $x<0$, $\mathrm{e}(v,x)$ does not depend on its first argument locally around x. Therefore, $\partial_F\mathrm{e}(v,x)=\{0\}$ for all $x<0$. Now, consider some $x\in U$ with $x\geq 0$ and $x^*\in\partial_x^F\mathrm{e}(v,x)$. If $v\in I(x)$, then $\partial_F\mathrm{e}(v,x)\subseteq\{0\}$ (see Corollary 7.8(ii)). If, in contrast, $v\in F(x)$, then, by

Theorem 7.7 (putting $w := \pm 1$ there and observing that, by continuous differentiability of g, the partial Clarke subdifferentials reduce to partial gradients),

$$x^* = \frac{-\chi\left(\rho(v,x)\right)\nabla_x g\left(x,\rho(v,x)Lv\right)}{\left\langle \nabla_z g\left(x,\rho(v,x)Lv\right),Lv\right\rangle} = \frac{-2x\mathrm{e}^{h(\rho(v,x)v_1)}\chi\left(\rho(v,x)\right)}{\left\langle \nabla_z g\left(x,\rho(v,x)Lv\right),Lv\right\rangle} \leq 0.$$

Here, the inequality relies on $x \geq 0$, on χ being positive as a density and on

$$\left\langle \nabla_z g\left(x,\rho(v,x)Lv\right),Lv\right\rangle \geq \frac{-g(x,0,0)}{\rho(v,x)} > 0$$

by Lemma 7.3. Altogether, we have shown that $\partial_F \mathrm{e}(v,x) \subseteq (-\infty,0]$ for all $x \in U$. This entails that also $\mathbb{B}\mathrm{e}(x,0) \subseteq (-\infty,0]$. Since $v \in \mathbb{S}^1$ has been fixed arbitrarily, the desired relation (7.33) follows.

In order to show 6. we provide first a formula for the probability function φ. If $t \leq 0$, then, by definition of g,

$$\varphi(t) = \mathbb{P}\left(g\left(x,\xi_1,\xi_2\right) \leq 0\right) = \mathbb{P}\left(\xi_2 \leq 1\right) = \Phi(1)$$

because $\xi_2 \sim \mathcal{N}(0,1)$ by the distribution assumption on ξ in Example 7.15. If $t > 0$, then, again by the assumed distribution of ξ,

$$\begin{aligned}
\varphi(t) &= \mathbb{P}\left(\xi_2 \leq 1 - t^2\mathrm{e}^{h(\xi_1)}\right) = \frac{1}{2\pi}\int\limits_{-\infty}^{\infty}\left(\int\limits_{-\infty}^{1-t^2\mathrm{e}^{h(z_1)}} \mathrm{e}^{-\left(z_1^2+z_2^2\right)/2}\mathrm{d}z_2\right)\mathrm{d}z_1 \\
&= \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty}\mathrm{e}^{-z_1^2/2}\cdot\frac{1}{\sqrt{2\pi}}\left(\int\limits_{-\infty}^{1-t^2\mathrm{e}^{h(z_1)}} \mathrm{e}^{-z_2^2/2}\mathrm{d}z_2\right)\mathrm{d}z_1 \\
&= \frac{1}{\sqrt{2\pi}}\int\limits_{-\infty}^{\infty}\mathrm{e}^{-s^2/2}\cdot\Phi\left(1-t^2\mathrm{e}^{h(s)}\right)\mathrm{d}s.
\end{aligned}$$

Now, we are going to show that φ fails to be locally Lipschitz around 0. Observe first that, since Φ is increasing as a distribution function, h is increasing too by its definition. Then, for any s,t satisfying $s \geq \Phi^{-1}\left(1-\sqrt{t}\right)$ (recall that Φ is strictly increasing and so its inverse exists) it holds that

$$h(s) \geq h\left(\Phi^{-1}\left(1-\sqrt{t}\right)\right) = -1 - \log t^2.$$

Therefore, $t^2\mathrm{e}^{h(s)} \geq \mathrm{e}^{-1}$. Thus, we have shown that

$$\Phi(1) - \Phi\left(1-t^2\mathrm{e}^{h(s)}\right) \geq \Phi(1) - \Phi\left(1-\mathrm{e}^{-1}\right) =: \varepsilon \quad \forall s,t: s \geq \Phi^{-1}\left(1-\sqrt{t}\right).$$

With Φ being strictly increasing, we have that $\varepsilon > 0$. Now, for any $t > 0$, we calculate

$$
\begin{aligned}
\varphi(0) - \varphi(t) &= \Phi(1) - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{e}^{-s^2/2} \cdot \Phi\left(1 - t^2 \mathrm{e}^{h(s)}\right) \mathrm{d}s \\
&= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathrm{e}^{-s^2/2} \cdot \left(\Phi(1) - \Phi\left(1 - t^2 \mathrm{e}^{h(s)}\right)\right) \mathrm{d}s \\
&\geq \varepsilon \frac{1}{\sqrt{2\pi}} \int_{\Phi^{-1}\left(1-\sqrt{t}\right)}^{\infty} \mathrm{e}^{-s^2/2} \mathrm{d}s = \varepsilon \left(1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\Phi^{-1}\left(1-\sqrt{t}\right)} \mathrm{e}^{-s^2/2} \mathrm{d}s \right) \\
&= \varepsilon \left(1 - \Phi\left(\Phi^{-1}\left(1 - \sqrt{t}\right)\right)\right) = \varepsilon\sqrt{t}.
\end{aligned}
$$

Since $\varepsilon > 0$, φ fails to be locally Lipschitz around 0, which finally shows 6.

7.6.2 Example 7.16

We verify in this Appendix properties 1.-6. in Example 7.16. The continuous differentiability of g stated in 1. is obvious. By nonnegativity of α it is sufficient to check that $h(t) := \frac{\exp(t^2/2)}{t^2+4}$ is convex in order to verify 2., it is enough to compute the second derivative of the function h, then one gets $h''(t) = \frac{\mathrm{e}^{t^2/2}(t^6+5t^4+14t^2+8)}{(t^2+4)^3} > 0$ for every $t \in \mathbb{R}$ and it implies the convexity of h. Statement 3., 4. and 5. follow directly from the definition of g. In order to prove 6. we have to verify the hypotheses of Corollary 7.21, more precisely, we check that the constant R must be zero and condition (7.27) holds true in some neighborhood of $x = 0$.

Indeed, using the calculus with the same estimate of Proposition 7.19, one gets

$$
\partial_F \, \mathrm{e}(x, v) \begin{cases} = \{0\} & \text{if } x < 0, \\ \subseteq \{0\} & \text{if } v \in I(x) \\ = \frac{-\chi(\rho(x,v))}{\langle \nabla_z g(x, \rho(x,v)v), v\rangle} \nabla_x g(x, \rho(x,v)v)) & \text{if } v \in F(x), \end{cases}
$$

where

$$
\nabla_x g(x, z) = \begin{cases} 2x \frac{\exp(z_1^2/2)}{z_1^2+4} & \text{if } x \geq 0, \\ 0 & \text{if } x < 0, \end{cases}
$$

$$
\nabla_z g(x, z) = (\alpha(x) \frac{\exp(z_1^2/2) z_1 (z_1^2 + 2))}{(z_1^2 + 4)^2}, 1).
$$

Then (7.27) holds true for every $x \neq 0$. We are going to prove that $\partial_M \, \mathrm{e}(0, v) = \{0\}$ for almost all $v \in \mathbb{S}^1$; the previous is trivial for $v \in F(0)$. Now consider $v = (v_1, v_2) \in I(0)$; it is enough to verify that for any sequence $x_n > 0$ with $x_n \to 0$ and $v \in F(x_n)$ one has

$$
\lim_{n \to +\infty} \frac{-\chi(\rho(x_n, v))}{\langle \nabla_z g(x_n, \rho(x_n, v)v), v\rangle} \nabla_x g(x_n, \rho(x, v)v)) = 0.
$$

Define

$$x_n^* := \frac{-\chi(\rho(x_n,v))}{\langle \nabla_z g(x_n, \rho(x_n,v)v), v\rangle} \nabla_x g(x_n, \rho(x,v)v)). \tag{7.34}$$

From Lemma 7.3 $\frac{-1}{\langle \nabla_z g(x_n,\rho(x_n,v)v),v\rangle} \geq \frac{\rho(x_n,v}{g(x_n,0)}$ and from equation $g(x_n, \rho(x_n,v)Lv) = 0$ we obtain $x_n = \sqrt{\frac{1-v_2\rho(x_n,v)}{h(v_1\rho(x_n,v))}}$. Thus, replacing in (7.34)

$$0 \geq x_n^* \geq \frac{2}{x_n/2-1}\rho(x_n,v)\exp((-1+v_1^2)\rho^2(x_n,v)/2)\sqrt{\frac{(1-v_2\rho(x_n,v))(v_1^2\rho(x_n,v)^2+4)}{\exp(v_1^2\rho^2(x_n,v)/2)}}.$$

Since $\rho(x_n,v) \to \infty$ (see Lemma 7.1) the right side of the above equation goes to zero, because for every $(v_1,v_2) \in \mathbb{S}^1$

$$\lim_{r\to\infty} r^2(1-v_2)(v_1^2r^2+4)\exp((-1+v_1^2/2)r^2) = 0.$$

Conclusions

In what follows we give the main achievements of this thesis:

1. In this work we have provided a generalization of the variational characterization of convexity given in [107, Theorem 10]. This characterization relies on the epi-pointedness property and techniques of convex analysis. Our generalization corresponds to Theorem 2.5 and it is a relation between the closed convex hull of the function and the inf-convolution of the function and the support function of the domain of its conjugate. Particularly, this relation gives us the convexity of the function provided that the domain of the conjugate function is dense (see Corollary 2.8).

2. We have shown that the class of convex proper and lower-semicontinuous epi-pointed convex functions satisfies useful variational properties in any locally convex space. More precisely, we have shown that this class of functions enjoys many important properties similar to the ones of convex and lower semicontinuous functions in Banach spaces. In this part of the thesis we have proved that the class of epi-pointed lower-semicontinuous convex functions, defined on any locally convex space, satisfy the Brøndsted-Rockafellar theorem. We also obtain other important results in the same spirit, Theorem 3.10 for the maximal monotonicity of the subdifferential, Theorems 3.12 and 3.14 for the subdifferential limiting calculus rules for functions defined in locally convex spaces, and others.

3. We have introduce the definition of a family of subdifferential(see Definition 4.1), which allows us to extend to locally convex spaces some important results in the theory of subdifferentials for non-convex functions defined in Banach spaces. The main result of this investigation is a generalization of Zagrodny's Mean Value Theorem (see Theorem 4.9). Using this result we extend theorems of integration of subdifferentials (see Theorem 4.13) and a characterization of the convexity in terms of the monotonicity of the subdifferential (see Theorem 4.20).

4. We have investigated the subdifferential of some class of convex integral functionals. We have established a general formula (see Theorem 5.9), which is valid in any arbitrary locally convex space. This formula bypass the use of techniques of measurable selections, which are principally given in separable spaces. We have used this result to generalize many well-known formulas in the literature, for example Corollary 5.21 and Corollary 5.23.

5. We have studied the subdifferential of non-convex integral functionals. In this scenario

we have generalized the sequential formulas given in [64, Theorem 1 and 2] and [80]. Posteriorly, using this sequential formulas we investigated Limiting/Mordukhovich subdifferential, the G-subdifferential and the Clarke subdifferential. We introduce the notion of the *Integrable compact sole* property (see Definition 6.20). This property allows us to establish Theorem 6.21, which generalizes the classical result about the interchange between the Clarke subdifferential and the sign of integral (see e.g. [23, Theorem 2.7.2]).

6. We have applied our result to the calculus of the subdifferential of Gaussian probability functions. We have calculated a general upper-estimate for the Limiting/Murdokhovich of Gaussian probability functions (see Theorem 7.12). Later, this result has been applied to a finite system of smooth inequalities. In this setting we generalize the result of [123] about the Clarke subdifferential of probability functions.

Future works:

We propose to continue the present research in the framework of the following problems:

(i) Optimality conditions, duality theory and stability aspects of stochastic optimization problems.

(ii) Approximate subdifferential theory of convex and nonconvex integral functions outside separable Banach spaces. I will explore techniques based on separable reductions.

(iii) Variational properties of probability functions, involving possibly infinite smooth and nonsmooth inequalities.

(iv) Real-world applications, including models of eco-industrial parks.

References

[1] C. D. Aliprantis and K. C. Border. *Infinite-dimensional analysis*, volume 4 of *Studies in Economic Theory*. Springer-Verlag, Berlin, 1994. A hitchhiker's guide.

[2] H. Attouch. *Variational convergence for functions and operators*. Applicable Mathematics Series. Pitman (Advanced Publishing Program), Boston, MA, 1984.

[3] R. J. Aumann. Integrals of set-valued functions. *J. Math. Anal. Appl.*, 12:1–12, 1965.

[4] E. J. Balder and A. R. Sambucini. Fatou's lemma for multifunctions with unbounded values in a dual space. *J. Convex Anal.*, 12(2):383–395, 2005.

[5] J. Benoist and A. Daniilidis. Integration of Fenchel subdifferentials of epi-pointed functions. *SIAM J. Optim.*, 12(3):575–582 (electronic), 2002.

[6] J. Benoist and J.-B. Hiriart-Urruty. What is the subdifferential of the closed convex hull of a function? *SIAM J. Math. Anal.*, 27(6):1661–1679, 1996.

[7] V. I. Bogachev. *Measure theory. Vol. I, II.* Springer-Verlag, Berlin, 2007.

[8] J. M. Borwein. A note on ε-subgradients and maximal monotonicity. *Pacific J. Math.*, 103(2):307–314, 1982.

[9] J. M. Borwein. Proximality and Chebyshev sets. *Optim. Lett.*, 1(1):21–32, 2007.

[10] J. M. Borwein and W. B. Moors. Separable determination of integrability and minimality of the Clarke subdifferential mapping. *Proc. Amer. Math. Soc.*, 128(1):215–221, 2000.

[11] J M. Borwein, B. S. Mordukhovich, and Y. Shao. On the equivalence of some basic principles in variational analysis. *J. Math. Anal. Appl.*, 229(1):228–257, 1999.

[12] J. M. Borwein and L. Yao. Legendre-type integrands and convex integral functions. *J. Convex Anal.*, 21(1):261–288, 2014.

[13] J. M. Borwein and Q. J. Zhu. A survey of subdifferential calculus with applications. *Nonlinear Anal.*, 38(6, Ser. A: Theory Methods):687–773, 1999.

[14] J. M. Borwein and Q. J. Zhu. *Techniques of variational analysis*, volume 20 of *CMS*

Books in Mathematics/Ouvrages de Mathématiques de la SMC. Springer-Verlag, New York, 2005.

[15] N. Bourbaki. *Éléments de mathématique. VIII. Première partie: Les structures fondamentales de l'analyse. Livre III: Topologie générale. Chapitre IX: Utilisation des nombres réels en topologie générale*. Actualités Sci. Ind., no 1045. Hermann et Cie., Paris, 1948.

[16] N. Bourbaki. *Integration. I. Chapters 1–6*. Elements of Mathematics (Berlin). Springer-Verlag, Berlin, 2004. Translated from the 1959, 1965 and 1967 French originals by Sterling K. Berberian.

[17] N. Bourbaki. *Integration. II. Chapters 7–9*. Elements of Mathematics (Berlin). Springer-Verlag, Berlin, 2004. Translated from the 1963 and 1969 French originals by Sterling K. Berberian.

[18] J. S. Brauchart, E. B. Saff, I. H. Sloan, and R. S. Womersley. QMC designs: optimal order quasi Monte Carlo integration schemes on the sphere. *Math. Comp.*, 83(290):2821–2851, 2014.

[19] A. Brøndsted and R. T. Rockafellar. On the subdifferentiability of convex functions. *Proc. Amer. Math. Soc.*, 16:605–611, 1965.

[20] B. Cascales, V. Kadets, and J. Rodríguez. The Gelfand integral for multi-valued functions. *J. Convex Anal.*, 18(3):873–895, 2011.

[21] C. Castaing. Sur les multi-applications mesurables. *Rev. Française Informat. Recherche Opérationnell*, 1(1):91–126, 1967.

[22] C. Castaing and M. Valadier. *Convex analysis and measurable multifunctions*. Lecture Notes in Mathematics, Vol. 580. Springer-Verlag, Berlin-New York, 1977.

[23] F. H. Clarke. *Optimization and nonsmooth analysis*, volume 5 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, second edition, 1990.

[24] B. Cornet, Victor-Filipe Martins-da Rocha, et al. Fatou's lemma for unbounded Gelfand integrable mappings. *preprint*, 109, 2002.

[25] R. Correa, Y. García, and A. Hantoute. Integration formulas via the Fenchel subdifferential of nonconvex functions. *Nonlinear Anal.*, 75(3):1188–1201, 2012.

[26] R. Correa and A. Hantoute. New formulas for the Fenchel subdifferential of the conjugate function. *Set-Valued Var. Anal.*, 18(3-4):405–422, 2010.

[27] R. Correa and A. Hantoute. Subdifferential of the conjugate function in general banach spaces. *TOP*, 20(2):328–346, 2012.

[28] R. Correa and A. Hantoute. Lower semicontinuous convex relaxation in optimization.

SIAM J. Optim., 23(1):54–73, 2013.

[29] R. Correa, A. Hantoute, and P. Pérez-Aros. Complete characterization of the subdifferential of convex integral functionals.

[30] R. Correa, A. Hantoute, and P. Pérez-Aros. On Brøndsted–Rockafellar's Theorem for convex lower semicontinuous epi-pointed functions in locally convex spaces,.

[31] R. Correa, A. Hantoute, and P. Pérez-Aros. On the Klee–Saint Raymond's Characterization of Convexity. *SIAM J. Optim.*, 26(2):1312–1321, 2016.

[32] R.. Correa, A. Hantoute, and D. Salas. Integration of nonconvex epi-pointed functions in locally convex spaces. *J. Convex Anal.*, 23(2), 2016. To appear.

[33] R. Correa, A. Hantoute, and D. Salas. Integration of nonconvex epi-pointed functions in locally convex spaces. *J. Convex Anal.*, 23(2):511–530, 2016.

[34] R. Correa and A. Jofré. Some properties of semismooth and regular functions in nonsmooth analysis. In *Recent advances in system modelling and optimization (Santiago, 1984)*, volume 87 of *Lect. Notes Control Inf. Sci.*, pages 69–85. Springer, Berlin, 1986.

[35] R. Correa, A. Jofré, and L. Thibault. Subdifferential characterization of convexity. In *Recent advances in nonsmooth optimization*, pages 18–23. World Sci. Publ., River Edge, NJ, 1995.

[36] M. Cúth and M. Fabian. Asplund spaces characterized by rich families and separable reduction of Fréchet subdifferentiability. *J. Funct. Anal.*, 270(4):1361–1378, 2016.

[37] M. Cúth and M. Fabian. Rich families and projectional skeletons in Asplund WCG spaces. *J. Math. Anal. Appl.*, 448(2):1618–1632, 2017.

[38] I. Deák. Subroutines for computing normal probabilities of sets—computer experiences. *Ann. Oper. Res.*, 100:103–122 (2001), 2000. Research in stochastic programming (Vancouver, BC, 1998).

[39] G. Debreu. *Theory of value: an axiomatic analysis of economic equilibrium.* Cowles Foundation for Research in Economics at Yale University, Monograph 17. John Wiley & Sons, Inc., New York; Chapman & Hall, Ltd., London, 1959.

[40] J.-P. Dedieu. Cônes asymptotes d'ensembles non convexes. *Bull. Soc. Math. France Mém.*, (60):31–44, 1979. Analyse non convexe (Proc. Colloq., Pau, 1977).

[41] R. Deville, G. Godefroy, and V. Zizler. *Smoothness and renormings in Banach spaces*, volume 64 of *Pitman Monographs and Surveys in Pure and Applied Mathematics*. Longman Scientific & Technical, Harlow; copublished in the United States with John Wiley & Sons, Inc., New York, 1993.

[42] J. Diestel and J. J. Uhl, Jr. *Vector measures.* American Mathematical Society, Providence, R.I., 1977. With a foreword by B. J. Pettis, Mathematical Surveys, No. 15.

[43] J. Dieudonné. Sur la séparation des ensembles convexes. *Math. Ann.*, 163:1–3, 1966.

[44] O. Ditlevsen and H. O. Madsen. *Structural reliability methods*, volume 178. Wiley New York, 1996.

[45] I. Ekeland. Nonconvex minimization problems. *Bull. Amer. Math. Soc. (N.S.)*, 1(3):443–474, 1979.

[46] M. Fabian, P. Habala, P. Hájek, V. Montesinos, and V. Zizler. *Banach space theory.* CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC. Springer, New York, 2011.

[47] M. Fabian and A. D. Ioffe. Separable reductions and rich families in the theory of Fréchet subdifferentials. *J. Convex Anal.*, 23(3):631–648, 2016.

[48] S. P. Fitzpatrick and S. Simons. The conjugates, compositions and marginals of convex functions. *J. Convex Anal.*, 8(2):423–446, 2001.

[49] A. Genz and F. Bretz. *Computation of multivariate normal and t probabilities*, volume 195 of *Lecture Notes in Statistics.* Springer, Dordrecht, 2009.

[50] E. G. Gol'shteĭ. Problems of best approximation by elements of a convex set and some properties of the support functionals. *Dokl. Akad. Nauk SSSR*, 173:995–998, 1967.

[51] E. G. Gol'shteĭn. *Duality theory in mathematical programming and its applications.* Izdat. "Nauka", Moscow, 1971.

[52] T. González Grandón, H. Heitsch, and R. Henrion. A joint model of probabilistic/robust constraints with application to stationary gas networks. submitted.

[53] C. Gotzes, H. Heitsch, R. Henrion, and R. Schultz. On the quantification of nomination feasibility in stationary gas networks with random load. *Math. Methods Oper. Res.*, 84(2):427–457, 2016.

[54] A. Hantoute, M. A. López, and C. Zălinescu. Subdifferential calculus rules in convex analysis: a unifying approach via pointwise supremum functions. *SIAM J. Optim.*, 19(2):863–882, 2008.

[55] R. Henrion and A. Möller. A gradient formula for linear chance constraints under Gaussian distribution. *Math. Oper. Res.*, 37(3):475–488, 2012.

[56] J.-B. Hiriart-Urruty, M. Moussaoui, A. Seeger, and M. Volle. Subdifferential calculus without qualification conditions, using approximate subdifferentials: a survey. *Nonlinear Anal.*, 24(12):1727–1754, 1995.

[57] J.-B. Hiriart-Urruty and R. R. Phelps. Subdifferential calculus using ε-subdifferentials. *J. Funct. Anal.*, 118(1):154–166, 1993.

[58] A. D. Ioffe. Survey of measurable selection theorems: Russian literature supplement.

SIAM J. Control Optim., 16(5):728–732, 1978.

[59] A. D. Ioffe. On subdifferentiability spaces. In *Fifth international conference on collective phenomena*, volume 410 of *Ann. New York Acad. Sci.*, pages 107–119. New York Acad. Sci., New York, 1983.

[60] A. D. Ioffe. Approximate subdifferentials and applications. I. The finite-dimensional theory. *Trans. Amer. Math. Soc.*, 281(1):389–416, 1984.

[61] A. D. Ioffe. Approximate subdifferentials and applications. II. *Mathematika*, 33(1):111–128, 1986.

[62] A. D. Ioffe. Approximate subdifferentials and applications. III. The metric theory. *Mathematika*, 36(1):1–38, 1989.

[63] A. D. Ioffe. Proximal analysis and approximate subdifferentials. *J. London Math. Soc. (2)*, 41(1):175–192, 1990.

[64] A. D. Ioffe. Three theorems on subdifferentiation of convex integral functionals. *J. Convex Anal.*, 13(3-4):759–772, 2006.

[65] A. D. Ioffe. On the theory of subdifferentials. *Adv. Nonlinear Anal.*, 1(1):47–120, 2012.

[66] A. D. Ioffe and V. L. Levin. Subdifferentials of convex functions. *Trudy Moskov. Mat. Obšč.*, 26:3–73, 1972.

[67] A. D. Ioffe and V. M. Tikhomirov. Duality of convex functions and extremum problems. *Russian Mathematical Surveys*, 23(6):53, 1968.

[68] A. Ionescu Tulcea and C. Ionescu Tulcea. *Topics in the theory of lifting.* Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 48. Springer-Verlag New York Inc., New York, 1969.

[69] M. Ivanov. Sequential representation formulae for G-subdifferential and Clarke subdifferential in smooth Banach spaces. *J. Convex Anal.*, 11(1):179–196, 2004.

[70] M. Ivanov and N. Zlateva. Maximal monotonicity of the subdifferential of a convex function: a direct proof. *J. Convex Anal.*, 24(4), 2017. to appear.

[71] F. Jules and M. Lassonde. Formulas for subdifferentials of sums of convex functions. *J. Convex Anal.*, 9(2):519–533, 2002. Special issue on optimization (Montpellier, 2000).

[72] F. Jules and M. Lassonde. Subdifferential estimate of the directional derivative, optimality criterion and separation principles. *Optimization*, 62(9):1267–1288, 2013.

[73] J. L. Kelley. *General topology.* Springer-Verlag, New York-Berlin, 1975. Reprint of the 1955 edition [van Nostrand, Toronto, Ont.], Graduate Texts in Mathematics, No. 27.

[74] A. Kibzun and S. Uryasev. Differentiability of probability function. *Stochastic Anal. Appl.*, 16(6):1101–1128, 1998.

[75] V. Klee. Convexity of chebyshev sets. *Mathematische Annalen*, 142(3):292–304, 1961.

[76] A. Ya. Kruger. On Fréchet subdifferentials. *J. Math. Sci. (N. Y.)*, 116(3):3325–3358, 2003. Optimization and related topics, 3.

[77] P.-J. Laurent. *Approximation et optimisation.* Hermann, Paris, 1972.

[78] V. L. Levin. Some properties of support functionals. *Mathematical notes of the Academy of Sciences of the USSR*, 4(6):900–906, 1968.

[79] V. L. Levin. Convex integral functionals and the theory of lifting. *Russian Mathematical Surveys*, 30(2):119, 1975.

[80] O. Lopez and L. Thibault. Sequential formula for subdifferential of integral sum of convex functions. *J. Nonlinear Convex Anal.*, 9(2):295–308, 2008.

[81] S. Marcellin and L. Thibault. Integration of ε-Fenchel subdifferentials and maximal cyclic monotonicity. *J. Global Optim.*, 32(1):83–91, 2005.

[82] M. Marques Alves and B. F. Svaiter. A new proof for maximal monotonicity of subdifferential operators. *J. Convex Anal.*, 15(2):345–348, 2008.

[83] G. J. Minty. On the monotonicity of the gradient of a convex function. *Pacific J. Math.*, 14:243–247, 1964.

[84] B. S. Mordukhovich. *Variational analysis and generalized differentiation. I*, volume 330 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2006. Basic theory.

[85] B. S. Mordukhovich. *Variational analysis and generalized differentiation. II*, volume 331 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2006. Applications.

[86] B. S. Mordukhovich and N. Sagara. Subdifferentials of nonconvex integral functionals in banach spaces with applications to stochastic dynamic programming. *J. Convex Anal.*, 25(2):643–673, 2018.

[87] J.-J. Moreau. Proximité et dualité dans un espace hilbertien. *Bull. Soc. Math. France*, 93:273–299, 1965.

[88] J. J. Moreau. *Fonctionnelles convexes.* Number 2. Lecture notes Séminaire "Equations aux dérivée partialles", 1966.

[89] J.-P. Penot. Subdifferential calculus without qualification assumptions. *J. Convex Anal.*, 3(2):207–219, 1996.

[90] J.-P. Penot. Image space approach and subdifferentials of integral functionals. *Optimization*, 60(1-2):69–87, 2011.

[91] J.-P. Penot. *Calculus without derivatives*, volume 266 of *Graduate Texts in Mathemat-*

ics. Springer, New York, 2013.

[92] P. Pérez-Aros. Propiedades variacionales de funciones convexas desde el análisis epi-puntado. *Engineering thesis, Universidad de Chile*, 2014.

[93] G. C.. Pflug and H. Weisshaupt. Probability gradient estimation by set-valued calculus and applications in network design. *SIAM J. Optim.*, 15(3):898–914, 2005.

[94] R. R. Phelps. *Convex functions, monotone operators and differentiability*, volume 1364 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, second edition, 1993.

[95] A. Prékopa. *Stochastic programming*, volume 324 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1995.

[96] P. Pucci and G. Vitillaro. A representation theorem for Aumann integrals. *J. Math. Anal. Appl.*, 102(1):86–101, 1984.

[97] B. Ricceri. Singular points of non-monotone potential operators. *J. Nonlinear Convex Anal.*, 16(6):1123–1129, 2015.

[98] R. T. Rockafellar. Characterization of the subdifferentials of convex functions. *Pacific J. Math.*, 17:497–510, 1966.

[99] R. T. Rockafellar. Integrals which are convex functionals. *Pacific J. Math.*, 24:525–539, 1968.

[100] R. T. Rockafellar. On the maximal monotonicity of subdifferential mappings. *Pacific J. Math.*, 33:209–216, 1970.

[101] R. T. Rockafellar. Convex integral functionals and duality. In *Contributions to nonlinear functional analysis (Proc. Sympos., Math. Res. Center, Univ. Wisconsin, Madison, Wis., 1971)*, pages 215–236. Academic Press, New York, 1971.

[102] R. T. Rockafellar. Integrals which are convex functionals. II. *Pacific J. Math.*, 39:439–469, 1971.

[103] R. T. Rockafellar. Integral functionals, normal integrands and measurable selections. pages 157–207. Lecture Notes in Math., Vol. 543, 1976.

[104] R. T. Rockafellar. Generalized directional derivatives and subgradients of nonconvex functions. *Canad. J. Math.*, 32(2):257–280, 1980.

[105] R. T. Rockafellar. *Convex analysis*. Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 1997. Reprint of the 1970 original, Princeton Paperbacks.

[106] R. T. Rockafellar and R. J.-B. Wets. *Variational analysis*, volume 317 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1998.

[107] J. Saint Raymond. Characterizing convex functions by variational properties. *J. Nonlinear Convex Anal.*, 14(2):253–262, 2013.

[108] J. Saint Raymond. Weak compactness and variational characterization of the convexity. *Mediterr. J. Math.*, 10(2):927–940, 2013.

[109] L. Schwartz. *Radon measures on arbitrary topological spaces and cylindrical measures*. Published for the Tata Institute of Fundamental Research, Bombay by Oxford University Press, London, 1973. Tata Institute of Fundamental Research Studies in Mathematics, No. 6.

[110] A. Shapiro, D. Dentcheva, and A. Ruszczyński. *Lectures on stochastic programming*, volume 9 of *MOS-SIAM Series on Optimization*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Optimization Society, Philadelphia, PA, second edition, 2014. Modeling and theory.

[111] S. Simons. A new proof of the maximal monotonicity of subdifferentials. *J. Convex Anal.*, 16(1):165–168, 2009.

[112] P. D. Taylor. Subgradients of a convex function obtained from a directional derivative. *Pacific J. Math.*, 44:739–747, 1973.

[113] L. Thibault. A generalized sequential formula for subdifferentials of sums of convex functions defined on Banach spaces. In *Recent developments in optimization (Dijon, 1994)*, volume 429 of *Lecture Notes in Econom. and Math. Systems*, pages 340–345. Springer, Berlin, 1995.

[114] L. Thibault. A note on the Zagrodny mean value theorem. *Optimization*, 35(2):127–130, 1995.

[115] L. Thibault. Sequential convex subdifferential calculus and sequential Lagrange multipliers. *SIAM J. Control Optim.*, 35(4):1434–1444, 1997.

[116] L. Thibault. Limiting convex subdifferential calculus with applications to integration and maximal monotonicity of subdifferential. In *Constructive, experimental, and nonlinear analysis (Limoges, 1999)*, volume 27 of *CRC Math. Model. Ser.*, pages 279–289. CRC, Boca Raton, FL, 2000.

[117] L. Thibault and D. Zagrodny. Integration of subdifferentials of lower semicontinuous functions on Banach spaces. *J. Math. Anal. Appl.*, 189(1):33–58, 1995.

[118] L. Thibault and D. Zagrodny. Subdifferential determination of essentially directionally smooth functions in Banach space. *SIAM J. Optim.*, 20(5):2300–2326, 2010.

[119] L. Thibault and N. Zlateva. Integrability of subdifferentials of directionally Lipschitz functions. *Proc. Amer. Math. Soc.*, 133(10):2939–2948, 2005.

[120] S. Uryas'ev. Derivatives of probability functions and integrals over sets given by inequalities. *J. Comput. Appl. Math.*, 56(1-2):197–223, 1994. Stochastic programming:

stability, numerical methods and applications (Gosen, 1992).

[121] C. Vallée and C. Zălinescu. Series of convex functions: subdifferential, conjugate and applications to entropy minimization. *J. Convex Anal.*, 23(4):1137–1160, 2016.

[122] W. van Ackooij and R. Henrion. Gradient formulae for nonlinear probabilistic constraints with Gaussian and Gaussian-like distributions. *SIAM J. Optim.*, 24(4):1864–1889, 2014.

[123] W. van Ackooij and R. Henrion. (Sub-)gradient formulae for probability functions of random inequality systems under Gaussian distribution. *SIAM/ASA J. Uncertain. Quantif.*, 5(1):63–87, 2017.

[124] W. van Ackooij, R. Henrion, A. Möller, and R. Zorgati. On probabilistic constraints induced by rectangular sets and multivariate normal distributions. *Math. Methods Oper. Res.*, 71(3):535–549, 2010.

[125] W. van Ackooij, R. Henrion, A. Möller, and R. Zorgati. On joint probabilistic constraints with Gaussian coefficient matrix. *Oper. Res. Lett.*, 39(2):99–102, 2011.

[126] W. van Ackooij, R. Henrion, A. Möller, and R. Zorgati. Joint chance constrained programming for hydro reservoir management. *Optim. Eng.*, 15(2):509–531, 2014.

[127] D. H. Wagner. Survey of measurable selection theorems. *SIAM J. Control Optimization*, 15(5):859–903, 1977.

[128] D. H. Wagner. Survey of measurable selection theorems: an update. In *Measure theory, Oberwolfach 1979 (Proc. Conf., Oberwolfach, 1979)*, volume 794 of *Lecture Notes in Math.*, pages 176–219. Springer, Berlin-New York, 1980.

[129] R. J.-B. Wets. Stochastic programming. In *Optimization*, volume 1 of *Handbooks Oper. Res. Management Sci.*, pages 573–629. North-Holland, Amsterdam, 1989.

[130] D. Zagrodny. Approximate mean value theorem for upper subderivatives. *Nonlinear Anal.*, 12(12):1413–1428, 1988.

[131] Xiyin Zheng. A series of convex functions on a Banach space. *Acta Math. Sinica (N.S.)*, 14(1):77–84, 1998.

[132] C. Zălinescu. *Convex analysis in general vector spaces.* World Scientific Publishing Co., Inc., River Edge, NJ, 2002.

[133] C. Zălinescu. A new proof of the maximal monotonicity of the sum using the Fitzpatrick function. In *Variational analysis and applications*, volume 79 of *Nonconvex Optim. Appl.*, pages 1159–1172. Springer, New York, 2005.

Index of Notation

Index of Topics