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

“DISCRETIZATION-REDUCTION IN SEMI-INFINITE OPTIMIZATION AND SUBDIFFERENTIAL CALCULUS FOR SUPREMUM FUNCTIONS”

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

STEPHANIE PAOLA CARO TORRES

PROFESOR GUÍA:
ABDERRAHIM HANTOUTE

MIEMBROS DE LA COMISIÓN:
RAFAEL CORREA FONTECILLA
FABIÁN FLORES BAZÁN
MARCO ANTONIO LÓPEZ CERDÁ
JAIME ORTEGA PALMA

Este trabajo ha sido parcialmente financiado por CMM ANID BASAL FB210005, ANID-CHILE Beca Doctorado Nacional 2021-21211837, MICIU de España y Universidad de Alicante (Contract Beatriz Galindo BEA-GAL 18/00205), y por el proyecto AICO/2021/165 de Generalitat, Valenciana.

SANTIAGO DE CHILE
2023

RESUMEN DE TESIS
PARA OPTAR AL GRADO DE
DOCTORA EN CIENCIAS DE LA INGENIERÍA,
MENCIÓN EN MODELACIÓN MATEMÁTICA
POR: STEPHANIE PAOLA CARO TORRES
FECHA: 2023
PROF. GUÍA: ABDERRAHIM HANTOUTE

Discretización y reducción en optimización semi-infinita y cálculo subdiferencial para funciones supremo

El principal objetivo de esta tesis es proporcionar esquemas generales de discretización para problemas de optimización semi-infinita que permitan reescribirlos en problemas de optimización ordinaria. Los argumentos de reducción permitirán a su vez que el número de restricciones involucradas se reduzca como máximo a la dimensión del espacio más uno, siempre que el problema original se define en un contexto finito-dimensional. Las principales herramientas de esta parte son un nuevo teorema de tipo minimax (de dimensión finita), por un lado, y, por otro, nuevas reglas de cálculo subdiferencial para la función supremo definida en espacios localmente convexos (de dimensión infinita). Estas últimas reglas se dan de forma explícita y exclusiva a través de los datos. Nuestros resultados incluyen nuevos logros y también diferentes extensiones de resultados existentes en la literatura, como los establecidos para la teoría de la discretización en [6], [24], [42], [53], y [55] entre muchos otros. En particular, obtenemos generalizaciones de [42] y [24].

Nuestro enfoque también conduce a nuevas caracterizaciones del subdiferencial de la función supremo que amplían algunos resultados recientes en [16], [35] y [43]. Nuestro enfoque permite fórmulas más explícitas ya que no apelamos al cono normal del dominio de esta función surpemo sino directamente a las funciones involucradas. Aplicados a problemas de optimización, estos resultados nos proporcionan nuevas y generales condiciones de optimalidad, de tipo Fritz-John y KKT, que resaltan el papel que desempeñan las funciones cuasi-activas y no activas en el punto de referencia.

Aplicamos los argumentos de discretización anteriores a problemas de optimización multiobjetivo semi-infinita para proporcionar caracterizaciones de soluciones débilmente eficientes utilizando subproblemas que involucran funciones objetivo con un número finito de funciones de dato. Más precisamente, el número de dichas funciones no debe exceder la dimensión del espacio más uno. Esto nos permitirá dar extensiones de algunos de los resultados en [44], [49], [50] y [59]. La parte final de esta tesis trata de problemas de optimización cuadrática que generalizan la optimización cuadrática estándar, donde el simplejo n-dimensional usual se reemplaza por una base convexa compacta de un cono convexo cerrado puntiagudo. Estableceremos un resultado de dualidad fuerte y proporcionaremos una forma explícita de la función valor (perturbada) asociada con estos problemas, proporcionando así una generalización de [3, Teorema 4].

Palabras claves: Optimización semi-infinita, funciones convexas, procesos de discretización y reducción, función supremo, teoría del subdiferencial y condiciones de optimalidad, teoría de minimax.

RESUMEN DE TESIS
PARA OPTAR AL GRADO DE
DOCTORA EN CIENCIAS DE LA INGENIERÍA,
MENCIÓN EN MODELACIÓN MATEMÁTICA
POR: STEPHANIE PAOLA CARO TORRES
FECHA: 2023
PROF. GUÍA: ABDERRAHIM HANTOUTE

Discretization-reduction in semi-infinite optimization and subdifferential calculus for supremum functions

The main objective of this thesis is to provide general discretization schemes for semi-infinite optimization problems that allow them to be rewritten into ordinary optimization problems. The reduction arguments will in turn allow the number of constraints involved to be reduced to at most the dimension of the space plus one. The main tools of this part are some new (finite-dimensional) minimax-type theorems, on the one hand, and, on the other, new subdifferential calculus rules for the subdifferential of pointwise suprema defined in (infinite dimensional) locally convex spaces. These last rules are given explicitly and exclusively through the data provided. Our results include new achievements and also different extensions of existing results in the literature, such as those established for the discretization theory in [6], [24], [42], [53] and [55] among many others. In particular, we obtain some consistent generalizations of [42] and [24].

Our approach also gives rise to new characterizations of the subdifferential of pointwise suprema that extend some recent results in [16], [35] and [43], and which constitute the main tool to obtain the aforementioned minimax theorems. Our analysis allows more explicit formulas since we do not appeal to the normal cone of the domain of these suprema but directly to the functions involved. Applied to optimization problems, these results provide us new and general optimality conditions of Fritz-John and KKT types, which highlight the role played by almost active and non-active constraints functions.

We apply the above discretization arguments to multi-objective optimization problems to provide characterizations of weakly efficient solutions using subproblems involving finitely many objective functions. More precisely, the number of such functions must not exceed the dimension of the space plus one when the underlying setting is a finite dimensional. This will allow us to give extensions of some of the results in [44], [49], [50] and [59]. The final part of this thesis deals with quadratic optimization problems that generalize standard quadratic optimization, where the usual n-dimensional simplex is replaced by a compact convex base of a pointed closed convex cone. We will establish a strong duality result for this kind of problems and provide an explicit form of the associated value (or perturbed) function that results in a useful generalization of [3, Theorem 4].

Keywords: Semi-infinite optimization, convex functions, discretization and reduction processes, supremum functions, subdifferential and optimality theory, minimax theory.

To my family, especially to my mom.

Agradecimientos

Quisiera comenzar agradeciendo a los académicos que me han formado durante este proceso. Al profesor Fabián Flores Bazán, por sus conocimientos entregados en el desarollo del pregrado, por permitirnos tener experiencias significativas en los cursos desarrollados y por mantener en alto el área de optimización. Al Profesor Abderrahim Hantoute, a quien conocí en la defensa de mi tesis de pre-grado y quien ahora intercambia roles con el profesor Flores. A él le agradezco enormemente su amabilidad, sus buenas vibras y por sobre todo agradezco su curso dictado el 2018-2, fue un placer ser su alumna, además para mí fue ese curso el punto de partida de esta tesis. También quisiera agradecer al profesor Michal Kowalczyk, por su compromiso y pasión a la hora de enseñar, para mí el programa de doctorado no sería lo mismo sin usted, gracias.

Me permito también agradecer a las instituciones me han acogido a lo largo de este proceso: Departamento de Ingeniería Matemática y Centro de Modelamiento Matemático de la Universidad de Chile, Departamento de Ingeniería Matemática de la Universidad de Concepción y Departamento de Matemática de la Universidad de Alicante, España. Por supuesto que cuando me refiero a instituciones, implícitamente pienso en las personas que las conforman.

A las personas que conocí durante el periodo del doctorado, con quienes compartimos intereses, a pesar de las diversas áreas, agradezco los encuentros en las oficinas, los almuerzos en grupo y las salidas a tomar café, gracias.

Mención a parte para los miembros del grupo de optimización de la Universidad de Alicante, entre ellos los profesores: Marco Antonio López, Miguel Ángel Goberna, Francisco Aragón y Abderrahim Hantoute, quienes me acogieron durante mi pasantía de investigación y a quienes guardo mucho aprecio y respeto.

Cada uno de los lugares mencionados, tiene personas con las que me he relacionado compartido experiencias, reído muchas veces y sin duda he aprendido, sin ellas este camino que he recorrido no sería el mismo.

Finalmente, agradecer en conjunto a los profesores Rafael Correa, Fabián Flores, Marco Antonio López y Jaime Ortega, por aceptar ser parte de mi comisión y estar en esta etapa final del programa.

Mis sinceros agradecimientos a cada persona que ha sido, de alguna forma, parte de mi viaje. Quisiera compartir un poema que me acompañado desde el colegio, cuyo autor es Konstantino Kavafis.

Viaje a Ítaca

Cuando emprendas tu viaje a Ítaca pide que el camino sea largo, lleno de aventuras, lleno de experiencias. No temas a los lestrigones ni a los cíclopes ni al colérico Poseidón, seres tales jamás hallarás en tu camino, si tu pensar es elevado, si selecta es la emoción que toca tu espíritu y tu cuerpo. Ni a los lestrigones ni a los cíclopes ni al salvaje Poseidón encontrarás, si no los llevas dentro de tu alma, si no los yergue tu alma ante ti.

Pide que el camino sea largo. Que muchas sean las mañanas de verano en que llegues -¡con qué placer y alegría!- a puertos nunca vistos antes. Detente en los emporios de Fenicia y hazte con hermosas mercancías, nácar y coral, ámbar y ébano y toda suerte de perfumes sensuales, cuantos más abundantes perfumes sensuales puedas. Ve a muchas ciudades egipcias a aprender, a aprender de sus sabios.

Ten siempre a Ítaca en tu mente. Llegar allí es tu destino. Mas no apresures nunca el viaje. Mejor que dure muchos años y atracar, viejo ya, en la isla, enriquecido de cuanto ganaste en el camino sin aguantar a que Ítaca te enriquezca.

Ítaca te brindó tan hermoso viaje. Sin ella no habrías emprendido el camino. Pero no tiene ya nada que darte.

Aunque la halles pobre, Ítaca no te ha engañado. Así, sabio como te has vuelto, con tanta experiencia, entenderás ya qué significan las Ítacas.

TABLE OF CONTENT

1 Introduction **1**
1.1 English version 1
1.2 Spanish version 5

2 Notation and preliminary results **9**
2.1 Convex analysis 10
2.2 Topology and convexity 14
2.3 Subdifferential analysis 22
2.4 Quadratic forms 26

3 Discretization and reduction of Infsup and SIP optimization problems **28**
3.1 Introduction 28
3.2 Reduction of Infsup problems 30
3.3 Reduction of SIP problems 37
3.4 Consequences 41

4 Subdifferential calculus: characterizations of the normal cone to the domain of the supremum function **45**
4.1 Introduction 45
4.2 General characterizations of the normal cone 46
4.3 Characterization of the normal cone in the compact-continuous setting . . . 58

5 Subdifferential calculus for pointwise suprema **62**
5.1 Introduction 62
5.2 Subdifferential of the supremum function 64
5.3 Reduction process and subdifferential of the supremum 71
5.4 Danskin's Theorem 73
5.5 Optimality conditions 76

6 Multiobjective optimization **78**
6.1 Introduction 78
6.2 Reduction processes in multiobjective optimization 79
6.3 Characterization of weakly efficient solutions 82

7 Strong duality for some quadratic problems **84**
7.1 Introduction 84

7.2 Lagrangian duality 85
7.3 General problem 89
7.4 Standard quadratic problems 94

8 Future work **96**

Bibliography **96**

List of Figures

2.1 $f \in \Gamma_0(X)$ is minorized by a continuous affine mapping 15

3.1 Graphs of $f_m,\ m \in \mathbb{N},\ f_\infty$ 37
3.2 Constraint half-spaces for each $t \in (0,1]$. 41

4.1 $\partial_\varepsilon(\rho_{t,\varepsilon} f_t)(0)$ for each $t \in \mathbb{R}_+$ 51

6.1 $\mathrm{E}(\{1,2\}, S)$ in red, $\mathrm{WE}_{\{1,2\}}(S)$ in yellow. 80

7.1 Perturbed Function, ψ. 90
7.2 Epigraph of perturbed Function, $\operatorname{epi}\psi$. 91

Chapter 1

Introduction

1.1 English version

This thesis delves into the theory of Convex Analysis, visiting different topics dealing with optimization problems that are studied from many points of view. The analysis carried out is done through several steps organized into chapters, each with its own interest but all aimed at better understanding the arguments of discretization and reduction in optimization theory, namely in semi-infinite optimization and subdifferential theory.

The objectives of the thesis include the following items:

- Provide new minimax-type theorems of the form $\max_{t \in T} \inf_{x \in \mathbb{R}^n} f(x,t) = \inf_{x \in \mathbb{R}^n} \max_{t \in T} f(x,t)$ that allow reducing the number of the data functions involved, $f_t(\cdot) := f(\cdot, t)$, $t \in T$ (with T possibly infinite), to at most $n+1$ functions, where n is the dimension of the underlying Euclidean space. It is worth observing that the classical minimax theorem cannot be used here, at least not directly, since our model lacks the necessary hypotheses required by the classical minimax theorem. In a second step, we analyze the relationship between the optimal solutions of a given semi-infinite programming problem and the optimal solutions of its sub-problems given with finitely many constraints and whose number does not exceed n.
- Provide general characterizations of the subdifferential of the supremum of an arbitrary family of convex functions defined on locally convex spaces, and indexed in arbitrary sets. The desired characterizations will be given by means exclusively of the data functions, without involving additional concepts like the effective domain of the supremum or finite-dimensional sections of it.
- Provide applications to multi-objective and quadratic optimization.

The importance of minimax-type theorems, semi-infinite programming and Infsup problems is a well-known fact. There exists a lot of works in the literature related with this three topics (see, e.g., [42], [5] and [4]). For finite-dimensional spaces, there are many results in the literature dealing with the idea of reducing the number of sets or functions involved in optimization problems either within the objective (or multi-objective) function or the constraints. The most emblematic achievements are the theorems of Helly and Carathéodory and their extensions; see, for example, [53], [42], [6], [24] and [55] to name a few. More specifically, the problem of characterizing the subdifferential of the supremum function has been raised since the 1960s, and there are several contributions from different authors; see, for instance, Brøndsted & Rockafellar [8], Valadier [57], Brøndsted [7], Hantoute, López & Zălinescu [35] among many others.

These minimax problems and subdifferential calculus of the supremum are also useful in multiobjective optimization problems; see, Jahn [38] and Sawaragi et al. [54] where one can find the development of this theory and applications around this type of problems.

The topics studied in this thesis respond to the main objectives cited above and are undertaken through five main chapters, each dedicated to responding to a particular objective. These are given after Chapters 2 which is dedicated to fix the notations and provide some preliminary results.

Chapter 3, entitled *Discretization and reduction of Infsup and SIP optimization problems*, provides new minimax type results together with a reduction process for semi-infinite optimization problems. Chapter 4, entitled *Subdifferential calculus: characterizations of the normal cone to the domain of the supremum function*, gives the first step towards the study of the subdifferential of pointwise suprema. Chapter 5, entitled *Subdifferential calculus for pointwise suprema*, presents new characterizations of the subdifferential of pointwise suprema given explicitly by means of the data functions. Chapter 6, entitled *Multiobjective optimization*, focuses on convex multiobjective optimization problems where the multiobjective functions includes infinitely many functions. In this case, we provide characterizations of the set of weakly efficient solutions by means of appropriately chosen sub-problems. Chapter 7, entitled *Strong duality for some quadratic problems*, deals with a generalized quadratic problem that contains as a particular case the well-known standard quadratic problem. Finally, in Chapter 8 we present some future works that will arise from this thesis.

Below we describe the contents of each chapter.

In **Chapter 3** we develop a minimax theorem which takes into account the reduction of the number of the functions involved, allowing to rewrite Infsup type problems into Infmax problems. Applied to optimization theory, these arguments will allow to transform semi-infinite programming problem (SIP, for short) into an ordinary optimization problem. The main result of this chapter is Theorem 3.1, which establishes that

$$\inf_{\mathbb{R}^n} \sup_{t\in T} f_t(x) = \max_{S\subset T,\ |S|\leq n+1} \inf_{\mathbb{R}^n} \max\{f_t + \mathrm{I}_D,\ t \in S\},$$

for an appropriately chosen set D that takes into account the geometry of the effective domains of the f_t's or of their supremum. This, in particular, covers some of the results in [24, Theorem 1] (see Theorem 3.3). The reduction process above is applied in Theorem 3.8 to SIP problems, generalizing Levin's Theorem in [42]. Some other extensions are given in this chapter such as an alternative-type theorem (Theorem 3.10), Lagrangian reformulation of SIP problems, and less restrictive qualifications conditions (Slater conditions for constraint blocks (Corollary 3.13)). The results contained in this chapter are part of the following pre-print,

Caro, S. & Hantoute, A. *Discretization and reduction of Infsup and SIP optimization problem*, pre-print.

Chapter 4 is dedicated to establish a characterization of the normal cone to the effective domain of pointwise suprema. Indeed, such a normal cone plays an important role in establishing subdifferential calculus rules for the subdifferential of the supremum as shown in different works like Hantoute, López & Zălinescu [35], López & Volle [43] and Correa, Hantoute & López [16]. Thus, its characterization by means of the data functions would allow for complete characterizations of the subdifferential of the supremum. The main result of this chapter is given in Theorem 4.7, stating that

$$\mathrm{N}_{\operatorname{dom} f}(x) = \left[\overline{\operatorname{co}}\left(\left(\bigcup_{t\in T_0} \partial_\varepsilon(\rho_{t,\varepsilon} f_t)(x)\right) \cup \left(\bigcup_{t\in T\setminus T_0} \mathrm{N}^\varepsilon_{\operatorname{dom} f_t}(x)\right)\right)\right]_\infty, \text{ for all } \varepsilon > 0,$$

where $T_0 := \{t \in T : \operatorname{cl} f_t \text{ is proper}\}$.

Chapter 5 presents a characterization of the subdifferential of the pointwise supremum function. Now, with the help of Chapter 4, we will be able to provide a formula for such a subdifferential which is written only by means of the data functions involved. For instance, given an arbitrary set T we show in Theorem 5.1 that

$$\partial f(x) = \bigcap_{\varepsilon>0} \overline{\operatorname{co}}\left(\bigcup_{t\in T_\varepsilon(x)} \partial_\varepsilon f_t(x) + \left(\varepsilon \bigcup_{t\in T_0\setminus T_\varepsilon(x)} \partial_\varepsilon(\rho_{t,\varepsilon} f_t)(x)\right) \cup \left(\bigcup_{t\in T\setminus T_0} \mathrm{N}^\varepsilon_{\operatorname{dom} f_t}(x) \cup \{\theta\}\right)\right).$$

The special case of compact index sets is also studied in Theorem 5.3 and Corollary 5.4. Our results provide direct characterizations of the recent papers Correa, Hantoute & López [12] and Hantoute & López [34]. The results contained in Chapters 4 and 5 are part of the following pre-print,

Caro, S. & Hantoute, A. *Highlighting the role of active and non-active functions in optimization and subdifferential calculus*, pre-print.

Chapter 6 is devoted to multiobjective semi-infinite-programming problems given in the form

$$\min_{x\in S}\{f_t(x) : t \in T\},$$

where $f_t : \mathbb{R}^n \to \mathbb{R}_\infty$, $t \in T$, is a family of extended real-valued functions with T being an arbitrary set, and $S \subset \mathbb{R}^n$ is a nonempty convex set. First, we focus on the description of the set of weakly efficient solutions of the problem above by using the information provided by the associated sub-problems that involve no more that $n+1$ objective functions. This results has been also analyzed elsewhere (for instance, in Lowe et al. [44] and Ward [59]) when T is finite. Our approach uses the arguments developed in Chapter 3. The main result of this chapter is Theorem 6.4 that states that

$$\mathrm{WE}_T(S) = \bigcup_{B\subset T, |B|\leq n+1} \mathrm{WE}_B(S \cap D).$$

In particular, we get Corollary 6.5 which recovers similar results in Plastria & Carrizosa [50].

Also, we provide characterizations of the set of weakly efficient solutions by means of the subdifferential of the associated (supremum) objective function and the normal cone to the constraints set S. Namely, in Theorem 6.6 we obtain the following equivalence that generalizes some of the results in [50]:

$$\bar{x} \in \mathrm{WE}_T(S) \Longleftrightarrow \theta \in \mathrm{N}_S(\bar{x}) + \bigcap_{\varepsilon>0} \overline{\mathrm{co}} \left(\bigcup_{t\in T} \partial_\varepsilon f_t(\bar{x}) \right).$$

The results contained in this chapter are part of the following work which is in progress,

> **Caro, S.** & Hantoute, A. *Reduction arguments in multi-objective optimization and subdifferential calculus.*

Finally, in **Chapter 7** we consider a generalization of the standard quadratic problem, which is given by

$$\mu = \inf_{g(x)=0,\ x\in C} \frac{1}{2} x^\top A x + 2r^\top x,$$

where $C \subset \mathbb{R}^n$ is a closed convex cone, $e \in \operatorname{int} C^*$ and $g(x) := e^\top x - 1$. The chapter begins with a review of the Lagrangian duality theory, emphasizing the importance of the relations given by Flores-Bazán & Mastroeni in [31]. We also study the equivalence between the solutions of the dual/primal problems and the associated perturbation function given by Flores-Bazán, Jourani & Mastroeni in [30]. Our main result, given in Theorem 7.6, proves a strong duality theorem for the given quadratic problem through an explicit formulation of the perturbed function, ψ,

$$\psi(a) = \begin{cases} (1+a)^2\mu, & \text{if } a \geq -1, \\ +\infty, & \text{if } a < -1. \end{cases}$$

This analysis reveals a hidden convexity in the given quadratic problem, namely in view of the fact that the set

$$\{(f(x), g(x)) : x \in C\} + \mathbb{R}_+ \times \{0\}$$

is convex. In Theorem 7.8, we present a generalization to some of the results in [3], showing an explicit form of the function $\Theta(\lambda) = \inf_{x\in C} f(x) + \lambda(e^\top x - 1)$. The results contained in this chapter are part of the paper,

> Flores-Bazán, F., Carcamo, G. & **Caro, S.** *Extensions of the Standard quadratic optimization problem: strong duality, optimality, hidden convexity and S-lemma*, Appl. Math. Optim. 81 (2020), no. 2, 383-408.

1.2 Spanish version

Esta tesis profundiza en la teoría del Análisis Convexo, visitando diferentes temas que tratan problemas de optimización estudiados desde muchos puntos de vista. El análisis realizado se organiza a través de varias etapas organizadas en capítulos, cada uno con su propio interés pero todos dirigidos a comprender mejor los argumentos de la discretización y la reducción en la teoría de la optimización, concretamente en la optimización semi-infinita y la teoría subdiferencial.

Los objetivos de la tesis incluyen los siguientes puntos:

- Proporcionar nuevos teoremas de tipo minimax de la forma siguiente $\max_{t\in T} \inf_{x\in\mathbb{R}^n} f(x,t) = \inf_{x\in\mathbb{R}^n} \max_{t\in T} f(x,t)$ que nos permitan reducir el número de funciones de dato involucradas, $f_t(\cdot) := f(\cdot,t),\ t \in T$ (con T posiblemente infinitas) hasta a lo más $n+1$ funciones, donde n es la dimensión del espacio Euclidiano subyacente. Vale la pena señalar que el teorema clásico del minimax no se puede utilizar aquí, al menos no directamente, ya que nuestro modelo carece de las hipótesis necesarias que requiere el teorema clásico del minimax. En un segundo paso, analizamos la relación entre las soluciones óptimas de un problema de programación semi-infinito y la solución óptima de sus subproblemas con un número finito de restricciones y cuyo número no exceda n.
- Proporcionar caracterizaciones generales del subdiferencial del supremo de una familia arbitraria de funciones convexas definidas en espacios localmente convexos e indexadas en conjuntos arbitrarios. Las caracterizaciones deseadas se daran exclusivamente mediante las funciones dato, sin involucrar conceptos adicionales como el dominio efectivo del supremo o secciones finito-dimensional de este conjunto.
- Proporcionar aplicaciones para optimización cuadrática y multiobjetivo.

La importancia de los teoremas de tipo minimax, la programación semi-infinita y los problemas Infsup es un hecho bien conocido. Existe una gran variedad de trabajos en la literatura relacionados con estos tres temas (ver, por ejemplo, [42], [5] y [4]). Para espacios de dimensión finita, hay muchos resultados en la literatura que tratan con la idea de reducir el número de conjuntos o funciones involucradas en problemas de optimización, ya sea dentro de la función objetivo (o multiobjetivo) o de las restricciones. Los resultados más emblemáticos son los teoremas de Helly y Carathéodory, y sus extensiones; véase, por ejemplo, [53], [42], [6], [24] y [55] por nombrar algunos. Más específicamente, el problema de caracterizar la subdiferencial de la función suprema se plantea desde la década de 1960, y existen varias contribuciones de distintos autores; véase, por ejemplo, Brøndsted & Rockafellar [8], Valadier [57], Brøndsted [7], Hantoute, López & Zălinescu [35] entre muchos otros.

Estos problemas de tipo minimax y el cálculo subdiferencial del supremo también son útiles en problemas de optimización multiobjetivo; ver Jahn [38] y Sawaragi et al. [54], donde se puede encontrar el desarrollo de esta teoría y aplicaciones alrededor de este tipo de problemas.

Los temas estudiados en esta tesis responden a los principales objetivos citados anteriormente y se abordan a través de cinco capítulos principales, cada uno de ellos dedicado a dar respuesta a un objetivo particular. Estos se dan después del Capítulo 2, que está dedicado a fijar las notaciones y proporcionar algunos resultados preliminares.

El capítulo 3, titulado *Discretization and reduction of Infsup and SIP optimization problems*, proporciona nuevos resultados de tipo minimax junto con un proceso de reducción para problemas de optimización semi-infinita. El capítulo 4 *Subdifferential calculus: characterizations of the normal cone to the domain of the supremum function*, da el primer paso hacia el estudio del subdiferencial de la función supremo puntual. El capítulo 5, titulado *Subdifferentials calculus for pointwise suprema*, presenta una nueva caracterización del subdiferencial de suprema puntual dada explícitamente por medio de las funciones de dato. El capítulo 6, titulado *Multiobjective optimization*, se centra en problemas de optimización multiobjetivo convexos donde las funciones objetivo pueden ser infinitas. En este caso, proporcionamos caracterizaciones del conjunto de soluciones débilmente eficientes mediante subproblemas elegidos adecuadamente. Capítulo 7 de nombre *Strong duality for some quadratic problems*, trata de un problema cuadrático generalizado que contiene como caso particular el conocido problema cuadrático estándar. Finalmente, en el Capítulo 8 presentamos algunos trabajos futuros que surgirán de esta tesis.

A continuación describimos el contenido de cada capítulo.

En **Capítulo 3** desarrollamos un teorema minimax que tiene en cuenta la reducción del número de funciones involucradas, permitiendo reescribir problemas de tipo Infsup en problemas de Infmax. Aplicados a la teoría de la optimización, estos argumentos permitirán transformar un problema de programación semi-infinita (SIP, para abreviar) en un problema de optimización ordinario. El principal resultado de este capítulo es el Teorema 3.1, el cual establece que

$$\inf_{\mathbb{R}^n} \sup_{t\in T} f_t(x) = \max_{S\subset T,\ |S|\leq n+1} \inf_{\mathbb{R}^n} \max\{f_t + \mathrm{I}_D,\ t \in S\},$$

para un conjunto D apropiado, que tenga en cuenta la geometría de los dominios efectivos de las funciones f_t's o de su supremo. Esto, en particular, cubre algunos de los resultados en [24, Teorema 1] (ver Teorema 3.3). El proceso de reducción anterior se aplica en el Teorema 3.8 a problemas SIP, generalizando el Teorema de Levin en [42, Teorema 1]. En este capítulo se dan algunas otras extensiones, como un teorema de tipo alternativo (Teorema 3.10), la reformulación lagrangiana de problemas SIP y condiciones de calificación menos restrictivas (condiciones de Slater para bloques de restricciones (Corolario 3.13). Los resultados contenidos en este capítulo son parte del siguiente pre-print,

> **Caro, S.** & Hantoute, A. *Discretization and reduction of Infsup and SIP optimization problem*, pre-print.

Capítulo 4 está dedicado a establecer una caracterización del cono normal al dominio efectivo de la función supremo puntual. De hecho, dicho cono normal juega un papel importante en el establecimiento de reglas de cálculo subdiferencial para el subdiferencial de la función supremo, como se muestra en diferentes trabajos, ver por ejemplo, Hantoute, López & Zălinescu [35], López & Volle [43] y Correa, Hantoute & López [16]. Así, su caracterización mediante funciones dato permitiría caracterizaciones completas del subdiferencial de la función supremo. El principal resultado de este capítulo se da en Teorema 4.7, y establece que

$$\mathrm{N}_{\operatorname{dom} f}(x) = \left[\overline{\mathrm{co}}\left(\left(\bigcup_{t\in T_0}\partial_\varepsilon(\rho_{t,\varepsilon}f_t)(x)\right)\bigcup\left(\bigcup_{t\in T\setminus T_0}\mathrm{N}^{\varepsilon}_{\operatorname{dom} f_t}(x)\right)\right)\right]_\infty, \text{ para todo } \varepsilon>0,$$

donde $T_0 := \{t\in T : \operatorname{cl} f_t \text{ es propia}\}$.

Capítulo 5 presenta una caracterización del subdiferencial de la función suprema puntual. Ahora, con la ayuda del Capítulo 4, podremos proporcionar una fórmula para dicho subdiferencial que se escribe solo mediante las funciones dato involucradas. Por ejemplo, dado un conjunto arbitrario T mostramos en Teorema 5.1 que

$$\partial f(x) = \bigcap_{\varepsilon>0}\overline{\mathrm{co}}\left(\bigcup_{t\in T_\varepsilon(x)}\partial_\varepsilon f_t(x) + \left(\varepsilon\bigcup_{t\in T_0\setminus T_\varepsilon(x)}\partial_\varepsilon(\rho_{t,\varepsilon}f_t)(x)\right)\cup\left(\bigcup_{t\in T\setminus T_0}\mathrm{N}^{\varepsilon}_{\operatorname{dom} f_t}(x)\cup\{\theta\}\right)\right).$$

El caso especial de conjuntos de índices compactos también se estudia en Teorema 5.3 y en Corolario 5.4. Nuestros resultados proporcionan caracterizaciones directas de los artículos recientes de Correa, Hantoute & López [12] y Hantoute & López [34]. Los resultados contenidos en los Capítulos 4 y 5 son parte del siguiente pre-print,

> **Caro, S.** & Hantoute, A. *Highlighting the role of active and non-active functions in optimization and subdifferential calculus*, pre-print.

Capítulo 6 está dedicado a problemas de programación multiobjetivo semi-infinita, presentados en la forma

$$\min_{x\in S}\{f_t(x) : t\in T\},$$

donde $f_t : \mathbb{R}^n \to \mathbb{R}_\infty$, $t\in T$, es una familia de funciones a valores reales extendidos, donde T es un conjunto arbitrario y $S\subset\mathbb{R}^n$ es un conjunto convexo no vacío. Primero, nos centramos en la descripción del conjunto de soluciones débilmente eficientes del problema anterior utilizando la información proporcionada por los subproblemas asociados que involucran no más de $n+1$ funciones objetivo. Estos resultados también se han analizado en otros trabajos (por ejemplo, en Lowe et al. [44] y Ward [59]) cuando T es finito. Nuestro enfoque utiliza los argumentos desarrollados en el Capítulo 3. El resultado principal de este capítulo es Teorema 6.4 el cual establece que

$$\mathrm{WE}_T(S) = \bigcup_{B\subset T, |B|\le n+1}\mathrm{WE}_B(S\cap D).$$

En particular, obtenemos el Corolario 6.5 que recupera el resultado de Plastria & Carrizosa en [50, Corollary 2.1]. Además, proporcionamos caracterizaciones del conjunto de soluciones débilmente eficientes mediante el subdiferencial de la función objetivo asociada (supremo) y el cono normal del conjunto de restricciones S. Es decir, en el Teorema 6.6 obtenemos la siguiente equivalencia que generaliza algunos de los resultados en [50]:

$$\bar{x}\in\mathrm{WE}_T(S) \Longleftrightarrow \theta\in\mathrm{N}_S(\bar{x}) + \bigcap_{\varepsilon>0}\overline{\mathrm{co}}\left(\bigcup_{t\in T}\partial_\varepsilon f_t(\bar{x})\right).$$

Los resultados contenidos en este capítulo son parte del siguiente trabajo que está en progreso

Caro, S. & Hantoute, A. *Reduction arguments in multi-objective optimization and subdifferential calculus.*

Finalmente, en el **Capítulo 7** consideramos una generalización del problema cuadrático estándar, que viene dado por

$$\mu = \inf_{g(x)=0,\ x\in C} \frac{1}{2}x^\top Ax + 2r^\top x,$$

donde $C \subset \mathbb{R}^n$ es un cono convexo cerrado, $e \in \operatorname{int} C^*$ y $g(x) := e^\top x - 1$. El capítulo comienza con una revisión de dualidad teoría lagrangiana, enfatizando la importancia de las relaciones dadas por Flores-Bazán & Mastroeni en [31]. También estudiamos la equivalencia entre las soluciones de los problemas duales/primales y la función perturbada asociada, dada por Flores-Bazán, Jourani & Mastroeni en [30]. Nuestro resultado principal, dado en Teorema 7.6, demuestra dualidad fuerte para el problema cuadrático dado a través de una formulación explícita de la función perturbada, ψ,

$$\psi(a) = \begin{cases} (1+a)^2\mu, & \text{if } a \geq -1, \\ +\infty, & \text{if } a < -1. \end{cases}$$

De hecho, este análisis revela una convexidad oculta en el problema cuadrático dado, en vista de que el conjunto

$$\{(f(x), g(x)) : x \in C\} + \mathbb{R}_+ \times \{0\}$$

es convexo. En Teorema 7.8 presentamos una generalización de un resultado en [3, Theorem 4], mostrando una forma explícita de la función $\Theta(\lambda) = \inf_{x\in C} f(x) + \lambda(e^\top x - 1)$. Los resultados contenidos en este capítulo son parte del siguiente artículo,

F. Flores-Bazán, G. Carcamo & **S. Caro**, *Extensions of the Standard quadratic optimization problem: strong duality, optimality, hidden convexity and S-lemma*, Appl. Math. Optim. 81 (2020), no. 2, 383-408.

Chapter 2

Notation and preliminary results

In this chapter, we present the basic notions and properties that we will use throughout this manuscript.

Let X be a (real) separated locally convex space (lcs, for short), whose topological dual space, written X^*, is endowed with any compatible topology for duality pair associate to the bilinear form $(x^*, x) \in X^* \times X \mapsto \langle x^*, x \rangle := x^*(x)$. The w^*-topology, the Mackey topology, and the norm topology when X is a Banach space, are typical examples of such compatible topologies. The main feature of these topologies is that $X^{**} := (X^*)^* \equiv X$. The zero vectors in X and X^* are denoted by θ. The basis of closed, convex and balanced neighborhoods of θ, in both X and X^*, called θ-neighborhoods, is represented by $\mathcal{N}$. We use the notation $\mathbb{R}_+ = \{x \in \mathbb{R} : x \geq 0\}$, $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$ and $\mathbb{R}_\infty := \mathbb{R} \cup \{+\infty\}$, and adopt the conventions $(+\infty) + (-\infty) = (-\infty) + (+\infty) = +\infty$ and $0 \cdot (+\infty) := +\infty$.

Given a possibly infinite set T, we denote by $\mathbb{R}^T$ the locally convex product space of functions from T with values in $\mathbb{R}$. We call support of $\alpha \in \mathbb{R}^T$ to the set $\operatorname{supp} \alpha := \{t \in T : \alpha_t = \alpha(t) \neq 0\}$. Then the topological dual space of $\mathbb{R}^T$, written $\mathbb{R}^{(T)}$, is formed by the mappings with finite support. We denote $\mathbb{R}_+^{(T)} := \{\alpha : T \to \mathbb{R}_+ : |\operatorname{supp} \alpha| < +\infty\}$, where $|\cdot|$ stands for the cardinality of the set. The extended unit simplex is

$$\Delta(T) := \left\{ \lambda \in \mathbb{R}_+^{(T)} : \sum_{\lambda_t > 0} \lambda_t = 1 \right\}.$$

In particular, when T is finite, the corresponding unit simplex is compact and is simply denoted $\Delta_{|T|}$.

For the following preliminary results we rely on the books [13], [52] and [60].

2.1 Convex analysis

Definition 2.1 *We consider two sets $A,\ B \subset X$ (or X^*).*

1. *A is convex if, for all $x, y \in A$,*

$$\lambda x + (1-\lambda)y \in A, \text{ for all } \lambda \in [0,1].$$

2. *The Minkowski sum of A and B is*

$$A + B := \{a + b : \ a \in A,\ b \in B\}, \quad \text{with } A + \emptyset = \emptyset + A = \emptyset. \tag{2.1}$$

3. *The product of A by $\Lambda \subset \mathbb{R}$ is*

$$\Lambda A := \{\lambda a : \lambda \in \Lambda,\ a \in A\}, \lambda A := \{\lambda\} A, \lambda \in \mathbb{R} \text{ and } \Lambda\emptyset = \emptyset A = \emptyset.$$

4. *The convex hull of A is*

$$\operatorname{co}(A) = \left\{ \sum_{i=1}^{n} \lambda_i x_i : n \in \mathbb{N}, \lambda \in \Delta_n, (x_i) \subset Abn \right\}.$$

5. *The conical convex hull of A is*

$$\operatorname{cone}(A) = \{\lambda x : \lambda \geq 0, x \in A\} = \mathbb{R}_+ A, \text{ with } \operatorname{cone}(\emptyset) = \theta.$$

6. *The interior of A, denoted $\operatorname{int}(A)$ is the largest open set contained in C.*

7. *The closed hull of A, denoted $\operatorname{cl}(A)$ or $\overline{A}$, is the smallest closed set containing A.*

8. *Let A be a convex cone in $\mathbb{R}^n$, $n \geq 1$. We say that A is a pointed cone if $A \cap (-A) = \{0_n\}$, where 0_n is the zero vector in $\mathbb{R}^n$.*

Theorem 2.2 (Carathéodory's Theorem) *Let S be any subset of $\mathbb{R}^n$. Then $x \in \operatorname{co}(S)$ if and only if x can be expressed as a convex combination of at most $n+1$ elements of S.*

Theorem 2.3 (Helly's Theorem) *Consider a finite family of convex sets $C_i \subset \mathbb{R}^n$, $i \in T := \{1, \cdots, k\}$, $k \geq n$. If $\bigcap_{i\in S} C_i \neq \emptyset$, for all $S \subset T$ such that $|S| \leq n+1$, then we have $\bigcap_{i\in S} C_i \neq \emptyset$.*

Definition 2.4 *Let be given a nonempty set $A \subset X$ (or X^*) and $\varepsilon \geq 0$.*

1. *The negative dual cone of A is*

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

(observe that $A^- = (\overline{\operatorname{cone}}(A))^-$). We also write

$$A^{--} := (A^-)^-.$$

2. *The orthogonal subspace of* A *is*

$$A^{\perp} := (-A^{-}) \cap A^{-} = \{x^* \in X^* : \langle x^*, x\rangle = 0, \text{ for all } x \in A\}.$$

3. *The* ε*-normal cone to* A *at* x $(x \in A)$ *is*

$$\mathrm{N}_A^{\varepsilon}(x) := \{x^* \in X^* : \langle x^*, z - x\rangle \leq \varepsilon, \text{ for all } z \in A\}.$$

4. *The normal cone to* A *at* x *is* $\mathrm{N}_A(x) := \mathrm{N}_A^0(x)$.

5. *When* A *is closed and convex, its recession cone is*

$$A_{\infty} := \{y \in X : x + \lambda y \in A, \text{ for all } \lambda > 0\},$$

where x *is any point in* A.

We give some important facts about the operations defined above.

Proposition 2.5 *Let be given two nonempty sets* $A, B \subset X$.

1. *The Bipolar Theorem states that*

$$A^{--} = \overline{\mathrm{cone}}(A).$$

2. *If* $A \subset B$, *then* $B^{-} \subset A^{-}$.

3. *If* $0 \in A \cap B$, *then* $(A + B)^{-} = (A \cup B)^{-} = A^{-} \cap B^{-}$.

The following two lemmas comes from [12].

Lemma 2.6 *Consider nonempty sets* A *and* $A_1, \cdots, A_k$ *in* X, $k \geq 2$. *Then,*

$$\left[\overline{\mathrm{co}}(A \cup (\cup_{i=1,\cdots,k} A_k))\right]_{\infty} = \left[\overline{\mathrm{co}}(A \cup (A_1 + \cdots + A_k))\right]_{\infty}. \tag{2.2}$$

Lemma 2.7 *Consider a family of nonempty sets* $\{A_t,\ t \in T_1 \cup T_2\} \subset X$, *where* T_1 *and* T_2 *are disjoint nonempty sets.* *Then for every* $m > 0$ *we have*

$$\begin{aligned} \left[\overline{\mathrm{co}}\left(\cup_{t \in T_1 \cup T_2} A_t\right)\right]_{\infty} &= \left[\overline{\mathrm{co}}\left((\cup_{t \in T_1} A_t) \cup (\cup_{t \in T_2} m A_t)\right)\right]_{\infty} \\ &= \left[\overline{\mathrm{co}}\left(\cup_{t_1 \in T_1, t_2 \in T_2}\left(A_{t_1} + m A_{t_2}\right)\right)\right]_{\infty}. \end{aligned} \tag{2.3}$$

Lemma 2.8 *Let* A *be a set in* X^* *and* $x \in X$, *we have that*

$$\bigcap_{L \in \mathcal{F}(x)} \left[\overline{\mathrm{co}}(A + L^{\perp})\right]_{\infty} = \left[\overline{\mathrm{co}}(A)\right]_{\infty},$$

where $\mathcal{F}(x) = \{L \subset X : L \text{ is a finite-dimensional linear subspace such that } x \in L\}$.

Proof. Let $x \in X$ and fix any subspace $L \in \mathcal{F}(x)$. First we notice that $\overline{\text{co}}(A + L^\perp) = \text{cl}(\text{co}(A + L^\perp)) \subset \text{cl}(\text{co}(A) + L^\perp)$. Using the characterization of the recession cone given in [60, pp. 6], for every $a \in \overline{\text{co}}(A + \theta)(\subset \overline{\text{co}}(A + L^\perp)$, for every $L \in \mathcal{F}(x))$ we obtain

$$\begin{aligned}
\bigcap_{L\in\mathcal{F}(x)} \left[\overline{\text{co}}(A + L^\perp)\right]_\infty &= \bigcap_{L\in\mathcal{F}(x)} \bigcap_{t>0} t\left[\overline{\text{co}}(A + L^\perp) - a\right] \\
&= \bigcap_{L\in\mathcal{F}(x)} \bigcap_{t>0} \left[\overline{\text{co}}(tA + tL^\perp) - ta\right] \\
&= \bigcap_{t>0} \bigcap_{L\in\mathcal{F}(x)} \left[\text{cl}\left(\text{co}(tA) + L^\perp\right) - ta\right] \\
&= \bigcap_{t>0} \left(\bigcap_{L\in\mathcal{F}(x)} \left[\text{cl}\left(\text{co}(tA) + L^\perp\right)\right] - ta \right) \\
&= \bigcap_{t>0} (\overline{\text{co}}(tA) - ta) \quad \text{(by [15, Lemma 3])} \\
&= \left[\overline{\text{co}}(A)\right]_\infty .
\end{aligned}$$

□

Definition 2.9 *Let be given a function $f : X \to \overline{\mathbb{R}}$.*

1. *The effective domain and the epigraph of f are, respectively,*

$$\text{dom}\, f := \{x \in X : \ f(x) < +\infty\} \textit{ and } \text{epi}\, f := \{(x, \lambda) \in X \times \mathbb{R} : \ f(x) \le \lambda\}.$$

2. *The function f is proper if $\text{dom}\, f \neq \emptyset$ and $f(x) > -\infty$ for all $x \in X$.*

3. *The function f is convex if* epi f *is convex.*

4. *The function f is strictly convex if for all $x, y \in X$ and for all $0 < \alpha < 1$, satisfies*

$$f(\alpha x + (1 - \alpha)y) < \alpha f(x) + (1 - \alpha)f(y).$$

5. *When X is a normed space, with a norm $\|\cdot\|$. The function f is coercive if*

$$f(x) \to +\infty \textit{ as } \|x\| \to +\infty.$$

Definition 2.10 *Let $f, g : X \to \overline{\mathbb{R}}$ be two functions.*

1. *The positive part of f is the function defined as*

$$f^+(x) := \max\{0, f(x)\}, \textit{ for all } x \in X.$$

2. *The convex hull of f, $\text{co}\, f : X \to \overline{\mathbb{R}}$, is the largest among all convex functions dominated*

by f. Equivalently,

$$(\operatorname{co} f)(x) = \inf \{\mu : \ (x,\mu) \in \operatorname{co}(\operatorname{epi} f)\} \tag{2.4}$$

$$= \inf \left\{ \sum_{i=1}^{k} \lambda_i f(x_i) : \ \lambda \in \Delta_k, \ x_i \in X, \ \sum_{i=1}^{k} \lambda_i x_i = x, \ k \in \mathbb{N} \right\}. \tag{2.5}$$

3. *The Fenchel-Legendre conjugate of f is the function $f^* : X^* \to \overline{\mathbb{R}}$ is*

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

4. *The inf-convolution of f and g is the function $f \Box g : X \to \overline{\mathbb{R}}$ defined as*

$$(f \Box g)(x) = \inf\{f(x_1) + g(x_2) : x_1 + x_2 = x\}, \ \textit{for all } x \in X.$$

5. *When X is a normed space with a norm $\|\cdot\|$, the Moreau-Yosida regularization of f with parameter $\lambda > 0$ is the function f^λ defined on X by*

$$f^\lambda(x) = \left(f \,\Box\, \frac{1}{2\lambda} \|\cdot\|^2 \right)(x).$$

The following definitions presents some typical functions, which are frequently used in the sequel.

Definition 2.11 *Let A be a subset of X (or X^*).*

1. *The support function of A is the function $\sigma_A : X^* \to \overline{\mathbb{R}}$ defined by*

$$\sigma_A(x^*) := \sup\{\langle x^*, x\rangle : \ x \in A\}, \ x^* \in X^*, \ \textit{with } \sigma_\emptyset \equiv -\infty,$$

2. *The indicator function of A is*

$$\mathrm{I}_A(x) := \begin{cases} 0, & \textit{if } x \in A, \\ +\infty, & \textit{if } x \in X \setminus A. \end{cases}$$

Definition 2.12 *Given A be a subset of X. The Minkowski function (Minkowski gauge) of the set A is defined by*

$$p_A : X \to \mathbb{R}_\infty, \quad p_A(x) := \inf\{\lambda \geq 0 : x \in \lambda A\},$$

with the convention $\inf \emptyset = +\infty$.

The following gives a description of the domain of the support function.

Lemma 2.13 *Let A be a convex and closed subset of X (or X^*). Then we have*

$$A_\infty = (\operatorname{dom} \sigma_A)^- \text{ and } (A_\infty)^- = \operatorname{cl}(\operatorname{dom} \sigma_A).$$

Lemma 2.14 *Let A and B be two nonempty subsets of X. Then,*

$$([\overline{\mathrm{co}}\,(A \cup B)]_\infty)^- \subset ([\overline{\mathrm{co}}\,(A)]_\infty)^- \cap ([\overline{\mathrm{co}}\,(B)]_\infty)^- .$$

PROOF. Since $\theta \in [\overline{\mathrm{co}}\,(A)]_\infty \cap [\overline{\mathrm{co}}\,(B)]_\infty$, by Proposition 2.5(3) we obtain

$$([\overline{\mathrm{co}}\,(A)]_\infty)^- \cap ([\overline{\mathrm{co}}\,(B)]_\infty)^- = ([\overline{\mathrm{co}}\,(A)]_\infty \cup [\overline{\mathrm{co}}\,(B)]_\infty)^- .$$

Moreover, the inclusion $[\overline{\mathrm{co}}\,(A)]_\infty \cup [\overline{\mathrm{co}}\,(B)]_\infty \subset [\overline{\mathrm{co}}\,(A) \cup \overline{\mathrm{co}}\,(B)]_\infty$ is always fulfilled and Proposition 2.5(2) this implies that

$$\begin{aligned} ([\overline{\mathrm{co}}(A \cup B)]_\infty)^- &= ([\overline{\mathrm{co}}(A) \cup \overline{\mathrm{co}}(B)]_\infty)^- \\ &\subset ([\overline{\mathrm{co}}\,(A)]_\infty \cup [\overline{\mathrm{co}}\,(B)]_\infty)^- \\ &= ([\overline{\mathrm{co}}\,(A)]_\infty)^- \cap ([\overline{\mathrm{co}}\,(B)]_\infty)^- . \end{aligned}$$

□

2.2 Topology and convexity

Definition 2.15 *Let be given a function $f : X \to \overline{\mathbb{R}}$.*

1. *The function f is said to be lower semicontinuous (lsc, in brief) at $x \in X$ if*

$$f(x) = \liminf_{x' \to x} f(x') = \sup_{V \in \mathcal{N}} \inf \{f(x') : \; x' \in x + V\} .$$

   We say that f is upper semicontinuous (usc, in brief) at x if $-f$ is lsc at x. The function f is lsc (usc, respectively), if f is lsc (usc, respectively) at every point of X. We sometimes refer to lsc functions as closed functions. If it is necessary to specify the topology with respect to which the function is lsc, say a given topology τ, then we say that such a function is $\tau-$lsc.

2. *The set of proper lsc convex functions on X is denoted $\Gamma_0(X)$.*

3. *A function f is continuous if and only if f is both upper and lower semicontinuous.*

Below is a very useful property for functions in $\Gamma_0(X)$, see for example [60, Theorem 2.2.6].

Lemma 2.16 *Given $f : X \to \overline{\mathbb{R}}$ belonging to $\Gamma_0(X)$, then f is bounded from below by a continuous affine function.*

Faced with the lack of semicontinuity/lower convexity, we will consider lsc and convex hulls or envelopes.

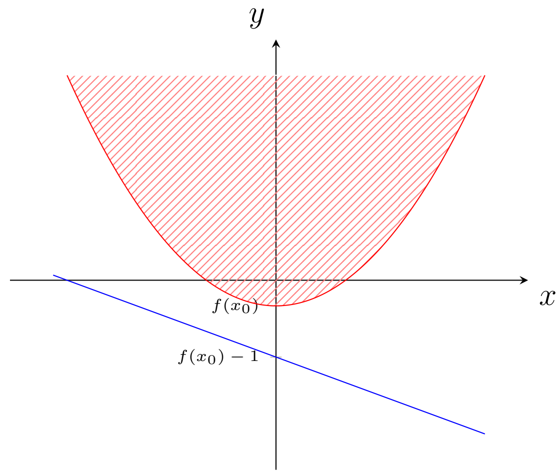


Figure 2.1: $f \in \Gamma_0(X)$ is minorized by a continuous affine mapping

Definition 2.17 *Let be given a function $f : X \to \overline{\mathbb{R}}$.*

1. *The closed hull of f is the function $\operatorname{cl} f : X \to \overline{\mathbb{R}}$ whose epigraph is $\operatorname{cl}(\operatorname{epi} f)$. Equivalently,*
$$(\operatorname{cl} f)(x) = \liminf_{x' \to x} f(x'), \quad \text{for all } x \in X. \tag{2.6}$$

2. *The closed convex hull of f, $\overline{\operatorname{co}} f : X \to \overline{\mathbb{R}}$, is the largest among all closed convex functions dominated by f. Equivalently, we have that*
$$\operatorname{epi}(\overline{\operatorname{co}} f) = \overline{\operatorname{co}}(\operatorname{epi} f).$$

Definition 2.18 *The set $\mathcal{C}(X, \mathbb{R})$ denotes the space of all continuous functions from X to $\mathbb{R}$ and given $k \in \mathbb{N}$, $\mathcal{C}^k(X, \mathbb{R})$ defines the k-times continuously differentiable functions.*

Theorem 2.19 (Bi-conjugate Theorem) *Let $f : X \to \overline{\mathbb{R}}$ be a function with a nonempty domain.*

1. *If $\overline{\operatorname{co}} f$ is proper, then $f^{**} = \overline{\operatorname{co}} f$.*

2. *If $\overline{\operatorname{co}} f$ is not proper, then $f^{**} = -\infty$.*

3. *Suppose that f is convex. If f is lsc at $\bar{x} \in \operatorname{dom} f$, then $f(\bar{x}) = f^{**}(\bar{x})$; moreover, if $f(\bar{x}) \in \mathbb{R}$, then $f^{**} = \overline{f}$ and $\overline{f}$ is proper.*

Lemma 2.20 *Let $f : X \to \mathbb{R}_\infty$ be a function.*

1. *If $A \subset X$, then*
$$\sigma_{\{0\}\cup A} = (\sigma_A)^+.$$

2. *If f is convex, then*
$$\mathrm{cl}(f^+) = (\mathrm{cl}\, f)^+.$$

3. *If f is convex but not proper, then*
$$\mathrm{cl}(f^+) = \mathrm{I}_{\mathrm{dom\,cl}\, f}.$$

PROOF. 1. Let $x^* \in X^*$. We have

$$\begin{aligned}\sigma_{\{0\}\cup A}(x^*) &= \sup\{\langle x^*, x\rangle : x \in \{0\} \cup A\} \\ &= \max\{0, \sup\{\langle x^*, x\rangle : x \in A\}\} \\ &= \max\{0, \sigma_A(x^*)\} = (\sigma_A)^+(x^*).\end{aligned}$$

2. See [13, Proposition 5.2.4(ii)].

3. If $x \in \mathrm{dom}(\mathrm{cl}\, f)$, then $(\mathrm{cl}\, f)(x) = -\infty$ and, by the previous item,
$$\mathrm{cl}(f^+)(x) = (\mathrm{cl}\, f(x))^+ = 0.$$
Otherwise, if $x \notin \mathrm{dom}(\mathrm{cl}\, f)$, then $(\mathrm{cl}\, f)(x) = +\infty$ and so, again by the previous item,
$$\mathrm{cl}(f^+)(x) = (\mathrm{cl}\, f(x))^+ = +\infty.$$

□

The following proposition gathers some important facts about the Fenchel conjugate.

Proposition 2.21 *Let $f, g : X \to \overline{\mathbb{R}}$ be two functions.*

1. *f^* is convex and w^*-lsc.*

2. *The Young-Fenchel inequality holds, that is, $f(x)+f^*(x^*) \geq \langle x, x^*\rangle$, $\forall x \in X, \forall x^* \in X^*$.*

3. *$f \leq g \Rightarrow g^* \leq f^*$.*

4. *If $\alpha > 0$, then we have $(\alpha f)^*(x^*) = \alpha f^*(\alpha^{-1}x^*)$, $\forall x^* \in X^*$.*

5. *We have*
$$f^*(\theta) = -\inf_{x\in X} f(x). \tag{2.7}$$

6. *If f is convex and lsc, then the function f^* is proper if and only if f is so.*

7. *If X is a normed space, with norm $\|\cdot\|$ and dual norm $\|\cdot\|_*$, and $f = \frac{1}{2}\|\cdot\|^2$, then*

$$f^* = \frac{1}{2}\|\cdot\|_*^2.$$

8. *The conjugate of the inf-convolution $f\Box g$ of f and g is*

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

9. *When $f, g \in \Gamma_0(X)$, the conjugate of the sum, $f + g$, is*

$$(f+g)^* = \mathrm{cl}(f^*\Box g^*).$$

   Additionally, if f is continuous at some point of $\operatorname{dom} g$*, then*

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

10. *The conjugate of the Moreau-Yosida regularization f^λ, $\lambda > 0$, is*

$$(f^\lambda)^*(x^*) = f^*(x^*) + \frac{\lambda}{2}\|x^*\|^2. \tag{2.8}$$

11. *If $f = \inf_{i\in I} f_i$ for functions $f_i : X \to \mathbb{R}_\infty$, $i \in I$, then*

$$f^* = \sup_{i\in I} f_i^*.$$

12. *If $f = \sup_{i\in I} f_i$ for proper lsc convex functions $f_i : X \to \mathbb{R}_\infty$, $i \in I$, and f is proper, then*

$$f^* = \overline{\mathrm{co}}\left(\inf_{i\in I} f_i^*\right). \tag{2.9}$$

Proof. Items $1-4$ and 8 can be found in [60, Theorem 2.3.1].

5. By definition, we have $f^*(\theta) = \sup\{\langle\theta, x\rangle - f(x) : x \in X\} = -\inf_{x\in X} f(x)$.

6. See [13, Proposition 3.1.4].

7. We start by proving the formula when X is a reflexive Banach space. Fix $x^* \in X^*$. If $x^* = \theta$, then the desired statement obviously holds. Otherwise, $x^* \neq \theta$ and James' theorem yields some $x \in X \setminus \{\theta\}$ such that $\langle x^*, x\rangle = \|x^*\|\|x\|$. Then the point $z := x\|x^*\|/\|x\|$ satisfies

$$\langle x^*, z\rangle = \langle x^*, x\|x^*\|/\|x\|\rangle = \|x^*\|^2, \quad \|z\| = \|x^*\|,$$

   and we deduce

$$\left(\frac{1}{2}\|\cdot\|^2\right)^*(x^*) \geq \langle x^*, z\rangle - \frac{1}{2}\|z\|^2 = \|x^*\|^2 - \frac{1}{2}\|z\|^2 = \frac{1}{2}\|x^*\|^2.$$

Finally, the opposite inequality follows as,

$$\left(\frac{1}{2}\|\cdot\|^2\right)^*(x^*) \leq \frac{1}{2}\|x^*\|_*^2 + \frac{1}{2}\|z\|^2 - \frac{1}{2}\|z\|^2 = \frac{1}{2}\|x^*\|_*^2.$$

Next, in the general case, given $x^* \in X^*$, we fix $\varepsilon > 0$ and $x_\varepsilon \in B_X$ (the closed unit ball in X) such that

$$\left|\frac{1}{2}\|x^*\|_*^2 - \frac{1}{2}\langle x^*, x_\varepsilon\rangle^2\right| \leq \varepsilon.$$

Take a finite-dimensional subspace L of X such that $x_\varepsilon \in L$, and denote by $(f + \mathrm{I}_L)_L$ and x_L^* the restriction of $f + \mathrm{I}_L$ and x^*, respectively, to L. Then we easily verify that

$$((f + \mathrm{I}_L)_L)^*(x_L^*) = (f + \mathrm{I}_L)^*(x^*).$$

But $(f+\mathrm{I}_L)_L$ is nothing else but the function $\frac{1}{2}\|x\|_L^2$, $x \in L$, with $\|\cdot\|_L$ being the norm in L, which is the restriction of the initial norm $\|\cdot\|$ to L. Thus, since L is reflexive, by the reasoning above we deduce that

$$((f + \mathrm{I}_L)_L)^*(x_L^*) = \frac{1}{2}\|x_L^*\|_{L^*}^2.$$

Therefore, observing that $f \leq f + \mathrm{I}_L$, we have $f^*(x^*) \geq (f + \mathrm{I}_L)^*(x^*)$ by item 3., and we obtain

$$\begin{aligned} f^*(x^*) &\geq (f + \mathrm{I}_L)^*(x^*) = (1/2)\|x_L^*\|_{L^*}^2 \geq \frac{1}{2}\langle x_L^*, x_\varepsilon\rangle^2 = (1/2)\langle x^*, x_\varepsilon\rangle^2 \\ &\geq (1/2)\|x^*\|_*^2 - \varepsilon. \end{aligned}$$

Consequently, as $\varepsilon \downarrow 0$, we get the first inequality $f^*(x^*) \geq (1/2)\|x^*\|_*^2$. To show the opposite inequality we use the Young and Cauchy-Schwartz inequalities,

$$\begin{aligned} f^*(x^*) &= \sup_{x\in X}\{\langle x^*, x\rangle - (1/2)\|x\|^2\} \\ &\leq \sup_{x\in X}\{(1/2)\|x\|^2 + (1/2)\|x^*\|_*^2 - (1/2)\|x\|^2\} \\ &= (1/2)\|x^*\|_*^2. \end{aligned}$$

9. We apply the previous item to the functions f^* and g^* to obtain

$$f^{**} + g^{**} = (f^* \Box g^*)^*$$

Since $f, g \in \Gamma_0(X)$, $f^{**} = f, g^{**} = g$ and $f \,\Box\, g$ is convex. Then, applying the conjugate to the last equality, we get

$$(f + g)^* = (f^* \,\Box\, g^*)^{**} = \mathrm{cl}(f^* \,\Box\, g^*).$$

If we assume that f is continuous at $x_0 \in \operatorname{dom} g$, then the conclusion holds due to [60, Corollary 2.3.5].

10. Since $f^\lambda(x) = \left(f \,\square\, \frac{1}{2\lambda}\|\cdot\|^2\right)(x)$, by items $6-7$ we have for all $x^* \in X^*$

$$\begin{aligned}
(f^\lambda)^*(x^*) &= f^*(x^*) + \left(\frac{1}{2\lambda}\|\cdot\|^2\right)^*(x^*) \\
&= f^*(x^*) + \left(\frac{1}{\lambda}\left(\frac{1}{2}\|\cdot\|^2\right)\right)^*(x^*) \\
&= f^*(x^*) + \frac{1}{\lambda}\cdot\frac{1}{2}\|\lambda x^*\|_*^2 \\
&= f^*(x^*) + \frac{\lambda}{2}\|x^*\|_*^2.
\end{aligned}$$

11. For all $x^* \in X^*$ we have

$$\begin{aligned}
\left(\inf_{i\in I} f_i\right)^*(x^*) &= \sup_{x\in X}\left\{\langle x^*, x\rangle - \inf_{i\in I} f_i(x)\right\} \\
&= \sup_{x\in X}\sup_{i\in I}\{\langle x^*, x\rangle - f_i(x)\} = \sup_{i\in I} f_i^*(x^*).
\end{aligned}$$

12. Using the bi-conjugate theorem, Theorem 2.19, together with item 9, for all $x \in X$ we have

$$f(x) = \sup_{i\in I} f_i(x) = \sup_{i\in I} f_i^{**}(x) = \left(\inf_{i\in I} f_i^*\right)^*(x).$$

Thus, taking the conjugate of each side, for all $x^* \in X^*$ we obtain that

$$f^*(x^*) = \left(\inf_{i\in I} f_i^*\right)^{**}(x^*). \tag{2.10}$$

Now, since f is assumed proper, its conjugate f^* is proper too, and we deduce that the function $(\inf_{i\in I} f_i^*)^{**}$ is also proper. Thus, due to the inequalities

$$\left(\inf_{i\in I} f_i^*\right)^{**} \le \inf_{i\in I} f_i^* \le f_i^*, \quad \text{for all } i \in I,$$

the function $\overline{\text{co}}(\inf_{i\in I} f_i^*)$ is proper. Finally, thanks to Theorem 2.19, the relation in (2.10) yields

$$f^*(x^*) = \overline{\text{co}}(\inf_{i\in I} f_i^*).$$

□

Proposition 2.22 *Given $f : \mathbb{R}^n \to \mathbb{R}_\infty$ a proper and convex function and $x \in \mathbb{R}^n$. The following properties hold*

1. *For all $\lambda > 0$, $\inf_{x\in\mathbb{R}^n} f(x) = \inf_{x\in\mathbb{R}^n} f^\lambda(x)$.*

2. *For all $\lambda > 0$ the function f^λ is $\mathcal{C}^1$.*

3. As $\lambda \to 0$, f^λ *converges pointwise to* $\mathrm{cl}(f)$*.*

4. $\sup_{\lambda>0} f^\lambda = \mathrm{cl}\, f$

Proof. Items 1-3 can be found in [48], page 52, Proposition 3.39 and Proposition 3.41, respectively. To prove item 4 we use the previous item. For all $\lambda > 0$, the function f^λ is lsc, so $\sup_{\lambda>0} f^\lambda$ is lsc too. Moreover, for all $x \in \mathbb{R}^n$

$$f^\lambda(x) \leq f(x) \Rightarrow \sup_{\lambda>0} f^\lambda(x) \leq \mathrm{cl}\, f(x) \leq f(x)$$

and together with item 2, we obtain the desired conclusion. □

Lemma 2.23 *Let* $\{A_\varepsilon,\ \varepsilon > 0\}$ *be a family of nonempty closed convex sets in* X*.*

1. $\sigma_{\cup_{\varepsilon>0} A_\varepsilon} = \sup_{\varepsilon>0} \sigma_{A_\varepsilon}$*,*

2. $\sigma_{\cap_{\varepsilon>0} A_\varepsilon} = \overline{\mathrm{co}}\,(\inf_{\varepsilon>0} \sigma_{A_\varepsilon})$*, provided that* $\cap_{\varepsilon>0} A_\varepsilon \neq \emptyset$*.*

In particular, if the family of functions $\{\sigma_{A_\varepsilon}, \varepsilon > 0\}$ *is non-increasing, then the last equality above becomes*

$$\sigma_{\cap_{\varepsilon>0} A_\varepsilon} = \mathrm{cl}(\inf_{\varepsilon>0} \sigma_{A_\varepsilon}).$$

Proof. 1. This easily follows from the definition of the support function.

2. Fix $x^* \in X^*$. We have

$$\sigma_{\cap_{\varepsilon>0} A_\varepsilon}(x^*) = (\mathrm{I}_{\cap_{\varepsilon>0} A_\varepsilon})^*(x^*) = (\sup_{\varepsilon>0} \mathrm{I}_{A_\varepsilon})^*(x^*) = \overline{\mathrm{co}}(\inf_{\varepsilon>0} (\mathrm{I}_{A_\varepsilon})^*)(x^*),$$

where the last equality comes from Proposition 2.21(11.). Finally, we are done because the infimum of a non-increasing family of convex functions is convex.

□

Below, we present the classical Minimax Theorem (see, for instance, [60, Theorem 2.10.2]).

Theorem 2.24 (Minimax Theorem) *Let* X *be an lcs,* Y *be a linear space,* $A \subset X$ *be a nonempty convex compact set and* $B \subset Y$ *be a nonempty convex set. Let also* $F : A \times B \to \mathbb{R}$ *be a function with the property that* $F(\cdot, y)$ *is concave and usc, for every* $y \in B$*, and* $F(x, \cdot)$ *is convex for every* $x \in A$*. Then*

$$\max_{x\in A} \inf_{y\in B} F(x, y) = \inf_{y\in B} \max_{x\in A} F(x, y).$$

If moreover Y *is an lcs,* B *is compact and* $F(x, \cdot)$ *is lsc for every* $x \in A$*, then*

$$\max_{x\in A} \min_{y\in B} F(x, y) = \min_{y\in B} \max_{x\in A} F(x, y);$$

in particular F *has saddle points.*

The following result is a consequence of the previous theorem.

Proposition 2.25 *Given convex functions $f_1, \cdots, f_m : \mathbb{R}^n \to \mathbb{R}_\infty$, $m, n \geq 1$, we have that*

$$\inf_{\mathbb{R}^n} \max\{f_1, \cdots, f_m\} = \max_{\alpha \in \Delta_m} \inf_{\mathbb{R}^n} \left\{\sum_{1 \leq i \leq m} \alpha_i f_i\right\}.$$

Proof. We start by checking the following equality,

$$\inf_{\mathbb{R}^n} f = \inf_{\mathbb{R}^n} \max_{\alpha \in \Delta_m} \left\{\sum_{1 \leq i \leq m} \alpha_i f_i\right\},$$

where $f := \max\{f_1, \cdots, f_m\}$. Since $f \geq f_i$, for all $i = 1, \cdots, m$, we obtain the inequality $\inf_{\mathbb{R}^n} f \geq \inf_{\mathbb{R}^n} \max_{\alpha \in \Delta_m} \{\sum_{1 \leq i \leq m} \alpha_i f_i\}$. Conversely, we take $\lambda \in \mathbb{R}$ such that

$$\inf_{\mathbb{R}^n} \max_{\alpha \in \Delta_m} \left\{ \sum_{1 \leq i \leq m} \alpha_i f_i \right\} < \lambda.$$

Then there exists $x_\lambda \in \mathbb{R}^n$ that satisfies the following inequality

$$\sum_{1 \leq i \leq m} \alpha_i f_i(x_\lambda) \leq \lambda, \text{ for all } \alpha \in \Delta_m.$$

In particular, $f_i(x_\lambda) \leq \lambda$ for all $i = 1, \cdots, m$. Therefore $f(x_\lambda) \leq \lambda$ and consequently, $\inf_{x \in \mathbb{R}^n} f(x) \leq \lambda$.

Let us also observe that if f takes the value $-\infty$ somewhere, then $\inf_{\mathbb{R}^n} f = -\infty$ and, so, for all $\alpha \in \Delta_m$,

$$\inf_{\mathbb{R}^n} \left\{ \sum_{1 \leq i \leq m} \alpha_i f_i \right\} \leq \inf_{\mathbb{R}^n} \left\{ \sum_{1 \leq i \leq m} \alpha_i f \right\} = \inf_{\mathbb{R}^n} f = -\infty;$$

that is, the conclusion holds. The same occurs if $\inf_{\mathbb{R}^n} f = +\infty$. Indeed, in such a case we can find some $\alpha \in \Delta_m$ such that $\inf_{\mathbb{R}^n} \{\sum_{1 \leq i \leq m} \alpha_i f_i\} = +\infty$; otherwise, we would find some $x \in \mathbb{R}^n$ such that $\sum_{1 \leq i \leq m} \frac{1}{m} f_i(x) < +\infty$, implying the contradiction $\inf_{\mathbb{R}^n} f \leq f(x) < +\infty$.

Next, according to the discussion above, we assume that f is proper and consider the function $F : \Delta_m \times \operatorname{dom} f \subset \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$, defined as

$$(\alpha, x) \longmapsto F(\alpha, x) := \sum_{1 \leq i \leq m} \alpha_i f_i(x).$$

Then we see that $F(\cdot, x)$ is concave and usc for every $x \in \operatorname{dom} f$, and $F(\alpha, \cdot)$ is convex for every $\alpha \in \Delta_m$. Thus, since Δ_m is a compact set in $\mathbb{R}^m$ and $\operatorname{dom} f$ is nonempty and convex, the minimax theorem (Theorem 2.24)) ensures the equality, that is,

$$\inf_{\mathbb{R}^n} f = \inf_{x \in \operatorname{dom} f} \max_{\alpha \in \Delta_m} F(\alpha, x) = \max_{\alpha \in \Delta_m} \inf_{x \in \operatorname{dom} f} F(\alpha, x). \tag{2.11}$$

Moreover, since for all $x \in \operatorname{dom} f$ and $\alpha \in \Delta_m$,

$$F(\alpha, x) \leq f(x) \leq \max_{\beta \in \Delta_m} F(\beta, x),$$

we have that $\max_{\alpha \in \Delta_m} F(\alpha, x) = f(x)$ and, so, (2.11) reads

$$\max_{\alpha \in \Delta_m} \inf_{x \in \operatorname{dom} f} F(\alpha, x) = \inf_{x \in \operatorname{dom} f} f(x) = \inf_{x \in X} f(x).$$

□

2.3 Subdifferential analysis

We introduce here the notions of directional derivatives and subdifferentials for convex functions defined on an lcs X.

Definition 2.26 *We give a function $f : X \to \overline{\mathbb{R}}$ and $x, u \in X$.*

1. *The directional derivative of f at x in the direction u is the function $f'(x;\cdot) : X \to \overline{\mathbb{R}}$ defined by*
$$f'(x;u) := \lim_{t\downarrow 0} \frac{f(x+tu)-f(x)}{t}.$$

2. *The ε-directional derivative of f at x in the direction u is the function $f'_\varepsilon(x;\cdot) : X \to \overline{\mathbb{R}}$ defined by*
$$f'_\varepsilon(x;u) := \inf_{t>0} \frac{f(x+tu)-f(x)+\varepsilon}{t}.$$

We give an example in which we relate the ε-directional derivative of the indicator of a given set in terms of the gauge function.

Lemma 2.27 *Let $C \subset X$ be a nonempty convex set. Then, for all $x \in C$, $u \in X$ and $\varepsilon \geq 0$ we have*
$$(\mathrm{I}_C)'_\varepsilon(x;u) = \varepsilon P_{(C-x)}(u)$$
and, in particular,
$$(\mathrm{I}_C)'(x;u) = \mathrm{I}_{\mathbb{R}_+(C-x)}(u).$$

Proof. Since $x \in C$, we have

$$\begin{aligned}
(\mathrm{I}_C)'_\varepsilon(x;u) &= \inf_{t>0} \frac{\mathrm{I}_C(x+tu)+\varepsilon}{t} \\
&= \inf\left\{\frac{\varepsilon}{t} : x+tu \in C,\ t>0\right\} \\
&= \varepsilon \inf\{t>0 : x + \frac{1}{t}u \in C\} \\
&= \varepsilon \inf\{t>0 : u \in t(C-x)\} \\
&= \varepsilon P_{(C-x)}(u).
\end{aligned}$$

In particular, when $\varepsilon = 0$, by the definition of the gauge function we obtain

$$(\mathrm{I}_C)'(x;u) = 0 \cdot P_{(C-x)}(u) = \mathrm{I}_{\mathbb{R}_+(C-x)}(u).$$

□

Definition 2.28 *Let* $f : X \to \overline{\mathbb{R}}$ *be a function,* $\varepsilon \in \mathbb{R}$ *and* $x \in X$*.*

1. *The* ε*-subdifferential (or the approximate subdifferential) of* f *at* x *is*
$$\partial_\varepsilon f(x) = \{x^* \in X^* : \ f(y) \geq f(x) + \langle x^*, y - x\rangle - \varepsilon \ \text{ for all } y \in X\}, \tag{2.12}$$
when $x \in \operatorname{dom} f$*, and* $\partial_\varepsilon f(x) := \emptyset$ *when* $f(x) \notin \mathbb{R}$ *or* $\varepsilon < 0$*.*

2. *The subdifferential of* f *at* x *is* $\partial f(x) := \partial_0 f(x)$*. Equivalently,* $\partial f(x) = \bigcap_{\varepsilon>0} \partial_\varepsilon f(x)$*.*

The results below specify the concepts above to convex functions (see, e.g., [13] or [60]).

Theorem 2.29 *Let* $f : X \to \mathbb{R}_\infty$ *be a proper convex function and take* $x \in \operatorname{dom} f$*. Then, for every* $u \in X$*,* $f'(x;u)$ *exists in* $\overline{\mathbb{R}}$ *and we have that*
$$f'(x;u) = \inf_{t>0} \frac{f(x+tu) - f(x)}{t}.$$

Theorem 2.30 *Let* $f : X \to \mathbb{R}_\infty$ *be a proper convex function,* $x \in \operatorname{dom} f$ *and* $\varepsilon \in \mathbb{R}_+$*. Then the function* $f'_\varepsilon(x;\cdot)$ *is sublinear. Moreover, we have that*

1. $\operatorname{dom} f'_\varepsilon(x;\cdot) = \operatorname{cone}(\operatorname{dom} f - x)$*.*

2. $f'_\varepsilon(x;u) \leq f(x+u) - f(x) + \varepsilon$*, for all* $u \in X$*.*

3. $f'(x;u) = \lim_{\delta\downarrow 0} f'_\delta(x;u) = \inf_{\delta>0} f'_\delta(x;u)$*, for all* $u \in X$*.*

4. *If* f *is continuous at* x*, then* $f'_\varepsilon(x;u) \in \mathbb{R}$*, for every* $u \in X$*.*

Proof. See Theorems 2.1.14 and 2.4.9 in [60]. □

Lemma 2.31 *Let* $f : X \to \overline{\mathbb{R}}$ *be a convex function which is not proper. Then, for all* $x \in \operatorname{dom} f$*,* $u \in X$ *and* $\varepsilon \geq 0$ *we have*
$$(\operatorname{cl}(f^+))'_\varepsilon(x;u) = \varepsilon P_{(\operatorname{dom}(\operatorname{cl} f) - x)}(u).$$
In particular, we have
$$(\operatorname{cl}(f^+))'(x;u) = \mathrm{I}_{\mathbb{R}_+(\operatorname{dom}(\operatorname{cl} f) - x)}(u).$$

Proof. Since f is convex and not proper, from Lemma 2.20(3.) we have for all $z \in X$
$$\operatorname{cl}(f^+)(z) = \mathrm{I}_{\operatorname{dom}(\operatorname{cl} f)}(z).$$
Then, for all $x \in \operatorname{dom} f$ and $\varepsilon \geq 0$,
$$\begin{aligned}((\operatorname{cl} f)^+)'_\varepsilon(x;u) &= \inf_{t>0} \frac{(\operatorname{cl} f)^+(x+tu) - (\operatorname{cl} f)^+(x) + \varepsilon}{t} \\ &= \inf_{t>0} \frac{\mathrm{I}_{\operatorname{dom}(\operatorname{cl} f)}(x+tu) - \mathrm{I}_{\operatorname{dom}(\operatorname{cl} f)}(x) + \varepsilon}{t},\end{aligned}$$
and so we are done thanks to Lemma 2.27. □

Some properties of the subdifferential and ε-subdifferential come next.

Theorem 2.32 *The following statements hold for every function $f : X \to \overline{\mathbb{R}}$ and every $\bar{x} \in X$ such that $f(\bar{x}) \in \mathbb{R}$.*

1. *$\partial f(\bar{x}) \subset X^*$ is a convex and w^*-closed (possibly empty) set.*

2. *If $\partial f(\bar{x}) \neq \emptyset$, then*
$$(\overline{\mathrm{co}} f)(\bar{x}) = \overline{f}(\bar{x}) = f(\bar{x}) \text{ and } \partial(\overline{\mathrm{co}} f)(\bar{x}) = \partial \overline{f}(\bar{x}) = \partial f(\bar{x});$$
in particular, f is proper and $f^{**}(\bar{x}) = \overline{\mathrm{co}} f(\bar{x})$; hence, f is lsc at $\bar{x}$.

3. *$\partial f(\bar{x}) \neq \emptyset \Leftrightarrow f(\bar{x}) = \max_{x^* \in X^*}(\langle \bar{x}, x^* \rangle - f^*(x^*))$.*

4. *We have, for every $\varepsilon \geq 0$,*
$$\partial_\varepsilon f(\bar{x}) = \{x^* \in X^* : \ f(\bar{x}) + f^*(x^*) \leq \langle x^*, \bar{x} \rangle + \varepsilon\}. \tag{2.13}$$

5. *For $\lambda > 0$ $\partial_\varepsilon(\lambda f)(\bar{x}) = \lambda \partial_{\varepsilon/\lambda} f(\bar{x})$.*

6. *Given $\varepsilon_1, \varepsilon_2 \geq 0$ and $g : X \to \overline{\mathbb{R}}$ such that $g(\bar{x}) \in \mathbb{R}$, then*
$$\partial_{\varepsilon_1} f(\bar{x}) + \partial_{\varepsilon_2} g(\bar{x}) \subset \partial_{\varepsilon_1+\varepsilon_2}(f+g)(\bar{x}).$$

7. *If X is a normed space and ε, $\delta \geq 0$, then $\partial_\varepsilon \left(\frac{\delta}{2}\|\cdot\|^2\right)(0) = \sqrt{2\varepsilon\delta} B_{X^*}$, where B_{X^*} the closed unit ball in X^*.*

8. *Given C be any subset of X, $\varepsilon \geq 0$ and $x \in C$, then*
$$\partial_\varepsilon \mathrm{I}_C(x) = \mathrm{N}_C^\varepsilon(x).$$

Proof. For items $1-3$ see [60, Theorem 2.4.1], and [60, Theorem 2.4.2] for item $4-5$. while the inclusion of item 6 is direct from the definition of ε-subdifferential. Moreover, for item 8 see [13], page 98. To prove item 7 we use the definition of the ε-subdifferential:

$$\begin{aligned}
\partial_\varepsilon \left(\frac{\delta}{2}\|\cdot\|^2\right)(0) &= \left\{x^* \in X^* : \left(\frac{\delta}{2}\|\cdot\|^2\right)^*(x^*) + \left(\frac{\delta}{2}\|\cdot\|^2\right)(0) \leq \langle x^*, 0\rangle + \varepsilon\right\} \\
&= \left\{x^* \in X^* : \delta\left(\frac{1}{2}\|\cdot\|^2\right)^*\left(\frac{1}{\delta}x^*\right) \leq \varepsilon\right\} \text{ (by Proposition 2.21(4.))} \\
&= \left\{x^* \in X^* : \frac{1}{2\delta}\|x^*\|^2 \leq \varepsilon\right\} \text{(by Proposition 2.21(6.))} \\
&= \left\{x^* \in X^* : \|x^*\| \leq \sqrt{2\delta\varepsilon}\right\} = \sqrt{2\varepsilon\delta} B_{X^*}.
\end{aligned}$$

□

The following two results give the Moreau-Rockafellar and the Hiriart-Urruty & Phelps theorems, respectively. See, for example, [13, Proposition 4.1.16] and [60, Corollary 2.6.7], respectively.

Theorem 2.33 (Moreau-Rockafellar Theorem) *Let $f, g : X \to \mathbb{R}_\infty$ be two functions in $\Gamma_0(X)$ such that f is finite and continuous at some point in $\operatorname{dom} g$. Then, for every $x \in X$, we have*

$$\partial f(x) + \partial g(x) = \partial (f+g)(x).$$

Proposition 2.34 *Let $f, g \in \Gamma_0(X)$. If $x \in \operatorname{dom} f \cap \operatorname{dom} g$ and $\varepsilon \geq 0$, then we have the following rules.*

1. *If $\varepsilon > 0$, then* $\partial_\varepsilon (f+g)(x) = \operatorname{cl}\left(\bigcup_{\substack{\varepsilon_1, \varepsilon_2 \geq 0 \\ \varepsilon_1 + \varepsilon_2 = \varepsilon}} (\partial_{\varepsilon_1} f(x) + \partial_{\varepsilon_2} g(x)) \right)$.
2. $\partial (f+g)(x) = \bigcap_{\varepsilon > 0} \operatorname{cl}(\partial_\varepsilon f(x) + \partial_\varepsilon g(x))$.

Lemma 2.35 *Assume that $\{f_t,\ t \in T\} \subset \Gamma_0(X)$. Given two families of parameters $(\varepsilon_t)_{t\in T},\ (\alpha_t)_{t\in T} \subset\ ]0,1]$, we have*

$$\left[\overline{\text{co}} \left(\bigcup_{\substack{t\in T \\ 0\leq \alpha_t \leq 2\varepsilon_t}} \partial_\varepsilon (\alpha_t f_t)(x) \right) \right]_\infty \subset \left[\overline{\text{co}} \left(\bigcup_{t \in T} \partial_\varepsilon (\varepsilon_t f_t)(x) \right) \right]_\infty .$$

Proof. We have that

$$\begin{aligned}
\left[\overline{\text{co}} \left(\bigcup_{\substack{t\in T \\ 0\leq \alpha_t \leq 2\varepsilon_t}} \partial_\varepsilon (\alpha_t f_t)(x) \right) \right]_\infty &= \left[\overline{\text{co}} \left(\bigcup_{\substack{t\in T \\ 0\leq \alpha_t \leq 2\varepsilon_t}} \partial_\varepsilon \left(\frac{\alpha_t}{2\varepsilon_t} 2\varepsilon_t f_t \right)(x) \right) \right]_\infty \\
&\subset \left[\overline{\text{co}} \left(\bigcup_{\substack{t\in T \\ 0\leq \alpha_t \leq 2\varepsilon_t}} \left(\partial_\varepsilon \left(\frac{\alpha_t}{2\varepsilon_t} 2\varepsilon_t f_t \right)(x) \bigcup \partial_\varepsilon \left((1 - \frac{\alpha_t}{2\varepsilon_t}) 2\varepsilon_t f_t \right)(x) \right) \right) \right]_\infty \\
&= \left[\overline{\text{co}} \left(\bigcup_{\substack{t\in T \\ 0\leq \alpha_t \leq 2\varepsilon_t}} \left(\partial_\varepsilon \left(\frac{\alpha_t}{2\varepsilon_t} 2\varepsilon_t f_t \right)(x) + \partial_\varepsilon \left((1 - \frac{\alpha_t}{2\varepsilon_t}) 2\varepsilon_t f_t \right)(x) \right) \right) \right]_\infty \quad (2.14) \\
&\subset \left[\overline{\text{co}} \left(\bigcup_{\substack{t\in T \\ 0\leq \alpha_t \leq 2\varepsilon_t}} \partial_{2\varepsilon} \left(\frac{\alpha_t}{2\varepsilon_t} 2\varepsilon_t f_t + (1 - \frac{\alpha_t}{2\varepsilon_t}) 2\varepsilon_t f_t \right)(x) \right) \right]_\infty \quad (2.15) \\
&\subset \left[\overline{\text{co}} \left(\bigcup_{\substack{t\in T \\ 0\leq \alpha_t \leq 2\varepsilon_t}} \partial_{2\varepsilon} \left(2\varepsilon_t f_t \right)(x) \right) \right]_\infty = \left[\overline{\text{co}} \left(\bigcup_{t\in T} 2\partial_\varepsilon (\varepsilon_t f_t)(x) \right) \right]_\infty \\
&= \left[\overline{\text{co}} \left(\bigcup_{t\in T} \partial_\varepsilon (\varepsilon_t f_t)(x) \right) \right]_\infty .
\end{aligned}$$

The equality in (2.14) above comes from Lemma 2.6, while the inclusion in (2.15) is a simple version of Proposition 2.34. □

2.4 Quadratic forms

Definition 2.36 (Pointed cone) *Let P be a convex cone in $\mathbb{R}^n$, $n \geq 1$. We say that P is a pointed cone if $P \cap (-P) = \{0_n\}$, where 0_n is the zero vector in $\mathbb{R}^n$.*

Definition 2.37 *Let $A \in \mathcal{M}_{n\times n}(\mathbb{R})$, $n \geq 1$ be a matrix, and let $P \subseteq \mathbb{R}^n$ be a cone.*

1. *The matrix A is said to be positive semi-definite, written $A \succeq 0$, if $x^\top Ax \geq 0$ for all $x \in \mathbb{R}^n$.*

2. *The matrix A is said to be positive definite, written $A \succ 0$, if $x^\top Ax > 0$ for all $x \in \mathbb{R}^n \backslash \{0\}$.*

3. *The matrix A is said to be copositive in P if $x^\top Ax \geq 0$, for all $x \in P$.*

4. *The matrix A is said to be strictly copositive in P if $x^\top Ax > 0$, for all $x \in P\backslash\{0\}$.*

Let $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}^m$, $n, m > 1$ be two functions. We consider the minimization problem (called *primal problem*)

$$\begin{aligned} \mu \doteq \inf \quad & f(x) \\ \text{s.t} \quad & g(x) \in -P \\ & x \in C, \end{aligned} \tag{2.16}$$

where P is a convex cone in $\mathbb{R}^m$ and C is a convex subset of $\mathbb{R}^n$. This problem (2.16) is known as the standard quadratic problem when $f(x) = \frac{1}{2}x^\top Ax$ is a homogeneous quadratic function with $A \in S_n(\mathbb{R})$, $g(x) = \mathbb{1}^\top x - 1$, $\mathbb{1}^\top = (1, \cdots, 1) \in \mathbb{R}^n$, $P = \{0_m\}$ and $C = \mathbb{R}^n_+$. Then the associated feasible set becomes $K := \{x \in C : g(x) = 0\} = \Delta_n$, the canonical simplex in the Euclidean n-dimensional space, and our problem is written

$$\mu \doteq \min_{x\in\Delta_n} f(x). \tag{2.17}$$

As observed in [2, Chapter 1], the minimizers of problem (2.17) remain the same if we replace A by the new matrix $\tilde{A} = A + \gamma\mathbb{1} \cdot \mathbb{1}^\top$, where γ is an arbitrary constant. Indeed, if $\bar{x}$ is a minimizer of (2.17), then for every feasible point x of (2.17) we have

$$\begin{aligned} \bar{x}^\top \tilde{A}\bar{x} &= \bar{x}^\top A\bar{x} + \gamma\bar{x}^\top \left(\mathbb{1} \cdot \mathbb{1}^\top\right) \bar{x} = \bar{x}^\top A\bar{x} + \gamma \underbrace{(\bar{x}^\top \mathbb{1})}_{=1} \underbrace{(\mathbb{1}^\top \bar{x})}_{=1} \\ &= \bar{x}^\top A\bar{x} + \gamma \leq x^\top Ax + \gamma = x^\top Ax + \gamma \underbrace{(x^\top \mathbb{1})}_{=1} \underbrace{(\mathbb{1}^\top x)}_{=1} \\ &= x^\top Ax + \gamma x^\top \left(\mathbb{1} \cdot \mathbb{1}^\top\right) x = x^\top \left(A + \gamma\mathbb{1} \cdot \mathbb{1}^\top\right) x \\ &= x^\top \tilde{A}x. \end{aligned}$$

The main advantage of this change is that the new matrix $\tilde{A}$ has non-negative entries. Also, according to the same author ([2, Chapter 1]), if the set K of constraints of (2.17) is maintained but the objective function is modified to a non-necessarily homogeneous quadratic form, say $f(x) = x^\top Ax + 2r^\top x$, $r \in \mathbb{R}^n$ then one can "homogenize" by considering the matrix $\hat{A} = A + \mathbb{1} \cdot r^\top + r \cdot \mathbb{1}^\top$. Then we verify that the (non-homogeneous) problem (2.17) and its homogenized version have the same optimal value. In fact, for any element $x \in \triangle_n$, we would have that

$$\begin{aligned} x^\top \hat{A} x &= x^\top A\, x + x^\top (\mathbb{1} \cdot r^\top) x + x^\top (r \cdot \mathbb{1}^\top) x \\ &= x^\top A\, x + \underbrace{x^\top \mathbb{1}}_{=1} \cdot r^\top x + x^\top r \cdot \underbrace{\mathbb{1}^\top x}_{=1} \\ &= x^\top A\, x + 2 \cdot r^\top x. \end{aligned}$$

Chapter 3

Discretization and reduction of Infsup and SIP optimization problems

3.1 Introduction

In this chapter we work with extended real-valued convex functions defined on $\mathbb{R}^n$, $f_0, f_t : \mathbb{R}^n \to \overline{\mathbb{R}}$, $t \in T$, where the index set T is arbitrary (possibly infinite). To the space $\mathbb{R}^n$ we associate the Euclidean norm, $\|\cdot\|$, unless another option is explicitly indicated. We consider the semi-infinite programming problem (SIP, in brief)

$$\inf_{f_t(x)\leq 0,\ t\in T} f_0(x), \tag{3.1}$$

and the associated ordinary (finite) programming subproblems

$$\inf_{f_t(x)\leq 0,\ t\in S} f_0(x), \tag{3.2}$$

given for finite subsets $S \subset T$. We will answer the following questions: Can we reduce the number of constraints involved within problem (3.1)?; that is, do problems (3.1) and (3.2) have the same optimal value for some appropriately chosen sets S?

Some advantages of this type of results come from a numerical point of view, as we transform an infinite problem into a finite one. This could reduce the computational implementation time, which is an important issue in every optimization problem. More precisely, the reductions analysis would allow, for example, for a considerable simplification of the study of duality theory for problem (3.1). Moreover, constraints qualification conditions would be required to hold only for finite blocks of constraints.

Now, we consider Infsup optimization problems given in the following form

$$\inf_{x\in\mathbb{R}^n} \sup\{f_t(x), t \in T\} \tag{3.3}$$

and

$$\inf_{x\in\mathbb{R}^n} \max\{f_t(x), t \in S\}, \tag{3.4}$$

where $S \subset T$ and $|S| < +\infty$. We wonder if it is possible to reduce the number of functions involved in the objective function given as a supremum in (3.3). We will also show that there is a relationship between SIP and Infsup problems.

In 1969, Levin presented a reduction result in SIP problems as (3.1) [42, Theorem 1], under the hypotheses that T compact and the mappings $t \mapsto f_t(x)$, $x \in X$ are upper semi-continuous, together with some convexity and continuity assumptions on the functions f_0 and f_t, $t \in T$ and a Slater-type condition. The reduction in [42, Proposition 1] is done for semi-infinite optimization problems allowing the use of Helly's Theorem. For the Infsup problem as (3.3), where all the involved functions are convex and the index set T is finite, Drezner [24, Theorem 1] gives a short and direct proof using Helly's Theorem to reduce the problem (3.3) into a problem of type (3.4), providing an alternative proof to the one given for ordinary semi-infinite optimization in [42]. In the same direction, Shapiro [55, Section 3] proves a reduction for SIP problems but under the Slater condition.

It is also worth recalling that Rockafellar [53] and Borwein [6] also established a reduction processes for Infsup convex problems with constraints given as

$$\inf_{f_j(x) \leq 0,\ j \in J} f_0(x), \tag{3.5}$$

and where $f_0 := \sup\{f_i : i \in I\}$, and both I and J are arbitrary index sets. In this case, additionally, the authors used the so-called asymptotically regular conditions. We observe that the proofs in [53] are not based on the Helly theorem, which is indeed deduced as a corollary.

The main contribution of the chapter is to establish a process for reducing the number of functions involved in Infsup problems like (3.3). To this aim we use a fairly general condition that is based on the closures of the f_t's. condition $\mathrm{cl}(\sup_{t \in T} f_t) = \sup_{t \in T} \mathrm{cl}(f_t)$, and work in a compact-continuous setting. This result is a generalization of [42, Theorem 1]. Additionally, for SIP problems as (3.1) we also apply the mentioned reduction process but with the additional "blocks" Slater condition. Our result generalizes [24, Theorem 1], and does not use Helly's theorem.

Although the analysis for SIP and Infsup problems is done separately, it is important to notice that we can transform an Infsup problem from $\mathbb{R}^n$ into a SIP problem in $\mathbb{R}^{n+1}$, as

$$\inf_{x \in \mathbb{R}^n} \sup_{t \in T} f_t(x) = \inf_{f_t(x) - \gamma \leq 0,\ t \in T} \gamma,$$

which is an SIP problem whose objective function and constraints are $\tilde{f}_0(x, \gamma) := \gamma$ and $\tilde{f}_t(x, \gamma) := f_t(x) - \gamma, t \in T$, respectively. There are different ways to transform a SIP into a Infsup problem, for instance, using the Lagrangian reformulation

$$\inf_{f_t(x) \leq 0,\ t \in T} f_0(x) = \inf_{x \in \mathbb{R}^n} \sup_{\alpha \in \Delta(T)} (f_0(x) + \sum_{t \in T} \alpha_t f_t(x)).$$

However, when the optimal value of (3.1), μ, belongs to $\mathbb{R}$ we can show that

$$\inf_{x \in \mathbb{R}^n} \sup\{f_0(x) - \mu, f_t(x), t \in T\} = 0,$$

which in turn yields a reformulation of (3.1) that explicitly uses the data functions, f_0 and $f_t, t \in T$. The results presented here can be applied in minimax theorem, Chebyshev's norms, location problems [11] among others.

The structure of the chapter is the following: in section 3.2 we study the reduction process in Infsup problems (see Theorem 3.1), while in section 3.3 we provide reduction schemes for SIP problems (Theorem 3.8 and Lemma 3.6). At the same time, the relationship between SIP and Infsup problems is clarified. In section 3.4 we present some extensions of the previous results; for example, Theorem 3.10 provides an Alternative-type theorem that allow us to relate the ε-minima of the original problem with the ε-minima of the reduced problem. It also gives information about the Lagrange multipliers of the reduced problem and, as consequence, we obtain zero duality gaps under usual qualification conditions.

3.2 Reduction of Infsup problems

This section is devoted to reduce Infsup problems in $\mathbb{R}^n$, given by

$$\mu := \inf_{\mathbb{R}^n} \sup\{f_t,\ t \in T\}, \tag{3.6}$$

where $f_t : \mathbb{R}^n \to \mathbb{R}_\infty$, $t \in T$, are given proper convex functions. Throughout the section we consider the associated supremum function

$$f := \sup_{t\in T} f_t,$$

and assume the compact-continuous setting, that is, T is Hausdorff compact and the index mappings $t \mapsto f_t(x)$ are usc, for all x in a given convex set D such that

$$\operatorname{dom} f \subset D \subset \mathbb{R}^n. \tag{3.7}$$

In this compact-continuous setting we know that (see, e.g., [33])

$$f := \sup_{t\in T} f_t = \max_{t\in T} f_t \text{ and } \operatorname{dom} f = \textstyle\bigcap_{t\in T} \operatorname{dom} f_t.$$

We give the first result of this section allowing for a reduction of problem (3.6) to smaller Infmax problems. A variant of this result is given in Theorem 3.2 in which we remove the condition on the lsc hulls of the functions involved.

Theorem 3.1 *Consider problem (3.6), where T is Hausdorff compact and the index mappings $t \mapsto f_t(x)$ are usc, for all x in a set $D \subset X$ satisfying (3.7). If $\operatorname{cl} f = \sup_{t\in T}(\operatorname{cl} f_t)$, then*

$$\inf_{\mathbb{R}^n} f = \max_{S\subset T,\ |S|\le n+1} \inf_{\mathbb{R}^n} \max\{f_t + \mathrm{I}_D,\ t \in S\}$$

and, consequently,

$$\inf_{\mathbb{R}^n} f = \max_{\alpha\in\Delta(T),\ |\operatorname{supp}\alpha|\le n+1} \inf_{\mathbb{R}^n} \left\{ \sum_{t\in T} \alpha_t f_t + \mathrm{I}_D \right\}.$$

Proof. By reasoning as in the beginning of the proof of Proposition 2.25, we may suppose that $f := \sup_{t\in T} f_t$ $(= \max_{t\in T} f_t)$ is proper and $\inf_{\mathbb{R}^n} f > -\infty$. We proceed by proving the first statement step by step.

Step 1. We assume here that f is additionally continuous and attains its infimum at some point $\bar{x} \in X$, so that $D = \operatorname{dom} f = X$. Then, according to Charatheodory's Theorem, there exists some index set

$$I \subset T(\bar{x}) := \{t \in T : f_t(\bar{x}) = f(\bar{x})\},$$

such that $|I| \leq n+1$ and, using [16, Corollary 10],

$$0_n \in \partial f(\bar{x}) = \operatorname{co}\left\{\bigcup_{t\in I} \partial f_t(\bar{x})\right\}.$$

Moreover, because all the f_t's are also continuous, we have $\partial f_t(\bar{x}) \neq \emptyset$, for all $t \in I$. So, we may assume that $|I| = n+1$. Then there are $x_t^* \in \partial f_t(\bar{x})$ and $\alpha \in \Delta_{n+1}$ such that

$$0_n = \sum_{t\in I} \alpha_t x_t^* \in \sum_{t\in I} \alpha_t \partial f_t(\bar{x}) \subset \partial\left(\sum_{t\in I} \alpha_t f_t\right)(\bar{x}) \subset \partial\left(\max_{t\in I} f_t\right)(\bar{x}).$$

Then the convex functions $\sum_{t\in I} \alpha_t f_t$ and $\max_{t\in I} f_t$ attain their infimum at the point $\bar{x}$ and we verify that $\sum_{t\in I} \alpha_t f_t(\bar{x}) = (\max_{t\in I} f_t)(\bar{x}) = f(\bar{x})$. Thus

$$\inf_{\mathbb{R}^n} f = f(\bar{x}) = \sum_{t\in I} \alpha_t f_t(\bar{x}) \leq \max_{\alpha\in\Delta(T),\, |\operatorname{supp}\alpha|\leq n+1} \inf_{\mathbb{R}^n} \sum_{t\in I} \alpha_t f_t \leq \inf_{\mathbb{R}^n}\left(\max_{t\in I} f_t\right) \leq \inf_{\mathbb{R}^n} f,$$

and the conclusion follows in the present case.

Step 2. In this step we suppose that f is continuous but, possibly, it does not attain its infimum. We fix $\delta,\ \varepsilon > 0$ and pick $x_\varepsilon \in \operatorname{dom} f$ such that $0_n \in \partial_\varepsilon f(x_\varepsilon)$ (as $\inf_X f > -\infty$), in other words, x_ε is an ε-minimizer of f. Then we introduce the functions

$$\tilde{f}_t := f_t + \frac{\delta}{2}\left\|\cdot - x_\varepsilon\right\|^2, \quad t \in T,$$

together with the continuous convex function

$$\tilde{f} := \max_{t\in T} \tilde{f}_t = f + \frac{\delta}{2}\left\|\cdot - x_\varepsilon\right\|^2.$$

Since f is minorized by a continuous affine mapping (as $f \in \Gamma_0(\mathbb{R}^n)$), the function $\tilde{f}$ is coercive. So, because of the finite-dimensional setting, f attains its infimum at some point $\bar{x}_\delta \in \mathbb{R}^n$. Therefore, based on step 1, we find some $I_\delta \subset T$ and $\alpha_\delta \in \Delta(T)$ such that $|\operatorname{supp}\alpha_\delta| \leq n+1$,

$$\inf_{\mathbb{R}^n} \tilde{f} = \inf_{\mathbb{R}^n}\left(\max_{t\in I_\delta} \tilde{f}_t\right) = \inf_{\mathbb{R}^n}\left(\frac{\delta}{2}\left\|\cdot - x_\varepsilon\right\|^2 + \max_{t\in I_\delta} f_t\right),$$

and

$$f_t(\bar{x}_\delta) + \frac{\delta}{2}\left\|\bar{x}_\delta - x_\varepsilon\right\|^2 = f(\bar{x}_\delta) + \frac{\delta}{2}\left\|\bar{x}_\delta - x_\varepsilon\right\|^2 \leq f(x_\varepsilon) \leq f(\bar{x}_\delta) + \varepsilon, \text{ for all } t \in I_\delta,$$

that is,

$$f_t(\bar{x}_\delta) = f(\bar{x}_\delta) \geq f(x_\varepsilon) - \varepsilon \geq f_t(x_\varepsilon) - \varepsilon, \text{ for all } t \in I_\delta. \tag{3.8}$$

At the same time, we have

$$\begin{aligned}
\max_{t\in I_\delta} f_t(\bar{x}_\delta) &\leq \left(\frac{\delta}{2}\left\|\cdot - x_\varepsilon\right\|^2 + \max_{t\in I_\delta} f_t\right)(\bar{x}_\delta) \\
&\leq \left(\frac{\delta}{2}\left\|\cdot - x_\varepsilon\right\|^2 + f\right)(\bar{x}_\delta) = \inf_{\mathbb{R}^n} \tilde{f} \\
&= \inf_{\mathbb{R}^n}\left(\frac{\delta}{2}\left\|\cdot - x_\varepsilon\right\|^2 + \max_{t\in I_\delta} f_t\right) \\
&\leq \max_{t\in I_\delta} f_t(x_\varepsilon),
\end{aligned}$$

and (3.8) yields

$$-\max_{t\in I_\delta} f_t(x_\varepsilon) \leq -\max_{t\in I_\delta} f_t(\bar{x}_\delta) = -f(\bar{x}_\delta) \leq -f(x_\varepsilon) + \varepsilon. \tag{3.9}$$

Moreover, we get

$$\begin{aligned}
\left(\frac{\delta}{2}\left\|\cdot - x_\varepsilon\right\|^2 + \max_{t\in I_\delta} f_t\right)(x_\varepsilon) &= \max_{t\in I_\delta} f_t(x_\varepsilon) \leq f(x_\varepsilon) \leq \inf_{\mathbb{R}^n} f + \varepsilon \\
&\leq \inf_{\mathbb{R}^n} \tilde{f} + \varepsilon = \inf_{\mathbb{R}^n}\left(\frac{\delta}{2}\left\|\cdot - x_\varepsilon\right\|^2 + \max_{t\in I_\delta} f_t\right) + \varepsilon,
\end{aligned}$$

and by using [60, Theorem 2.8.3] we obtain

$$\begin{aligned}
0_n &\in \partial_\varepsilon\left(\frac{\delta}{2}\left\|\cdot - x_\varepsilon\right\|^2 + \max_{t\in I_\delta} f_t\right)(x_\varepsilon) \\
&\subset \partial_\varepsilon\left(\frac{\delta}{2}\left\|\cdot - x_\varepsilon\right\|^2\right)(x_\varepsilon) + \partial_\varepsilon\left(\max_{t\in I_\delta} f_t\right)(x_\varepsilon).
\end{aligned}$$

Next, by Theorem 2.32, there exists $x^*_{\delta,\varepsilon} \in \sqrt{2\varepsilon\delta}B_{X^*}$ such that

$$x^*_{\delta,\varepsilon} \in \partial_\varepsilon\left(\max_{t\in I_\delta} f_t\right)(x_\varepsilon),$$

and so, taking into account (3.9), for all $y \in D$ we have

$$\begin{aligned}
\left\langle x^*_{\delta,\varepsilon}, y - x_\varepsilon\right\rangle &\leq \max_{t\in I_\delta} f_t(y) - \max_{t\in I_\delta} f_t(x_\varepsilon) + \varepsilon \\
&\leq \max_{t\in I_\delta} f_t(y) - f(x_\varepsilon) + 2\varepsilon.
\end{aligned}$$

We may suppose, by taking the limit as $\delta \downarrow 0$, that $I_\delta \to I_\varepsilon$ and $\alpha_\delta \to \alpha_\varepsilon \in \Delta(T)$ such that $|\operatorname{supp}\alpha_\varepsilon| \leq n+1$. So, as $\delta \downarrow 0$, $x^*_{\delta,\varepsilon} \to 0_n$ and for all $y \in D$

$$0 \leq \max_{t\in I_\varepsilon} f_t(y) - f(x_\varepsilon) + 2\varepsilon,$$

showing that

$$\inf_{\mathbb{R}^n} f \leq \max_{t\in I_\varepsilon} f_t(y) + 2\varepsilon.$$

We also show that $I_\varepsilon \to I$ and $\alpha_\varepsilon \to \alpha \in \Delta(T)$ such that $|\operatorname{supp}\alpha| \leq n+1$. So, by taking the limit as $\varepsilon \downarrow 0$ in the last inequality, we obtain

$$\inf_{\mathbb{R}^n} f \leq \max_{t\in I} f_t(y), \text{ for all } y \in D,$$

that is, $\inf_{\mathbb{R}^n} f = \inf_{\mathbb{R}^n} (\max_{t\in I} f_t + \mathrm{I}_D)$.

Step 3. In this last step, we extend the analysis of step 2 to non-continuous functions. Given $\lambda,\ \delta > 0$, we consider the Moreau-Yosida regularizations of the functions $f_{t,\ \delta} := f_t + \mathrm{I}_D + \frac{\delta}{2}\|\cdot\|^2$

$$f_{t,\delta}^\lambda := f_{t,\delta} \,\square \frac{\|\cdot\|^2}{4\lambda}.$$

Hence, $f_{t,\delta}^\lambda$ is convex and continuous (even $\mathcal{C}^1$). Moreover, the mappings $t \mapsto f_{t,\delta}^\lambda(x)$ are usc for all $x \in X$. Indeed, given $x \in X$, we have that

$$\begin{aligned}\limsup_{s\to t} f_{s,\delta}^\lambda(x) &= \limsup_{s\to t} \inf_{y\in D}\left(f_s(y) + \frac{\delta}{2}\|y\|^2 + \frac{\|x-y\|^2}{4\lambda}\right)\\ &\leq \inf_{y\in D}\left(\limsup_{s\to t} f_s(y) + \frac{\delta}{2}\|y\|^2 + \frac{\|x-y\|^2}{4\lambda}\right)\\ &\leq \inf_{y\in D}\left(f_t(y) + \frac{\delta}{2}\|y\|^2 + \frac{\|x-y\|^2}{4\lambda}\right) = f_{t,\delta}^\lambda(x),\end{aligned}$$

where the last inequality uses the current upper semi-continuity of the mappings $t \mapsto f_t(x)$, $x \in D$. Consequently, $\sup_{t\in T} f_{t,\delta}^\lambda = \max_{t\in T} f_{t,\delta}^\lambda$ and this function is continuous. Thus, step 2 (with $D = X$) yields some $I_\lambda \subset T$ such that $|I_\lambda| \leq n+1$ and

$$\begin{aligned}\inf_{\mathbb{R}^n}\max_{t\in T} f_{t,\delta}^\lambda = \inf_{\mathbb{R}^n}\left\{\max_{t\in I_\lambda} f_{t,\delta}^\lambda\right\} &= \max_{I\subset T,\ |I|\leq n+1}\ \inf_{\mathbb{R}^n}\left\{\max_{t\in I} f_{t,\delta}^\lambda\right\}\\ &\leq \max_{I\subset T,\ |I|\leq n+1}\ \inf_{\mathbb{R}^n}\left\{\max_{t\in I} f_{t,\delta} + \mathrm{I}_D\right\}.\end{aligned}$$

So, taking the supremum over $\lambda > 0$ (equivalently, the limit as $\lambda \downarrow 0$), we find some $I \subset T$ such that $|I| \leq n+1$ and

$$\sup_{\lambda>0}\inf_{\mathbb{R}^n}\max_{t\in T} f_{t,\delta}^\lambda \leq \inf_{\mathbb{R}^n}\max_{t\in I}\left(f_{t,\delta} + \mathrm{I}_D\right). \tag{3.10}$$

At the same time, on the one hand, by (2.7) and (2.9) we have that

$$\begin{aligned}\sup_{\lambda>0}\inf_{\mathbb{R}^n}\max_{t\in T} f_{t,\delta}^\lambda &= -\inf_{\lambda>0}\left(\max_{t\in T} f_{t,\delta}^\lambda\right)^*(0_n)\\ &= -\inf_{\lambda>0}\overline{\mathrm{co}}\left(\inf_{t\in T}(f_{t,\delta}^\lambda)^*\right)(0_n)\\ &= -\inf_{\lambda>0}\overline{\mathrm{co}}\left(\inf_{t\in T}(f_{t,\delta})^* + \lambda\|\cdot\|^2\right)(0_n),\end{aligned}$$

where the last equality comes from (2.8). More specifically, since the mappings $t \mapsto (f_{t,\delta})^*(x^*)$, $x^* \in \mathbb{R}^n$, are lsc, we also have that

$$\sup_{\lambda>0} \inf_{\mathbb{R}^n} \max_{t\in T} f_{t,\delta}^\lambda = -\inf_{\lambda>0} \overline{\mathrm{co}} \left(\min_{t\in T} (f_{t,\delta})^* + \lambda \left\|\cdot\right\|^2 \right) (0_n).$$

On the other hand, we have

$$\begin{aligned}\left(\inf_{\lambda>0} \left(\max_{t\in T} f_{t,\delta}^\lambda \right)^* \right)^* (x) &= \sup_{x^*\in\mathbb{R}^n} \sup_{\lambda>0} \left\{ \langle x^*, x\rangle - \left(\max_{t\in T} f_{t,\delta}^\lambda \right)^* (x^*) \right\} \\ &= \sup_{\lambda>0} \sup_{x^*\in\mathbb{R}^n} \left\{ \langle x^*, x\rangle - \left(\max_{t\in T} f_{t,\delta}^\lambda \right)^* (x^*) \right\} \\ &= \sup_{\lambda>0} \left(\max_{t\in T} f_{t,\delta}^\lambda \right)^{**} (x),\end{aligned}$$

and, since the function $x \mapsto \left(\max_{t\in T} f_{t,\delta}^\lambda\right)(x)$ is proper lsc and convex, the biconjugate theorem (Theorem 2.19) entails

$$\left(\inf_{\lambda>0} \left(\max_{t\in T} f_{t,\delta}^\lambda \right)^* \right)^* = \sup_{\lambda>0} \max_{t\in T} f_{t,\delta}^\lambda = \sup_{t\in T} \sup_{\lambda>0} f_{t,\delta}^\lambda = \sup_{t\in T} (\mathrm{cl}\, f_{t,\delta}), \tag{3.11}$$

where $\mathrm{cl}\, f_{t,\delta}$ is the closed hull of $f_{t,\delta}$ with respect to the x-variable. Moreover, the function $\inf_{\lambda>0} \overline{\mathrm{co}} \left(\min_{t\in T} f_{t,\delta}^* + \lambda \left\|\cdot\right\|^2\right)$ is proper convex and satisfies, for all $t \in T$,

$$\inf_{\lambda>0} \overline{\mathrm{co}} \left(\min_{t\in T} f_{t,\delta}^* + \lambda \left\|\cdot\right\|^2 \right) \leq \inf_{\lambda>0} \overline{\mathrm{co}} \left(f_{t,\delta}^* + \lambda \left\|\cdot\right\|^2 \right) = \inf_{\lambda>0} (f_{t,\delta}^* + \lambda \left\|\cdot\right\|^2) = f_{t,\delta}^*.$$

Thus, since each $f_{t,\delta}^* = \mathrm{cl} \left(f_t^* \Box \frac{\|\cdot\|^2}{2\delta} \right) = f_t^* \Box \frac{\|\cdot\|^2}{2\delta}$ is continuous, the function

$$\inf_{\lambda>0} \overline{\mathrm{co}} \left(\min_{t\in T} f_{t,\delta}^* + \lambda \left\|\cdot\right\|^2 \right)$$

is also continuous, by [60, Theorem 2.2.9]. Therefore, taking the conjugates in (3.11) we deduce

$$\inf_{\lambda>0} \overline{\mathrm{co}} \left(\min_{t\in T} f_{t,\delta}^* + \lambda \left\|\cdot\right\|^2 \right) (0_n) = \overline{\mathrm{co}} \left(\min_{t\in T} f_{t,\delta}^* \right) (0_n),$$

and we derive that

$$\sup_{\lambda>0} \inf_{\mathbb{R}^n} \max_{t\in T} f_{t,\delta}^\lambda = -\overline{\mathrm{co}}(\min_{t\in T} f_{t,\delta}^*)(0_n) = -(\sup_{t\in T} (\mathrm{cl}\, f_{t,\delta}))^*(0_n) = \inf_{\mathbb{R}^n} \sup_{t\in T} (\mathrm{cl}\, f_{t,\delta}).$$

Hence, (3.10) yields

$$\inf_{\mathbb{R}^n} \sup_{t\in T} (\mathrm{cl}\, f_t) \leq \inf_{\mathbb{R}^n} \sup_{t\in T} (\mathrm{cl}\, f_{t,\delta}) \leq \max_{I\subset T, |I|\leq n+1} \inf_{\mathbb{R}^n} \max_{t\in I} \left(f_{t,\delta} + \mathrm{I}_D \right),$$

and we obtain, as $\delta \downarrow 0$,

$$\inf_{\mathbb{R}^n} \sup_{t\in T} (\mathrm{cl}\, f_t) \leq \max_{I\subset T, |I|\leq n+1} \inf_{\mathbb{R}^n} \max_{t\in I} \left(f_t + \mathrm{I}_D \right).$$

Finally, by the current assumption we have

$$\begin{aligned}\inf_{\mathbb{R}^n} f &= \inf_{\mathbb{R}^n} (\operatorname{cl} f) = \inf_{\mathbb{R}^n} \sup_{t\in T} (\operatorname{cl} f_t) \leq \max_{I\subset T, |I|\leq n+1} \inf_{\mathbb{R}^n} \max_{t\in I} (f_t + \mathrm{I}_D) \\ &\leq \inf_{\mathbb{R}^n} \max_{t\in T} (f_t + \mathrm{I}_D) = \inf_D f = \inf_{\mathbb{R}^n} f,\end{aligned}$$

and we are done with the first conclusion of the theorem. For the last conclusion, we apply Proposition 2.25 and get

$$\inf_{\mathbb{R}^n} \max_{t\in T} (f_t + \mathrm{I}_D) = \max_{\alpha\in\Delta(T), |\operatorname{supp}\alpha|\leq n+1} \inf_{\mathbb{R}^n} \left\{ \sum_{t\in\operatorname{supp}\alpha} \alpha_t f_t + \mathrm{I}_D \right\}.$$

□

The condition on the closures in Theorem 3.1, $\operatorname{cl} f = \sup_{t\in T}(\operatorname{cl} f_t)$, is obviously satisfied when the f_t's are lsc. This is also the case when the supremum function f is continuous somewhere or if all of the f_t's have the same effective domain (see [33] and [35]). However, there are convex proper functions that do not satisfy this closure condition. For instance, if we consider the functions $f, g : \mathbb{R} \to \mathbb{R}_\infty$ defined by

$$f(x) = \begin{cases} \dfrac{1}{1-x}, & \text{for } x < 0, \\ +\infty, & \text{for } x \geq 0 \end{cases} \quad \text{and } g(x) = \begin{cases} +\infty, & \text{for } x \leq 0, \\ \dfrac{1}{1+x}, & \text{for } x > 0. \end{cases}$$

At the same time, we have that $\overline{\operatorname{epi} f \cap \operatorname{epi} g} = \emptyset$ and $\overline{\operatorname{epi} f} \cap \overline{\operatorname{epi} g} = \{(0,t) : t \geq 1\}$, that is, $\operatorname{cl}(\max\{f,g\}) \neq \max\{\operatorname{cl} f, \operatorname{cl} g\}$.

Nevertheless, the following result rewrites Theorem 3.1 by avoiding the condition on the closures.

Corollary 3.2 *Consider problem (3.6), where T is Hausdorff compact and the index mappings $t \mapsto f_t(x)$ are usc, for all $x \in \operatorname{dom} f$. Then we have*

$$\inf_{\mathbb{R}^n} f = \max_{S\subset T,\ |S|\leq n+1} \inf_{\operatorname{dom} f} \max\{f_t,\ t\in S\} = \max_{\alpha\in\Delta(T),\ |\operatorname{supp}\alpha|\leq n+1} \inf_{\operatorname{dom} f} \sum_{t\in S} \alpha_t f_t.$$

Proof. Let us define the convex functions

$$\tilde{f}_t := f_t + \mathrm{I}_{\operatorname{dom} f},\ t\in T,\ \text{and } \tilde{f} := \sup_{t\in T} \tilde{f}_t.$$

Then, by [33] (see the proof of Proposition 6), we have that

$$\operatorname{cl} \tilde{f} = \sup_{t\in T} \left(\operatorname{cl} \tilde{f}_t \right),$$

and therefore, by applying Theorem 3.1 with $D = \operatorname{dom} f$, we obtain

$$\inf_{\mathbb{R}^n} f = \inf_{\mathbb{R}^n} \tilde{f} = \max_{S\subset T, |S|\leq n+1} \inf_{\operatorname{dom} f} \max_{t\in S} f_t = \max_{\alpha\in\Delta(T),\ |\operatorname{supp}\alpha|\leq n+1} \inf_{\operatorname{dom} f} \sum_{t\in T} \alpha_t f_t.$$

□

When T is a finite set, we have the following known result (see [24, Theorem 1]), which can be deduced from the general case of Corollary 3.2 (or directly from the Helly theorem; see the paragraph after Proposition 3.5 in the next section).

Corollary 3.3 *Consider a finite family of convex functions $f_1, \cdots, f_m : \mathbb{R}^n \to \mathbb{R}_\infty$, $m > n-1$, and denote $f := \max_{1\le t\le m} f_t$. Then we have*

$$\inf_{\mathbb{R}^n} f = \max_{S\subset\{1,\cdots,m\},\ |S|\le n+1} \inf_{\mathbb{R}^n} \max\{f_t : t \in S\} = \max_{\alpha\in\Delta_m,\ |\operatorname{supp}\alpha|\le n+1} \inf_{\mathbb{R}^n} \left\{ \sum_{t\in\operatorname{supp}\alpha} \alpha_t f_t \right\}.$$

Proof. We are obviously in the compact-continuous setting, and it suffices to apply Theorem 3.1 with $D = \mathbb{R}^n$. □

Additionally, for the sake of completeness, we give a direct proof of the last theorem, which is based on Helly's theorem (this has also been used in [24]). This approach cannot be directly extended to infinite families. We give a family of convex functions $f_1, \cdots, f_m : \mathbb{R}^n \to \mathbb{R}_\infty$, $m > n-1$, and $f := \max_{1\le t\le m} f_t$ such that $\inf_{\mathbb{R}^n} f < +\infty$. We also may suppose that $\inf_{\mathbb{R}^n} f > -\infty$; for otherwise, we easily conclude the desired relation. Next, we consider the nonempty convex sets

$$B := \left\{(x,\gamma) \in \mathbb{R}^{n+1} : \gamma < \inf_{\mathbb{R}^n} f\right\}$$

and

$$C_t := \operatorname{epi} f_t,\ t \in \{1, \cdots, m\},$$

so that $B \cap (\cap_{1\le t\le m} C_t) = \emptyset$. Since $\cap_{1\le t\le m} C_t \ne \emptyset$, as a consequence of the condition $\inf_{\mathbb{R}^n} f < +\infty$, by applying Helly's theorem in $\mathbb{R}^{n+1}$ we find $t_1, \cdots, t_{n+1} \in T$ such that

$$B \cap \left(\bigcap_{1\le i\le n+1} C_{t_i}\right) = \emptyset.$$

Therefore, given any $x \in X$, $\delta > 0$, there exists some $j \in \{1, \cdots, n+1\}$ such that $(x, \inf_{\mathbb{R}^n} f - \delta) \notin C_{t_j}$, that is,

$$\inf_{\mathbb{R}^n} f - \delta < f_{t_j}(x) \le \max_{1\le i\le n+1} f_{t_i}(x).$$

Thus, we get the desired conclusion when $\delta \downarrow 0$

$$\inf_{\mathbb{R}^n} f \le \inf_{\mathbb{R}^n} \max_{1\le i\le n+1} f_{t_i}.$$

The terms I_D and $\mathrm{I}_{\operatorname{dom} f}$ used in the previous results are essential in our analysis, and could not be removed in general as we show in the following example.

Example 3.4 *Consider the family of lsc proper convex functions $f_m, f_\infty : \mathbb{R} \to \mathbb{R}_\infty$, $m \ge 1$, given by (see Figure 3.1)*

$$f_m(x) := \begin{cases} 2 - \frac{1}{m}, & \text{if } -1 \le mx \le 1, \\ +\infty, & \text{if not,} \end{cases} \quad \text{and } f_\infty(x) := \begin{cases} +\infty, & x < 0, \\ -\frac{2}{3}x + 2, & x \in [0,3], \\ 0, & x > 3. \end{cases}$$

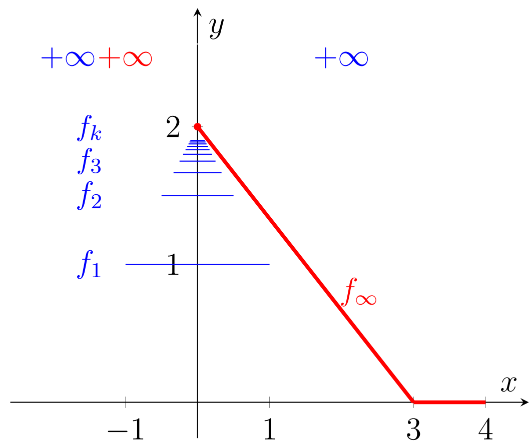


Figure 3.1: Graphs of $f_m,\ m \in \mathbb{N},\ f_\infty$

The index set here is the Hausdorff compact set $T = \mathbb{N} \cup \{+\infty\}$, *and we have that*

$$f(x) := \sup_{t \in T} f_t(x) = \mathrm{I}_{\{0\}}(x) + 2,$$

so that $\operatorname{dom} f = \{0\}$ *and*

$$\limsup_{m \to +\infty} f_m(0) = f_\infty(0) = 2,$$

that is, the conditions of Corollary 3.2 hold. Then we verify that $\inf_{\mathbb{R}} f = 2$ *and we would like to know whether there would exist* $t_1, t_2 \in T$ *such that* $\inf_{\mathbb{R}} f = \inf_{\mathbb{R}} \max\{f_{t_1}, f_{t_2}\}$. *Observe that*

$$\begin{aligned}\inf_{\mathbb{R}} f_m &= 2 - \frac{1}{m} < 2, \text{ for all } m \geq 1, \ \inf_{\mathbb{R}} f_\infty = 0,\\ \inf_{\mathbb{R}} \max\{f_{m_1}, f_{m_2}\} &= 2 - \frac{1}{\max\{m_1, m_2\}} < 2, \text{ for all } m_1, m_2 \geq 1.\end{aligned}$$

Also, since $\max\{f_m(x), f_\infty(x)\} = -\frac{2}{3}x + 2 + \mathrm{I}_{[0, \frac{1}{m}]}(x)$, *we get*

$$\inf_{\mathbb{R}} \max\{f_m, f_\infty\} = 2 - \frac{2}{3m} < 2.$$

However, we have that

$$\inf_{\mathbb{R}} \max\{f_\infty + \mathrm{I}_{\operatorname{dom} f}\} = f_\infty(0) = 2 = \inf_{\mathbb{R}} f,$$

and Corollary 3.2 is not true if we remove the term $\mathrm{I}_{\operatorname{dom} f}$ *from the reduction statement.*

3.3 Reduction of SIP problems

We provide in this section a reduction approach to SIP programming problems. We start by the case of ordinary convex optimization problems. The following result, already classical ([42, Proposition 1]), follows from Helly's Theorem (Theorem 2.3). We include the proof for completeness.

Proposition 3.5 *Given convex functions $f_0, f_1, \cdots, f_m : \mathbb{R}^n \to \mathbb{R}_\infty$, $m \geq 1$, we have*

$$\inf_{f_i \leq 0,\ 1 \leq i \leq m} f_0 = \inf_{f_i \leq 0,\ 1 \leq i \leq \min\{m,n\}} f_0,$$

provided the problem $\inf_{f_i \leq 0,\ 1 \leq i \leq m} f_0$ is feasible.

Proof. Because only the inequality "$\leq$" needs to be proved, we suppose that

$$\alpha := \inf_{f_i \leq 0,\ 1 \leq i \leq m} f_0 > -\infty.$$

Thus, since

$$[f_0 < \alpha] \bigcap \left(\bigcap_{1 \leq i \leq m} [f_i \leq 0] \right) = \emptyset \text{ and } \bigcap_{1 \leq i \leq m} [f_i \leq 0] \neq \emptyset,$$

Helly's Theorem yields some $S \subset T$, $|S| \leq n$ (and not $n+1$ as in Helly's theorem statement because of the feasibility assumption), such that

$$[f_0 < \alpha] \bigcap \left(\bigcap_{i \in S} [f_i \leq 0] \right) = \emptyset.$$

Hence, $\inf_{f_i \leq 0,\ i \in S} f_0 \geq \alpha$, and the desired inequality follows by the arbitrariness of α. □

It is worth observing that Proposition 3.5 also allows for a direct proof of Corollary 3.3. In fact, as in [42], it suffices to write

$$\inf_{\mathbb{R}^n} \max\{f_1, \cdots, f_m\} = \inf_{f_i(x) \leq \gamma,\ i \in \{1, \cdots, m\}} \gamma,$$

and, then, use the reduction approach in Proposition 3.5.

The following key lemma transforms semi-infinite optimization problems into different minimax problems.

Lemma 3.6 *The following statements hold true for every arbitrary family of convex functions $f_0,\ f_t : \mathbb{R}^n \to \mathbb{R}_\infty$, $t \in T$ $(0 \notin T)$.*

1. $\inf_{f_t \leq 0,\ t \in T} f_0 = \inf_{\mathbb{R}^n} \sup_{\alpha \in \mathbb{R}_+^{(T)}} (f_0 + \sum_{t \in T} \alpha_t f_t)$.

2. $\inf_{f_t \leq 0,\ t \in T} f_0 = \inf_{\mathbb{R}^n} \sup\{f_0;\ \mathrm{I}_{[f_t \leq 0]},\ t \in T\}$.

3. *If $\mu := \inf_{f_t \leq 0,\ t \in T} f_0 \in \mathbb{R}$, then*

$$\inf_{\mathbb{R}^n} \sup\{f_0 - \mu;\ f_t,\ t \in T\} = 0. \tag{3.12}$$

Proof. Assertions 1. and 2. are easily checked. To prove assertion 3., we assume that $\mu \in \mathbb{R}$. Then

$$\inf_{x \in C} \sup\{f_0(x) - \mu;\ f_t(x),\ t \in T\} = \inf_{x \in C} (f_0(x) - \mu) = 0, \tag{3.13}$$

where $C := [\sup_{t \in T} f_t \leq 0]$. Now observe that

$$\inf_{\mathbb{R}^n} \sup\{f_0(x) - \mu;\ f_t(x),\ t \in T\} \leq \inf_{x \in C} \sup\{f_0(x) - \mu;\ f_t(x),\ t \in T\} = 0.$$

Then, since

$$\inf_{x\in\mathbb{R}^n\setminus C} \sup\{f_0(x)-\mu,\ f_t(x);\ t\in T\} > 0,$$

relation (3.13) yields

$$\inf_{\mathbb{R}^n} \sup\{f_0-\mu,\ f_t;\ t\in T\} = \inf_{x\in C} \sup\{f_0(x)-\mu,\ f_t(x);\ t\in T\} = 0,$$

that is, $\inf_{\mathbb{R}^n} \sup\{f_0-\mu,\ f_t;\ t\in T\} = 0$. □

The following lemma gives the opposite statement in Lemma 3.6(3).

Lemma 3.7 *Given an arbitrary family of convex functions $f_0,\ f_t : \mathbb{R}^n \to \mathbb{R}_\infty$, $t \in T$ $(0 \notin T)$, such that $\inf_{f_t\le 0,\ t\in T} f_0 \in \mathbb{R}$, we suppose that Slater condition holds; that is, $\sup_{t\in T} f_t(x_0) < 0$ for some $x_0 \in \operatorname{dom} f_0$. Then we have that*

$$\inf_{\mathbb{R}^n} \sup\{f_0-\mu;\ f_t,\ t\in T\} = 0 \tag{3.14}$$

for some $\mu\in\mathbb{R}$ if and only if

$$\mu = \inf_{f_t\le 0,\ t\in T} f_0.$$

Consequently, both problems

$$\inf_{f_t\le 0,\ t\in T} f_0, \qquad \text{and} \qquad \inf_{\mathbb{R}^n} \sup\{f_0-\mu, f_t, t\in T\}$$

have the same set of ε-solutions.

Proof. According to Lemma 3.6(3), it suffices to prove the "only if" part. Let $\mu\in\mathbb{R}$ such that (3.14) holds. Then, if C is defined as in the proof of Lemma 3.6(3), we obtain

$$\inf_{\mathbb{R}^n} \sup\{f_0-\mu, f_t, t\in T\} = \inf_{C} \sup\{f_0-\mu, f_t, t\in T\} = 0. \tag{3.15}$$

Next, take $x\in C$ and denote $x_\lambda := \lambda x_0 + (1-\lambda)x \in C$, $\lambda\in[0,1]$, where x_0 is the given Slater point; hence, $x_\lambda \in C$ and $\sup_{t\in T} f_t(x_\lambda) < 0$. Consequently, (3.15) ensures that $f_0(x_\lambda)-\mu \ge 0$ and by the convexity assumption we deduce that

$$\mu \le f_0(x_\lambda) \le \lambda f_0(x_0) + (1-\lambda) f_0(x).$$

Thus, as $\lambda \downarrow 0$, we infer that $f_0(x) \ge \mu$ and the arbitrariness of $x\in C$ implies that $\inf_{f_t\le 0,\ t\in T} f_0 \ge \mu$. Conversely, taking into account (3.15), we choose a sequence $(x_k)\subset C$ such that

$$\sup\{f_0(x_k)-\mu;\ f_t(x_k), t\in T\} \to 0.$$

Then

$$\inf_{f_t\le 0,\ t\in T} (f_0-\mu) \le \limsup_{k\to+\infty} (f_0(x_k)-\mu) \le \limsup_{k\to+\infty} \sup\{f_0(x_k)-\mu;\ f_t(x_k), t\in T\} = 0,$$

and we get the remaining inequality, $\mu \ge \inf_{f_t\le 0,\ t\in T} f_0$. □

In the next result we reduce the number of constraints involved in the SIP problem introduced in (3.1), giving rise to a generalization of [42, Theorem 1]. We consider the reduced problems $(\mathcal{P}_S)$, $S \subset T$ finite, given as

$$\inf_{\substack{f_t(x)\leq 0,\ t\in S \\ x\in \operatorname{dom} f}} f_0(x), \tag{$\mathcal{P}_S$}$$

where $f := \sup_{t\in T} f_t$. We say that the subproblem $(\mathcal{P}_S)$, $S \subset T$, satisfies the Slater condition if there exists $x_0 \in \operatorname{dom} f_0 \cap \operatorname{dom} f$ such that

$$\max_{t\in S} f_t(x_0) < 0. \tag{3.16}$$

Theorem 3.8 *Consider an arbitrary family of convex functions $f_0,\ f_t : \mathbb{R}^n \to \mathbb{R}_\infty$, $t \in T$, and denote $f := \sup_{t\in T} f_t$. Assume that T is Hausdorff compact and the mappings $t \mapsto f_t(x), x \in \operatorname{dom} f$ are usc. If $\inf_{f_t(x)\leq 0,\ t\in T} f_0(x) < +\infty$ and, for all set $S \subset T$ with $|S| \leq n$, the associated problem $(\mathcal{P}_S)$ has a Slater Point. Then*

$$\inf_{f_t(x)\leq 0,\ t\in T} f_0(x) = \max_{S\subset T,\ |S|\leq n} \inf_{\substack{f_t(x)\leq 0,\ t\in S \\ x\in \operatorname{dom} f}} f_0(x).$$

Proof. We consider the SIP problem given in (3.1) and denote by μ its optimal value. We may assume that $\inf_{f_t(x)\leq 0,\ t\in T} f_0(x) \in \mathbb{R}$; otherwise, $\inf_{f_t(x)\leq 0,\ t\in T} f_0(x) = -\infty$ and the desired property obviously holds. Then, by Lemma 3.6(3), we have

$$\inf_{\mathbb{R}^n} \sup\{f_0(x) - \mu;\ f_t(x),\ t \in T\} = 0, \tag{3.17}$$

and Corollary 3.2 gives rise to some $S \subset T$ with $|S| \leq n$ such that

$$\inf_{\mathbb{R}^n} \sup\{f_0(x) - \mu + \mathrm{I}_{\operatorname{dom} f_0 \cap \operatorname{dom} f}(x);\ f_t(x) + \mathrm{I}_{\operatorname{dom} f_0 \cap \operatorname{dom} f}(x),\ t \in S\} = 0,$$

that is,

$$\inf_{\mathbb{R}^n} \sup\{f_0(x) - \mu;\ f_t(x) + \mathrm{I}_{\operatorname{dom} f}(x),\ t \in S\} = 0.$$

Moreover, since $(\mathcal{P}_S)$ satisfies the Slater condition by the current assumption, Lemma 3.7 entails that the optimal value of $(\mathcal{P}_S)$ is equal to μ, as we wanted to prove. □

The following example shows that Slater's condition is necessary to have the equivalence between SIP and Infsup problems. Otherwise, using only the information given in problem (3.12), one could not specify which function would be the objective function.

Example 3.9 *Consider the SIP problem, given in $\mathbb{R}^2$,*

$$(P) \quad \inf_{-t^3x_1 - tx_2 + 2t^2 \leq 0,\ t\in[0,1]} x_2,$$

having a 0 optimal value. Then, for every $t_1, \cdots, t_m \in [0, 1]$ and $m \geq 1$, from Figure 3.2 we can see that

$$\inf_{-t_i^3 x_1 - t_i x_2 + 2t_i^2 \leq 0,\ i=1,\cdots,m} x_2 = -\infty,$$

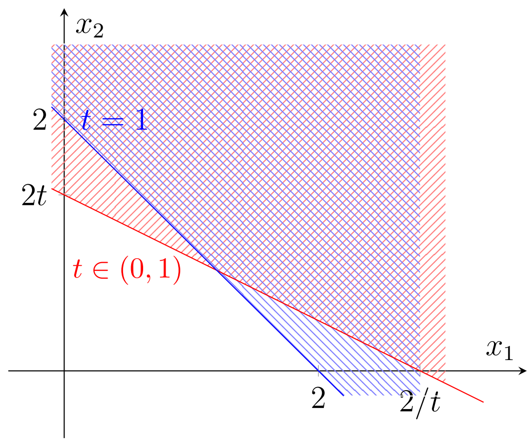


Figure 3.2: Constraint half-spaces for each $t \in (0, 1]$.

and (P) admits no discretization. At the same time, we verify that

$$\inf_{\mathbb{R}^n} \sup\{x_2; \ \ -t^3x_1 - tx_2 + 2t^2, \ t \in [0,1]\} = \inf_{\mathbb{R}^n} \max\{x_2, \ 0, \ \ -x_1 - x_2 + 2\} = 0. \quad (3.18)$$

Here we use the fact that the supremum $\sup_{t\in[0,1]}\{-t^3x_1 - tx_2 + 2t^2\}$ is attained at $t = 0$, $t = 1$, or at some $t_0 \in]0, 1[$. In the last case, t_0 would satisfy $-3t_0^2x_1 - x_2 + 4t_0 = 0$, so that $x_2 = -3t_0^2x_1 + 4t_0$ and

$$0 \leq \sup_{t\in[0,1]} \{-t^3x_1 - tx_2 + 2t^2\} = -t_0^3x_1 - t_0(-3t_0^2x_1 + 4t_0) + 2t_0^2 = 2t_0^2(t_0x_1 - 1).$$

Hence, $t_0x_1 \geq 1$, $x_1 > 0$, and t_0 cannot be a maximum point (because sufficiently small perturbations of it would provide larger value of the supremum). In other words, we find a reduction for this Infsup problem, while the SIP problem admits no discretization. In the current example, the Slater condition does not hold for subproblems having the the constraints 0 and $-x_1 - x_2 + 2$.

Consequently, every ε-solution of (P) ($\varepsilon > 0$) is an ε-solution of the problem $\inf_{\mathbb{R}^2} \max\{x_2,$ $0, -x_1 - x_2 + 2\}$.

3.4 Consequences

We present here some consequences of Theorem 3.1. We obtain an alternative-type result, which is a useful tool in optimization theory, namely within the analysis of Lagrangian duality and scalarization of vector optimization problems, among other applications. In a second step, we provide some relations between the solutions of a given SIP problem and its associated finite subproblems.

Here is the first alternative theorem.

Theorem 3.10 *Given a family of convex functions $f_t : \mathbb{R}^n \to \overline{\mathbb{R}}$, $t \in T$, such that $\operatorname{cl} f = \sup_{t\in T}(\operatorname{cl} f_t)$, we assume that T is Hausdorff compact and the mappings $t \mapsto f_t(x)$ are usc,* *for all x in a set $D \subset X$ satisfying (3.7). Then only one of the following two alternatives, (a) and (b), holds.*

a) There exists $x \in \mathbb{R}^n$ such that $\sup_{t\in T} f_t(x) < 0$.

b) There exists $\lambda \in \Delta(T)$ such that $|\operatorname{supp}\lambda| \leq n+1$ and

$$\sum_{t\in T} \lambda_t f_t(x) \geq 0, \text{ for all } x \in D.$$

Proof. We have the following equivalences:

$$\begin{aligned}
\text{Alternative } a) \text{ is not true} &\iff \sup_{t\in T} f_t(x) \geq 0, \text{ for all } x \in \mathbb{R}^n \\
&\iff \inf_{x\in\mathbb{R}^n} \sup_{t\in T} f_t(x) \geq 0 \\
&\iff \max_{\lambda\in\Delta(T),\ |\operatorname{supp}\lambda|\leq n+1} \inf_{x\in\mathbb{R}^n} \left\{ \sum_{t\in T} \lambda_t f_t(x) + \mathrm{I}_D(x) \right\} \geq 0 \\
&\qquad\qquad\qquad\qquad\qquad\qquad\text{(by Theorem 3.1)} \\
&\iff \text{there exists } \lambda \in \Delta_{(T)},\ |\operatorname{supp}\lambda| \leq n+1, \text{ such that} \\
&\qquad \sum_{t\in T} \lambda_t f_t(x) + \mathrm{I}_D(x) \geq 0, \text{ for all } x \in \mathbb{R}^n. \\
&\iff \text{Alternative } b) \text{ is true .}
\end{aligned}$$

□

The following Corollary gives another alternative theorem, which has been given in [6, Theorem 3.1] under an additional condition called *asymptotic regularity.*

Corollary 3.11 *Under the hypothesis of Theorem 3.10, exactly one of the following two alternatives holds.*

a) There is a solution in $\mathbb{R}^n$ to the system

$$f_t(x) \leq 0, \text{ for all } t \in T.$$

b) There exist $\varepsilon > 0$ and $\lambda \in \Delta(T)$, $|\operatorname{supp}\lambda| \leq n+1$, such that

$$\sum_{t\in T} \lambda_t f_t(x) \geq \varepsilon, \quad \text{for all } x \in D.$$

Proof. If the first alternative holds, that is, there exists a solution $x \in \mathbb{R}^n$ to the system $f_t(x) \leq 0$, $t \in T$, then for all $\varepsilon > 0$ there exists $0 < \varepsilon_0 < \varepsilon$ such that $f_t(x) \leq 0 < \varepsilon_0 < \varepsilon$. Thus, applying Theorem 3.10 to the functions $\tilde{f}_t(x) := f_t(x) - \varepsilon$, $t \in T$, we conclude that there is no $\lambda_\varepsilon \in \Delta(T)$ such that $|\operatorname{supp}\lambda_\varepsilon| \leq n+1$ and $\sum_{t\in T} \lambda_{\varepsilon,t}\tilde{f}_t(x) \geq 0$, for all $x \in D$.

Equivalently, there is no $\lambda_\varepsilon \in \Delta(T)$ such that $|\operatorname{supp} \lambda_\varepsilon| \leq n+1$ and $\sum_{t \in T} \lambda_{\varepsilon,t} f_t(x) \geq \varepsilon$, for all $x \in D$. Hence, the second alternative does not hold.

Conversely, if the second alternative holds, then there exist $\varepsilon > 0$ and $\lambda \in \Delta(T)$, $|\operatorname{supp} \lambda| \leq n+1$, such that

$$\sum_{t \in \operatorname{supp} \lambda} \lambda_t (f_t(x) - \varepsilon) \geq 0, \quad \forall x \in D.$$

Next, applying Theorem 3.10, the system $\sup_{t \in T}(f_t(x) - \varepsilon) < 0$ has no solution, and so does the system $f_t(x) \leq 0$, $t \in T$. □

The next two results propose duals optimization problems to (3.1).

Theorem 3.12 *Given a family of convex functions f_0, $f_t : \mathbb{R}^n \to \overline{\mathbb{R}}$, $t \in T$, and $f := \sup_{t \in T} f_t$, we suppose that $\operatorname{cl}(\max\{f, f_0\}) = \sup_{t \in T \cup \{0\}} \operatorname{cl} f_t$. We also assume that T is Hausdorff compact and the mappings $t \in T \mapsto f_t(x)$ are usc, for all x in a set $D \subset X$ satisfying (3.7). Then, for every $\varepsilon \geq 0$ and ε-minimum x_ε of problem* (3.1)*, there exists some $\lambda \in \Delta(T \cup \{0\})$ such that $|\operatorname{supp} \lambda| \leq n+1$,*

$$-\lambda_0 \varepsilon \leq \textstyle\sum_{t \in T} \lambda_t f_t(x_\varepsilon) \leq 0,$$

and x_ε is a $(\lambda_0 \varepsilon)$- minimum of the problem

$$\inf_{x \in D} \lambda_0 f_0(x) + \textstyle\sum_{t \in T} \lambda_t f_t(x).$$

Proof. Given $\varepsilon \geq 0$ and ε-minimum x_ε of (3.1), we easily verify that

$$\sup\{f_0(x) - f_0(x_\varepsilon) + \varepsilon; f_t(x), t \in T\} \geq 0, \text{ for all } x \in X.$$

Then, by Theorem 3.10, there exists $\lambda \in \Delta(T \cup \{0\})$ such that $|\operatorname{supp} \lambda| \leq n+1$ and

$$\lambda_0 \left(f_0(x) - f_0(x_\varepsilon) + \varepsilon\right) + \textstyle\sum_{t \in T} \lambda_t f_t(x) \geq 0, \text{ for all } x \in D \cap \operatorname{dom} f_0.$$

Hence, since x_ε is feasible, for all $x \in D \cap \operatorname{dom} f_0$ we get

$$\lambda_0 f_0(x) + \textstyle\sum_{t \in T} \lambda_t f_t(x) \geq \lambda_0 f_0(x_\varepsilon) - \lambda_0 \varepsilon \geq \lambda_0 f_0(x_\varepsilon) - \lambda_0 \varepsilon + \sum_{t \in T} \lambda_t f_t(x_\varepsilon),$$

and x_ε is a $(\lambda_0 \varepsilon)$- minimum of the problem $\inf_D \lambda_0 f_0 + \sum_{t \in T} \lambda_t f_t$. Moreover, the inequality in the middle of the relation above yields $-\lambda_0 \varepsilon \leq \sum_{t \in T} \lambda_t f_t(x_\varepsilon) \leq 0$. □

Corollary 3.13 *Under the hypothesis of Theorem 3.12, we suppose that all subproblems $(\mathcal{P}_S)$ with $|S| \leq n+1$ satisfies the Slater condition (see (3.16)). Then we have that*

$$\mu := \inf_{f_t \leq 0,\ t \in T} f_0 = \max_{S \subset T,\ |S| \leq n} \inf_{\substack{f_t \leq 0,\ t \in S \\ x \in D}} f_0 = \max_{\lambda \in \mathbb{R}_+^{(T)},\ |\operatorname{supp} \lambda| \leq n} \inf_{x \in D} \{f_0(x) + \sum_{t \in S} \lambda_t f_t(x)\},$$

where $D = \cap_{t \in T} \operatorname{dom} f_t$.

Proof. Lemma 3.6 (3.) implies that $\sup\{f_0(x)-\mu; f_t(x),\ t\in T\}=0$. So, using Theorem 3.10, there exists $\lambda\in\Delta(T\cup\{0\})$ such that $|\operatorname{supp}\lambda|\le n+1$ and

$$\lambda_0(f_0(x)-\mu)+\sum_{t\in T}\lambda_t f_t(x)\ge 0,\quad \forall x\in D\cap \operatorname{dom} f_0.$$

More precisely, we have $\lambda_0>0$ thanks to the Slater assumption and so, dividing on λ and taking $S:=\operatorname{supp}\lambda$,

$$\mu\ge \inf_{\substack{f_t\le 0,\ t\in S\\ x\in D}} f_0\ge \inf_{x\in D}\{f_0(x)+\sum_{t\in S}\lambda_t f_t(x)\}\ge\mu, \tag{3.19}$$

which yields the desired equalities. □

Chapter 4

Subdifferential calculus: characterizations of the normal cone to the domain of the supremum function

4.1 Introduction

As a first step towards the study of the subdifferential of supremum functions, in this chapter we begin by characterizing the normal cone to the domain of pointwise suprema of arbitrary families of convex functions. We then extend the results of [16], given in the continuous-compact setting, to provide explicit characterizations of such normal cones in terms of the underlying data without assuming any algebraic or topological conditions on either the index set or the index mappings.

It is well known that for every function $f \in \Gamma_0(X)$ and $x \in \operatorname{dom} f$, the normal cone to $\operatorname{dom} f$ is written by means of the ε-subdifferential as (see [60, Exercise 2.23])

$$\mathrm{N}_{\operatorname{dom} f}(x) = \left[\partial_\varepsilon f(x)\right]_\infty, \text{ for all } \varepsilon > 0. \tag{4.1}$$

Our objective is to extend this last relation to the case where f is the supremum of an arbitrary family of convex functions, and to provide a similar characterization which uses the ε-subdifferential of the data functions. First extensions to the continuous-compact setting have been provided recently in [12] and [34].

We consider a nonempty family of convex functions $f_t : X \to \overline{\mathbb{R}}$, $t \in T$, where X is a given lcs space, and the associated supremum function

$$f = \sup_{t \in T} f_t.$$

Instead of the lower semi-continuity, we shall use the following condition

$$\operatorname{cl} f = \sup_{t \in T} (\operatorname{cl} f_t), \tag{4.2}$$

which is called *closure condition*. Additionally, we shall consider parameters $(\rho_t)_{t\in T} \in \,]0,1]$ that satisfies

$$\inf_{t\in T} \rho_t f_t(x) > -\infty.$$

Given $x \in \operatorname{dom} f$ and $\varepsilon \geq 0$, remember that the ε-active set at x is

$$T_\varepsilon(x) := \{t \in T : \ f_t(x) \geq f(x) - \varepsilon\}, \text{ with } T(x) := T_0(x).$$

So, when $\varepsilon > 0$ the set $T_\varepsilon(x)$ is always nonempty. However, the set $T(x)$ is not necessarily nonempty. When

$$T \text{ is a Hausdorff compact and the mappings } t \mapsto f_t(x), x \in X, \text{are usc,} \tag{4.3}$$

we say that we have *compact-continuous*. In such a case, $T(x)$ is a nonempty compact subset of T.

The chapter is structured as follows: Section 4.2 is dedicated to characterizing the normal cone to the effective domain, when T is an arbitrary set of indices and the functions f_t do not necessarily belong to the family $\Gamma_0(X)$. Here it is enough to consider that each function f_t is convex and the family $\{f_t,\ t \in T\}$ satisfies the closure condition (4.2). The main result of this section is Theorem 4.7. In Section 4.3 we provide some simplifications of the results of [12] and [34] in the compact-continuous setting.

4.2 General characterizations of the normal cone

In this section we characterize the normal cone to the domain of the supremum of an arbitrary family of proper convex functions $f_t, t \in T$. For $x \in \operatorname{dom} f$ and $\varepsilon > 0$, we denote

$$\rho_{t,\varepsilon} := \begin{cases} \frac{-\varepsilon}{2f_t(x)-2f(x)+\varepsilon}, & \text{if } t \in T \setminus T_\varepsilon(x), \\ 1, & \text{if } t \in T_\varepsilon(x). \end{cases} \tag{4.4}$$

Observe that

$$\inf_{t\in T} \rho_{t,\varepsilon} f_t(x) > -\infty.$$

In fact, for all $t \in T_\varepsilon(x)$ we have that $\rho_{t,\varepsilon} f_t(x) = f_t(x) \geq f(x) - \varepsilon > -\infty$, while for all $t \in T \setminus T_\varepsilon(x)$ we obtain

$$\rho_{t,\varepsilon} f_t(x) = \frac{-\varepsilon f_t(x)}{2f_t(x)+\varepsilon} > -\varepsilon, \tag{4.5}$$

provided that $f(x) = 0$. Also, if $f(x) \neq 0$, applying the last inequality to the functions $f_t - f(x)$, we have that

$$\rho_{t,\varepsilon}(f_t - f(x))(x) = \frac{-\varepsilon(f_t(x) - f(x))}{2f_t(x) - 2f(x) + \varepsilon} > -\varepsilon,$$

and, hence,

$$\rho_{t,\varepsilon} f_t(x) > \rho_{t,\varepsilon} f(x) - \varepsilon > -\infty.$$

The first lemma of this section allows us to reduce the number of functions involved in the characterization of the subdifferential of the supremum function f.

Lemma 4.1 *Consider a family of lsc proper convex functions $f_1, \cdots, f_m : \mathbb{R}^n \to \mathbb{R}_\infty$, $m > n+1$, and denote $f := \max_{1\le i\le m} f_i$. Then, for every $x \in \operatorname{dom} f$ and $\varepsilon > 0$, we have*

$$\partial_\varepsilon f(x) \subset \bigcup \left\{\partial_\varepsilon f_\alpha(x) : f_\alpha(x) \geq f(x) - \varepsilon,\ \alpha \in \Delta_m,\ |\operatorname{supp}\alpha| \leq n+1\right\},$$

where $f_\alpha := \sum_{i\in\operatorname{supp}\alpha} \alpha_i f_i$.

Proof. We pick $x^* \in \partial_\varepsilon f(x)$. Then

$$f(x) - \langle x^*, x\rangle \leq \inf_{\mathbb{R}^n}(f - \langle x^*, \cdot\rangle) + \varepsilon, \tag{4.6}$$

and applying Corollary 3.3 to the family of proper lsc convex functions $\{f_1 - \langle x^*, \cdot\rangle, \cdots, f_m - \langle x^*, \cdot\rangle\}$ we find $\alpha \in \Delta_m$ such that $|\operatorname{supp}\alpha| \leq n+1$ and

$$\inf_{\mathbb{R}^n}(f - \langle x^*, \cdot\rangle) = \inf_{\mathbb{R}^n}\left(f_\alpha - \langle x^*, \cdot\rangle\right). \tag{4.7}$$

Thus,

$$\begin{aligned} \left(f_\alpha - \langle x^*, \cdot\rangle\right)(x) &\leq f(x) - \langle x^*, x\rangle && (4.8)\\ &\leq \inf_{\mathbb{R}^n}(f - \langle x^*, \cdot\rangle) + \varepsilon \\ &= \inf_{\mathbb{R}^n}\left(f_\alpha - \langle x^*, \cdot\rangle\right) + \varepsilon, && (4.9) \end{aligned}$$

and we deduce that $x^* \in \partial_\varepsilon f_\alpha(x)$. Moreover, from (4.8) and (4.9) we obtain the inequality $f_\alpha(x) \geq f(x) - \varepsilon$. □

Next, we give the characterization of the normal cone when all the convex functions f_t, $t \in T$, are lsc.

Theorem 4.2 *Assume that $\{f_t,\ t \in T\} \subset \Gamma_0(X)$ and denote $f := \sup_{t\in T} f_t$. Then, for every $x \in \operatorname{dom} f$,*

$$\mathrm{N}_{\operatorname{dom} f}(x) = \left[\overline{\operatorname{co}}\left(\bigcup\nolimits_{t\in T}\partial_\varepsilon(\rho_{t,\varepsilon} f_t)(x)\right)\right]_\infty, \text{ for all } \varepsilon > 0. \tag{4.10}$$

Proof. Fix $x \in \operatorname{dom} f$ and $\varepsilon > 0$. Since $T_\varepsilon(x)$ is always nonempty, we also choose $\bar{t} \in T_\varepsilon(x)$. Then, according to [34, Theorem 7] we have that

$$\mathrm{N}_{\operatorname{dom} f}(x) = \left[\overline{\operatorname{co}}\left(\bigcup\nolimits_{J\in\mathcal{T}}\partial_\varepsilon f_J(x)\right)\right]_\infty, \tag{4.11}$$

where $\mathcal{T} := \{J \subset T : \bar{t} \in J,\ |J| < +\infty\}$ and $f_J := \max_{t\in J} f_t$. We pick $J \in \mathcal{T}$, say $J = \{\bar{t}, t_1, \cdots, t_m\}$ for some $m \geq 1$.

Let us first suppose that X is a finite-dimensional space, say of dimension n so that, by Lemma 4.1,

$$\partial_\varepsilon f_J(x) \subset \bigcup\left\{\partial_\varepsilon f_\alpha(x) : f_\alpha(x) \geq f(x) - \varepsilon,\ \alpha \in \Delta_{|J|}, |\operatorname{supp}\alpha| \leq n+1\right\},$$

where $f_\alpha := \sum_{t\in\operatorname{supp}\alpha} \alpha_t f_t$. Observe that relation $f_\alpha(x) \geq f(x) - \varepsilon$ is equivalent to

$$\sum_{t\in\ \operatorname{supp}\alpha} \alpha_t(2f_t(x) - 2f(x) + \varepsilon) \geq -\varepsilon. \tag{4.12}$$

Additionally, on the one hand, for every $t \in T_\varepsilon(x)$ we have that

$$f(x) \geq f_t(x) \geq f(x) - \varepsilon,$$

so that

$$\varepsilon = 2f(x) - 2f(x) + \varepsilon \geq 2f_t(x) - 2f(x) + \varepsilon \geq -\varepsilon.$$

Hence,

$$\varepsilon \sum_{t \in T_\varepsilon(x)} \alpha_t \geq \sum_{t \in T_\varepsilon(x)} \alpha_t(2f_t(x) - 2f(x) + \varepsilon) \geq -\varepsilon \sum_{t \in T_\varepsilon(x)} \alpha_t.$$

On the other hand, for every $t \in T \backslash T_\varepsilon(x)$, we have $2f_t(x) - 2f(x) + \varepsilon < -\varepsilon$ and, so,

$$\alpha_t(2f_t(x) - 2f(x) + \varepsilon) \leq -\alpha_t \varepsilon \leq 0;$$

that is, for all $s \in T \backslash T_\varepsilon(x)$, we get

$$\alpha_s(2f_s(x) - 2f(x) + \varepsilon) \geq \sum_{t \in T \backslash T_\varepsilon(x)} \alpha_t(2f_t(x) - 2f(x) + \varepsilon). \tag{4.13}$$

Consequently, (4.12) yields

$$\sum_{t \in T \backslash T_\varepsilon(x)} \alpha_t(2f_t(x) - 2f(x) + \varepsilon) \geq -\varepsilon - \sum_{t \in T_\varepsilon(x)} \alpha_t(2f_t(x) - 2f(x) + \varepsilon) \geq -\varepsilon - \varepsilon \sum_{t \in T_\varepsilon(x)} \alpha_t.$$

Thus, since $0 \geq -\varepsilon \sum_{t \in T_\varepsilon(x)} \alpha_t \geq -\varepsilon$, by (4.13) we deduce that for all $t \in T \backslash T_\varepsilon(x)$

$$\alpha_t(2f_t(x) - 2f(x) + \varepsilon) + \varepsilon \geq -\varepsilon,$$

that is,

$$\alpha_t \leq \frac{-2\varepsilon}{2f_t(x) - 2f(x) + \varepsilon} = 2\rho_{t,\varepsilon}.$$

Therefore, using Proposition 2.34(1), by the definition of $\rho_{t,\varepsilon}$ we obtain

$$\begin{aligned}
\partial_\varepsilon f_J(x) &\subset \bigcup \left\{ \sum_{t \in \operatorname{supp} \alpha} \partial_\varepsilon (\alpha_t f_t)(x) : \alpha \in \Delta_{|J|}, |\operatorname{supp} \alpha| \leq n+1, \alpha_t \leq 2\rho_{t,\varepsilon} \right\} + \varepsilon B_{X^*} \\
&\subset (n+1) \operatorname{co} \left\{ \bigcup_{t \in \operatorname{supp} \alpha} \{\partial_\varepsilon (\alpha_t f_t)(x) : \alpha \in \Delta_{|J|}, |\operatorname{supp} \alpha| \leq n+1, \alpha_t \leq 2\rho_{t,\varepsilon}\} \right\} + \varepsilon B_{X^*},
\end{aligned}$$

where B_{X^*} is the unit ball in the dual X^* of X, which is here supposed to be finite-dimensional. Then, applying (4.11), we get

$$\begin{aligned}
\mathrm{N}_{\operatorname{dom} f}(x) = \left[\overline{\operatorname{co}}\left(\textstyle\bigcup_{J \in \mathcal{T}} \partial_\varepsilon f_J(x)\right)\right]_\infty &= \left[\overline{\operatorname{co}}\left(\textstyle\bigcup_{J \in \mathcal{T}} \frac{1}{n+1} \partial_\varepsilon f_J(x)\right)\right]_\infty \\
&\subset \left[\overline{\operatorname{co}}\left(\operatorname{co}\left\{\textstyle\bigcup_{t \in T,\ 0 \leq \alpha_t \leq 2\rho_{t,\varepsilon}} \partial_\varepsilon (\alpha_t f_t)(x)\right\} + \varepsilon B_{X^*}\right)\right]_\infty \\
&= \left[\overline{\operatorname{co}}\left(\textstyle\bigcup_{t \in T,\ 0 \leq \alpha_t \leq 2\rho_{t,\varepsilon}} \partial_\varepsilon (\alpha_t f_t)(x)\right)\right]_\infty \qquad (4.14) \\
&\subset \left[\overline{\operatorname{co}}\left(\bigcup_{t \in T} \partial_{2\varepsilon} (2\rho_{t,\varepsilon} f_t)(x)\right)\right]_\infty \qquad (4.15) \\
&= \left[\overline{\operatorname{co}}\left(\bigcup_{t \in T} \partial_\varepsilon (\rho_{t,\varepsilon} f_t)(x)\right)\right]_\infty,
\end{aligned}$$

where the inclusion in (4.15) follows from Lemma 2.35. Then we proved that the inclusion "$\subset$" in the statement of the theorem holds.

Now we give the proof in the general case, where X is an lcs. We pick a θ-neighborhood U and choose $L \in \mathcal{F}(x)$ such that $L^{\perp} \subset U$. We also denote $\tilde{f}_t := f_t + \mathrm{I}_L, t \in T$, so that the domain of each function f_t is included in the finite-dimensional subspace L. Consequently, the domain of the associated supremum function $\tilde{f} := \sup_{t\in T} \tilde{f}_t = f + \mathrm{I}_L$ is also included in L, and by applying the first part of the proof together with Proposition 2.34(1) we deduce that

$$\begin{aligned}\mathrm{N}_{\mathrm{dom}\, f}(x) &\subset \mathrm{N}_{L\cap \mathrm{dom}\, f}(x)\\ &\subset \left[\overline{\mathrm{co}}\left(\bigcup_{t\in T} \partial_{\varepsilon}(\rho_{t,\varepsilon}\tilde{f}_t)(x)\right)\right]_{\infty}.\end{aligned}$$

Moreover, by applying Proposition 2.34(1) to the function $\rho_{t,\varepsilon}\tilde{f}_t := \rho_{t,\varepsilon} f_t + \mathrm{I}_L$ we obtain

$$\begin{aligned}\partial_{\varepsilon}(\rho_{t,\varepsilon}\tilde{f}_t)(x) &= \mathrm{cl}\left(\bigcup_{\substack{\varepsilon_1,\varepsilon_2\geq 0,\\ \varepsilon=\varepsilon_1+\varepsilon_2}} \partial_{\varepsilon_1}(\rho_{t,\varepsilon} f_t)(x) + \partial_{\varepsilon_2}\mathrm{I}_L(x)\right)\\ &\subset \mathrm{cl}\left(\partial_{\varepsilon}(\rho_{t,\varepsilon} f_t)(x) + L^{\perp}\right),\end{aligned}$$

and, so

$$\begin{aligned}\mathrm{N}_{\mathrm{dom}\, f}(x) &\subset \left[\overline{\mathrm{co}}\left(\bigcup_{t\in T} \mathrm{cl}\left(\partial_{\varepsilon}(\rho_{t,\varepsilon} f_t)(x) + L^{\perp}\right)\right)\right]_{\infty}\\ &= \left[\overline{\mathrm{co}}\left(\bigcup_{t\in T} \partial_{\varepsilon}(\rho_{t,\varepsilon} f_t)(x) + L^{\perp}\right)\right]_{\infty}.\end{aligned}$$

In this way, we have proved that, for every finite-dimensional subspace $L \in \mathcal{F}(x)$,

$$\mathrm{N}_{\mathrm{dom}\, f}(x) \subset \left[\overline{\mathrm{co}}\left(\bigcup_{t\in T} \partial_{\varepsilon}(\rho_{t,\varepsilon} f_t)(x) + L^{\perp}\right)\right]_{\infty},$$

and so the desired inclusion comes from Lemma 2.8. Finally, to prove the opposite inclusion we follows the arguments used in the proof of [12, Theorem 5]. We denote $E_{\varepsilon} := \bigcup_{t\in T} \partial_{\varepsilon}(\rho_{t,\varepsilon} f_t)(x)$, and fix $x_0^* \in E_{\varepsilon}$. Then, given $x^* \in [\overline{\mathrm{co}}(E_{\varepsilon})]_{\infty}$, for all $\beta > 0$ we have $x_0^* + \beta x^* \in \overline{\mathrm{co}}(E_{\varepsilon})$ and, so, there are nets $(\lambda_{j,1}, \cdots, \lambda_{j,k_j}) \in \Delta_{k_j}^{+}$ (where $\Delta_{k_j}^{+} := \{\lambda \in \Delta_{k_j} : \lambda_t > 0, 1 \leq t \leq k_j\}$), elements $t_{j,1}, \cdots, t_{j,k_j} \in T$ and $x_{j,1}^* \in \partial_{\varepsilon}(\rho_{t_{j,1},\varepsilon} f_{t_{j,1}})(x), \cdots, x_{j,k_j}^* \in \partial_{\varepsilon}(\rho_{t_{j,k_j},\varepsilon} f_{t_{j,k_j}})(x)$ such that

$$x_0^* + \beta x^* = \lim_j (\lambda_{j,1} x_{j,1}^* + \cdots + \lambda_{j,k_j} x_{j,k_j}^*).$$

Hence, for every fixed $y \in \mathrm{dom}\, f$,

$$\begin{aligned}\langle x_0^* + \beta x^*, y - x\rangle &= \lim_j \langle \lambda_{j,1} x_{j,1}^* + \cdots + \lambda_{j,k_j} x_{j,k_j}^*, y - x\rangle\\ &\leq \limsup_j \sum_{i=1,\cdots,k_j} \lambda_{j,i}\left(\rho_{t_{j,i},\varepsilon} f_{t_{j,i}}(y) - \rho_{t_{j,i},\varepsilon} f_{t_{j,i}}(x) + \varepsilon\right)\\ &\leq \limsup_j \sum_{i=1,\cdots,k_j} \lambda_{j,i}\left(\rho_{t_{j,i},\varepsilon} f^{+}(y) - \rho_{t_{j,i},\varepsilon} f_{t_{j,i}}(x) + \varepsilon\right)\\ &\leq f^{+}(y) - \inf\{\rho_{t,\varepsilon} f_t(x), t \in T\} + \varepsilon,\end{aligned}$$

and, dividing by β and taking the limit $\beta \to +\infty$, the hypothesis over the parameter leads us to

$$\langle x^*, y - x \rangle \leq 0,$$

for all $y \in \operatorname{dom} f = \operatorname{dom} f^+$, that is, $x^* \in \mathrm{N}_{\operatorname{dom} f}(x)$. □

Remark 4.3 *As we can see in the proof above, the specification of the value of the parameters $\rho_{t,\varepsilon}$ is only used to verify the inclusion "$\supset$" in (4.10). For instance, the same proof as above shows that, for every $x \in \operatorname{dom} f$, we have that*

$$\mathrm{N}_{\operatorname{dom} f}(x) \subset \left[\overline{\operatorname{co}}\left(\textstyle\bigcup_{t\in T}\partial_\varepsilon f_t(x)\right)\right]_\infty, \ \textit{for all } \varepsilon > 0. \tag{4.16}$$

This observations leads to a simple characterization of the normal cone.

Corollary 4.4 *Assume that $\{f_t,\ t \in T\} \subset \Gamma_0(X)$, $f := \sup_{t\in T} f_t$ and $\inf_{t\in T} f_t(x) > -\infty$. Then, for every $x \in \operatorname{dom} f$,*

$$\mathrm{N}_{\operatorname{dom} f}(x) = \left[\overline{\operatorname{co}}\left(\textstyle\bigcup_{t\in T}\partial_\varepsilon f_t(x)\right)\right]_\infty, \ \textit{for all } \varepsilon > 0. \tag{4.17}$$

Proof. According to Remark 4.3, we only need to show the inclusion "$\supset$". This can be done by following the same arguments as those used at the end of the proof of Theorem 4.2. □

The following example shows that the family of parameters $\rho_{t,\varepsilon}$ cannot be ignored.

Example 4.5 *Consider the proper lsc convex functions $f_t : \mathbb{R} \to \mathbb{R}_\infty$, $t \in T := \mathbb{R}_+$, defined as $f_t(x) := tx - t$. Set*

$$f(x) := \sup_{t\in T} f_t(x) = \begin{cases} 0, & \textit{if } x \leq 1, \\ +\infty, & \textit{if } x > 1, \end{cases}$$

so that $f(1) = f(0) = 0$, $\operatorname{dom} f =]-\infty, 1]$ and, for each $\varepsilon > 0$ and $x \in \operatorname{dom} f$, $\partial_\varepsilon f_t(x) = \{t\}$. Then

$$\left[\overline{\operatorname{co}}\left(\textstyle\bigcup_{t\in T}\partial_\varepsilon f_t(0)\right)\right]_\infty = \left[\overline{\operatorname{co}}\left(\textstyle\bigcup_{t\in T}\partial_\varepsilon f_t(1)\right)\right]_\infty = \mathbb{R}_+.$$

If $x = 1$, then $\mathrm{N}_{\operatorname{dom} f}(1) = \mathbb{R}_+$ and we have the equality

$$\left[\overline{\operatorname{co}}\left(\textstyle\bigcup_{t\in T}\partial_\varepsilon f_t(1)\right)\right]_\infty = \mathrm{N}_{\operatorname{dom} f}(1).$$

Formula (4.17) can be applied in this case because $\inf_{t\in T} f_t(1) = 0 > -\infty$. On the other hand, for $x = 0$, we have that

$$\mathrm{N}_{\operatorname{dom} f}(0) = \{0\} \subsetneq \left[\overline{\operatorname{co}}\left(\textstyle\bigcup_{t\in T}\partial_\varepsilon f_t(0)\right)\right]_\infty.$$

Here, formula (4.17) cannot be applied because $\inf_{t\in T} f_t(0) = \inf_{t\geq 0}(-t) = -\infty$.

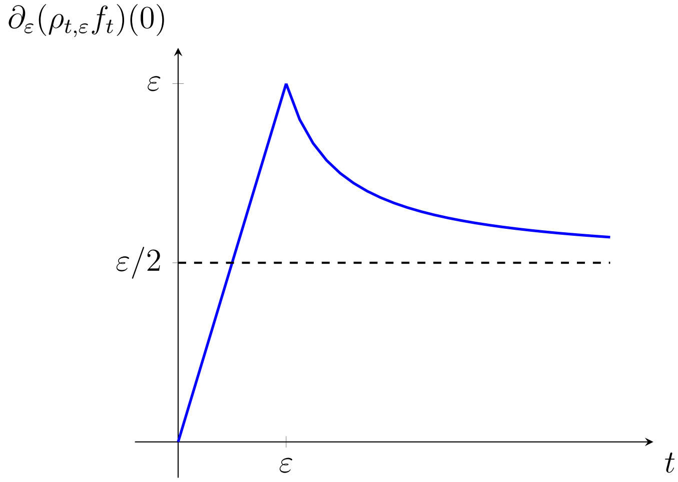


Figure 4.1: $\partial_\varepsilon(\rho_{t,\varepsilon} f_t)(0)$ for each $t \in \mathbb{R}_+$

In any case, using the parameters $\rho_{t,\varepsilon}$ for every $\varepsilon > 0$ formula (4.10) gives us

$$\begin{aligned} \mathrm{N}_{\mathrm{dom}\, f}(0) &= \left[\overline{\mathrm{co}}\left(\left(\bigcup_{t\in T_\varepsilon(0)} \partial_\varepsilon f_t(0)\right) \cup \left(\bigcup_{t\in T\setminus T_\varepsilon(0)} \partial_\varepsilon(\rho_{t,\varepsilon} f_t)(0)\right)\right)\right]_\infty \\ &= \left[\overline{\mathrm{co}}\left(\left(\bigcup_{0\le t\le\varepsilon}\{t\}\right) \cup \left(\bigcup_{t>\varepsilon}\left\{\frac{\varepsilon t}{2t-\varepsilon}\right\}\right)\right)\right]_\infty \\ &= \left[\overline{\mathrm{co}}\left([0,\varepsilon]\right)\right]_\infty = \{0\}\,. \end{aligned}$$

To give the main result of this section, we first establish the following technical lemma.

Lemma 4.6 *Consider a family of convex functions $f_t : X \to \overline{\mathbb{R}},\ t \in T$ and $f := \sup_{t\in T} f_t$ such that condition (4.2) holds, that is, $\mathrm{cl}\, f = \sup_{t\in T}(\mathrm{cl}\, f_t)$. Given the sets $T_1 := \{t \in T :$ $\mathrm{cl}\, f_t$ is proper $\}$ and $T_2 := T \setminus T_1$, we consider the functions*

$$\tilde{f}_t := \begin{cases} (f_t)^+, & \text{if } t \in T_2, \\ f_t, & \text{if } t \in T_1, \end{cases}$$

and the associated supremum $\tilde{f} := \sup_{t\in T} \tilde{f}_t$. Then we have

$$\mathrm{cl}\, \tilde{f} = \sup_{t\in T}(\mathrm{cl}\, \tilde{f}_t)$$

Proof. First, observe that

$$\tilde{f} := \sup_{t\in T} \tilde{f}_t = \max\{\sup_{t\in T_1} f_t, \sup_{t\in T_2}(f_t^+)\} \tag{4.18}$$

$$= \max\{\sup_{t\in T_1} f_t, \sup_{t\in T_2} f_t, 0\} = \max\{\sup_{t\in T} f_t, 0\} = (f)^+. \tag{4.19}$$

Then, on the other hand, by applying Lemma 2.20(2) twice we obtain that $\mathrm{cl}(f_t^+) = (\mathrm{cl}\, f_t)^+$ and

$$\sup_{t\in T_2} \mathrm{cl}(f_t^+) = (\sup_{t\in T_2} \mathrm{cl}\, f_t)^+.$$

Hence, by the current assumption,

$$\begin{aligned}\sup_{t\in T}\operatorname{cl}(\tilde{f}_t) &= \max\{\sup_{t\in T_1}\operatorname{cl}(f_t),\sup_{t\in T_2}\operatorname{cl}(\tilde{f}_t)\} = \max\{\sup_{t\in T_1}\operatorname{cl}(f_t),\sup_{t\in T_2}\operatorname{cl}(f_t^+)\}\\ &= \max\{\sup_{t\in T_1}\operatorname{cl}(f_t),(\sup_{t\in T_2}\operatorname{cl} f_t)^+\} = \max\{\sup_{t\in T}\operatorname{cl}(f_t),0\}\\ &= \max\{\operatorname{cl}(\sup_{t\in T} f_t),0\} = (\operatorname{cl} f)^+ = \operatorname{cl}(f^+).\end{aligned}$$

□

Theorem 4.7 *Consider a family of convex functions $\{f_t : X \to \overline{\mathbb{R}},\ t \in T\}$ and denote $f := \sup_{t\in T} f_t$. We assume that condition (4.2) holds, that is,*

$$\operatorname{cl} f = \sup_{t\in T}(\operatorname{cl} f_t).$$

If f is lsc at $x \in f^{-1}(\mathbb{R})$, then for all $\varepsilon > 0$ we have

$$\mathrm{N}_{\operatorname{dom} f}(x) = \left[\overline{\operatorname{co}}\left(\left(\bigcup_{t\in T_0}\partial_\varepsilon(\rho_{t,\varepsilon} f_t)(x)\right)\cup\left(\bigcup_{t\in T\setminus T_0}\mathrm{N}^\varepsilon_{\operatorname{dom} f_t}(x)\right)\right)\right]_\infty,$$

where $T_0 := \{t \in T : \operatorname{cl} f_t$ is proper$\}$.

Proof. We fix $x \in \operatorname{dom} f$ and $\varepsilon > 0$. We may assume, without loss of generality, that $f(x) = 0$ (it suffices to work with the functions $f_t - f(x)$). We start by supposing that

$$\{(\operatorname{cl} f_t),\ t \in T\} \subset \Gamma_0(X),$$

that is, all the functions $(\operatorname{cl} f_t)$, $t \in T$, are proper so that $T = T_0$. Then, since $x \in \operatorname{dom} f \subset \operatorname{dom}(\operatorname{cl} f)$, we have that

$$\mathrm{N}_{\operatorname{dom} f}(x) = \mathrm{N}_{\operatorname{cl}(\operatorname{dom} f)}(x) = \mathrm{N}_{\operatorname{cl}(\operatorname{dom}(\operatorname{cl} f))}(x) = \mathrm{N}_{\operatorname{dom}(\operatorname{cl} f)}(x),$$

and Theorem 4.2 applied to the family $\{(\operatorname{cl} f_t), t \in T\}$ yields

$$\mathrm{N}_{\operatorname{dom} f}(x) = \mathrm{N}_{\operatorname{dom}(\operatorname{cl} f)}(x) = \left[\overline{\operatorname{co}}\left(\bigcup_{t\in T}\partial_\varepsilon(\tilde{\rho}_{t,\varepsilon}(\operatorname{cl} f_t))(x)\right)\right]_\infty, \tag{4.20}$$

where (recall that $(\operatorname{cl} f)(x) = f(x) = 0$ by the assumption of the theorem)

$$\tilde{\rho}_{t,\varepsilon} := \begin{cases} \frac{-\varepsilon}{2(\operatorname{cl} f_t)(x)+\varepsilon}, & \text{if } t \in T\setminus \widetilde{T}_\varepsilon(x),\\ 1, & \text{if } t \in \widetilde{T}_\varepsilon(x),\end{cases}$$

and $\widetilde{T}_\varepsilon(x) := \{t \in T : (\operatorname{cl} f_t)(x) \geq -\varepsilon\}$. Observe that

$$\widetilde{T}_\varepsilon(x) \subset \{t \in T : f_t(x) \geq -\varepsilon\} = T_\varepsilon(x)$$

and, for all $t \in T \setminus T_\varepsilon(x)$ ($\subset T \setminus \widetilde{T}_\varepsilon(x)$), we have that

$$\tilde{\rho}_{t,\varepsilon} = \frac{\varepsilon}{-2f_t(x) - \varepsilon + 2f_t(x) - 2(\operatorname{cl} f_t)(x)} \leq \frac{\varepsilon}{-2f_t(x)-\varepsilon} = \rho_{t,\varepsilon}. \tag{4.21}$$

At the same time, for each $t \in \widetilde{T}_\varepsilon(x)$ we have that $(\operatorname{cl} f_t)(x) \geq -\varepsilon \geq f_t(x) - \varepsilon$ (as $f(x) = 0$) and so, on the one hand,

$$\partial_\varepsilon(\tilde{\rho}_{t,\varepsilon}(\operatorname{cl} f_t))(x) = \partial_\varepsilon(\operatorname{cl} f_t)(x) \subset \partial_{2\varepsilon} f_t(x).$$

On the other hand, using (4.5) and again the fact that $f(x) = 0$, for each $t \in T \setminus \widetilde{T}_\varepsilon(x)$ we have that $\tilde{\rho}_{t,\varepsilon}(\operatorname{cl} f_t)(x) > -\varepsilon \geq \tilde{\rho}_{t,\varepsilon} f_t(x) - \varepsilon$, and so

$$\partial_\varepsilon(\tilde{\rho}_{t,\varepsilon}(\operatorname{cl} f_t))(x) = \partial_\varepsilon(\operatorname{cl}(\tilde{\rho}_{t,\varepsilon} f_t))(x) \subset \partial_{2\varepsilon}(\tilde{\rho}_{t,\varepsilon} f_t)(x).$$

Thus, by (4.30),

$$\mathrm{N}_{\operatorname{dom} f}(x) \subset \left[\overline{\operatorname{co}}\left(\bigcup\nolimits_{t \in T} \partial_{2\varepsilon}(\tilde{\rho}_{t,\varepsilon} f_t)(x)\right)\right]_\infty.$$

Moreover, taking into account (4.31) and the fact that $\tilde{\rho}_{t,\varepsilon} \leq 1$, we have that $\tilde{\rho}_{t,\varepsilon} \leq \rho_{t,\varepsilon}$ for all $t \in T$, and Lemma 2.35 leads us to

$$\mathrm{N}_{\operatorname{dom} f}(x) \subset \left[\overline{\operatorname{co}}\left(\bigcup\nolimits_{t \in T} \partial_{2\varepsilon}(\rho_{t,\varepsilon} f_t)(x)\right)\right]_\infty.$$

More specifically, arguing as in the proof of Theorem 4.2, the last inclusion becomes an equality.

Now, we deal with the general case when not all the $\operatorname{cl} f_t$,'s are necessarily proper. We consider the new functions

$$\tilde{f}_t := \begin{cases} (f_t)^+, & \text{if } t \in T \setminus T_0, \\ f_t, & \text{if } t \in T_0, \end{cases}$$

together with the associated supremum function

$$\tilde{f} := \sup_{t \in T} \tilde{f}_t.$$

Observe that, due to the condition $f(x) = (\operatorname{cl} f)(x) = \sup_{t \in T}(\operatorname{cl} f_t)(x) = 0$, we have that $T_0 \neq \emptyset$. Also, since that (see Lemma 4.6)

$$\operatorname{cl} \tilde{f} = \sup_{t \in T}(\operatorname{cl} \tilde{f}_t)$$

and $(\operatorname{cl} \tilde{f})(x) = \tilde{f}(x) = 0$, by applying the paragraph above to the new family $\{\tilde{f}_t,\ t \in T\} \subset \Gamma_0(X)$ we deduce that

$$\mathrm{N}_{\operatorname{dom} f}(x) = \mathrm{N}_{\operatorname{dom} \tilde{f}}(x) = \left[\overline{\operatorname{co}}\left(\bigcup\nolimits_{t \in T} \partial_\varepsilon(\operatorname{cl}(\rho_{t,\varepsilon} \tilde{f}_t))(x)\right)\right]_\infty.$$

Since, $(\operatorname{cl}(\rho_{t,\varepsilon} \tilde{f}_t))'_\varepsilon(x; \cdot) = \sigma_{\partial_\varepsilon(\operatorname{cl}(\rho_{t,\varepsilon} \tilde{f}_t))(x)}(\cdot)$ by [60, Theorem 2.4.11] for all $t \in T$, by the bipolar theorem and taking into account Lemma 2.13 the previous inclusion reads

$$\begin{aligned}
\operatorname{cl}\left(\operatorname{dom} \sup_{t \in T}(\operatorname{cl}(\rho_{t,\varepsilon} \tilde{f}_t))'_\varepsilon(x; \cdot)\right) &\subset \operatorname{cl}\left(\operatorname{dom} \sup_{t \in T} \sigma_{\partial_\varepsilon(\operatorname{cl}(\rho_{t,\varepsilon} \tilde{f}_t))(x)}\right) \\
&= \operatorname{cl}\left(\operatorname{dom} \sigma_{\bigcup_{t \in T} \partial_\varepsilon(\operatorname{cl}(\rho_{t,\varepsilon} \tilde{f}_t))(x)}\right) \\
&= \left(\left[\overline{\operatorname{co}}\left(\bigcup\nolimits_{t \in T} \partial_\varepsilon(\operatorname{cl}(\rho_{t,\varepsilon} \tilde{f}_t))(x)\right)\right]_\infty\right)^- = \left(\mathrm{N}_{\operatorname{dom} f}(x)\right)^-.
\end{aligned}$$

Now observe that

$$
\begin{aligned}
&\operatorname{cl}\left(\operatorname{dom}\sup_{t\in T}(\operatorname{cl}(\rho_{t,\varepsilon}\tilde{f}_t))'_\varepsilon(x;\cdot)\right)\\
=\;&\operatorname{cl}\left(\operatorname{cl}\operatorname{dom}(\sup_{t\in T_0}(\operatorname{cl}(\rho_{t,\varepsilon}f_t))'_\varepsilon(x;\cdot))\cap\operatorname{cl}\operatorname{dom}(\sup_{t\in T\setminus T_0}\left((\operatorname{cl}(\rho_{t,\varepsilon}f_t)^+)'_\varepsilon(x;\cdot)\right))\right). \quad (4.22)
\end{aligned}
$$

If $t\in T_0$, then the associated functions $\operatorname{cl} f_t$ and $\rho_{t,\varepsilon}(\operatorname{cl} f_t)=(\operatorname{cl}\rho_{t,\varepsilon}f_t)$ belong to $\Gamma_0(X)$. Then $\operatorname{dom}(\rho_{t,\varepsilon}(\operatorname{cl} f_t))=\operatorname{dom}(\operatorname{cl} f_t)$, and the function $(\operatorname{cl}(\rho_{t,\varepsilon}f_t))'_\varepsilon(x;\cdot)$ is lsc by [60, Theorem 2.4.11]. So,

$$\sup_{t\in T_0}(\operatorname{cl}(\rho_{t,\varepsilon}f_t))'_\varepsilon(x;\cdot)=\sup_{t\in T_0}\sigma_{\partial_\varepsilon(\operatorname{cl}(\rho_{t,\varepsilon}f_t)(x)}=\sigma_{\cup_{t\in T_0}\partial_\varepsilon(\operatorname{cl}(\rho_{t,\varepsilon}f_t)(x)}=\sigma_A,$$

where

$$A=\cup_{t\in T_0}\partial_\varepsilon(\operatorname{cl}(\rho_{t,\varepsilon}f_t))(x),$$

that is, using Lemma 2.13,

$$\operatorname{cl}(\operatorname{dom}(\sigma_A))=[\overline{\operatorname{co}}(A)]^-_\infty. \quad (4.23)$$

Otherwise, if $t\in T\setminus T_0$, then by applying Lemma 2.20(3) to the non-proper convex function $\operatorname{cl}(\rho_{t,\varepsilon}\tilde{f}_t)$ we get

$$
\begin{aligned}
\operatorname{cl}(\rho_{t,\varepsilon}(f_t)^+) &= \mathrm{I}_{\operatorname{dom}(\operatorname{cl} f_t)},\\
(\operatorname{cl}(\rho_{t,\varepsilon}(f_t)^+))'_\varepsilon(x;\cdot) &= \sigma_{\partial_\varepsilon(\mathrm{I}_{\operatorname{dom}(\operatorname{cl} f_t)})(x)}=\sigma_{\mathrm{N}^\varepsilon_{\operatorname{dom}(\operatorname{cl} f_t)(x)}}=\sigma_{\mathrm{N}^\varepsilon_{\operatorname{dom} f_t}(x)},
\end{aligned}
$$

and

$$\sup_{t\in T\setminus T_0}\left((\operatorname{cl}(\rho_{t,\varepsilon}f_t)^+)'_\varepsilon(x;\cdot)\right)=\sigma_{\cup_{t\in T\setminus T_0}\mathrm{N}^\varepsilon_{\operatorname{dom} f_t}(x)}=\sigma_B,$$

where

$$B=\cup_{t\in T\setminus T_0}\mathrm{N}^\varepsilon_{\operatorname{dom} f_t}(x).$$

Moreover, again by Lemma 2.13 we have that

$$\operatorname{cl}(\operatorname{dom}(\sigma_B))=[\overline{\operatorname{co}}(B)]^-_\infty. \quad (4.24)$$

Now, since $\theta\in[\overline{\operatorname{co}}(A)]_\infty\cap[\overline{\operatorname{co}}(B)]_\infty$, Lemma 2.14 leads us to

$$([\overline{\operatorname{co}}(A\cup B)]_\infty)^-\subset[\overline{\operatorname{co}}(A)]^-_\infty\cap[\overline{\operatorname{co}}(B)]^-_\infty. \quad (4.25)$$

Therefore, combining (4.23), (4.24), and (4.25) in (4.22), we obtain that

$$([\overline{\operatorname{co}}(A\cup B)]_\infty)^-\subset(\mathrm{N}_{\operatorname{dom} f}(x))^-$$

and, thus,

$$\mathrm{N}_{\operatorname{dom} f}(x)\subset[\overline{\operatorname{co}}(A\cup B)]_\infty.$$

Conversely, by Lemma 2.6 and following the idea in the proof of Theorem 4.2 , we denote

$$E_\varepsilon:=A+B,$$

and fix $x_0^* \in E_\varepsilon$ since this last set is not empty. Then, given $x^* \in [\overline{\mathrm{co}}(E_\varepsilon)]_\infty$, for all $\beta > 0$ we have $x_0^* + \beta x^* \in \overline{\mathrm{co}}(E_\varepsilon)$ and, so, there are nets

$$(\lambda_{j,1}, \cdots, \lambda_{j,k_j}) \in \Delta_{k_j} \text{ and } t_{j,1}, \cdots, t_{j,k_j} \in T,\ k_j \geq 0,$$

$$u_{j,i}^* \in \partial_\varepsilon(\rho_{t_{j,i},\varepsilon} f_{t_{j,i}})(x) \text{ and } w_{j,i}^* \in \mathrm{N}^\varepsilon_{\mathrm{dom}\, f_{t_{j,i}}}(x),\ i = 1, \cdots, k_j,$$

such that the elements $x_{j,i}^* := u_{j,i}^* + w_{j,i}^*$ satisfy

$$x_0^* + \beta x^* = \lim_j (\lambda_{j,1} x_{j,1}^* + \cdots + \lambda_{j,k_j} x_{j,k_j}^*).$$

Hence, for every fixed $y \in \mathrm{dom}\, f$ $(\subset \mathrm{dom}\, f_t,\ \text{ for all } t \in T)$,

$$\begin{aligned} \langle x_0^* + \beta x^*, y - x \rangle &= \lim_j \langle \lambda_{j,1} x_{j,1}^* + \cdots + \lambda_{j,k_j} x_{j,k_j}^*, y - x \rangle \\ &= \lim_j \sum_{i=1,\cdots,k_j} \lambda_{j,i} \left(\langle u_{j,i}^*, y - x \rangle + \langle w_{j,i}^*, y - x \rangle \right) \\ &\leq \limsup_j \sum_{i=1,\cdots,k_j} \lambda_{j,i} \left(\rho_{t_{j,i},\varepsilon} f^+(y) - \rho_{t_{j,i},\varepsilon} f_{t_{j,i}}(x) + 2\varepsilon \right) \\ &\leq f^+(y) - \inf\{\rho_{t,\varepsilon} f_t(x),\ t \in T\} + 2\varepsilon, \end{aligned}$$

and, dividing by β and taking the limit when $\beta \to +\infty$, the hypothesis over the parameters leads us to

$$\langle x^*, y - x \rangle \leq 0, \quad \text{for all } y \in \mathrm{dom}\, f = \mathrm{dom}\, f^+,$$

that is, $x^* \in \mathrm{N}_{\mathrm{dom}\, f}(x)$. □

The following corollary provides a sharper characterization of the normal cone to dom f.

Corollary 4.8 *With the assumptions of Theorem 4.7 we have, for all $x \in f^{-1}(\mathbb{R})$ such that f is lsc at x and all $\varepsilon > 0$,*

$$\mathrm{N}_{\mathrm{dom}\, f}(x) = \left[\overline{\mathrm{co}} \left(\left(\bigcup_{t \in \hat{T}} \partial_\varepsilon(\rho_{t,\varepsilon} f_t)(x) \right) \cup \left(\bigcup_{t \in T \setminus T_0} \mathrm{N}^\varepsilon_{\mathrm{dom}\, f_t}(x) \right) \right) \right]_\infty,$$

where

$$\hat{T} := \{ t \in T : \partial_\varepsilon(\rho_{t,\varepsilon} f_t)(x) \neq \emptyset \}.$$

Proof. For all $t \in \hat{T}$ we have that

$$\partial_\varepsilon(\rho_{t,\varepsilon} f_t)(x) + \mathrm{N}^\varepsilon_{\mathrm{dom}\, f_t}(x) \subset \partial_{2\varepsilon}(\rho_{t,\varepsilon} f_t)(x).$$

Thus, by Theorem 4.7,

$$\begin{aligned} \mathrm{N}_{\mathrm{dom}\, f}(x) &\subset \left[\overline{\mathrm{co}} \left(\left(\bigcup_{t \in \hat{T}} \partial_\varepsilon(\rho_{t,\varepsilon} f_t)(x) \right) \cup \left(\bigcup_{t \in T \setminus T_0} \mathrm{N}^\varepsilon_{\mathrm{dom}\, f_t}(x) \right) \right) \right]_\infty \\ &\subset \left[\overline{\mathrm{co}} \left(\bigcup_{t \in \hat{T}} \left(\partial_\varepsilon(\rho_{t,\varepsilon} f_t)(x) \cup \mathrm{N}^\varepsilon_{\mathrm{dom}\, f_t}(x) \right) \cup \left(\bigcup_{t \in T \setminus T_0} \mathrm{N}^\varepsilon_{\mathrm{dom}\, f_t}(x) \right) \right) \right]_\infty. \end{aligned}$$

Hence, using Lemma 2.6,

$$\begin{aligned}\mathrm{N}_{\operatorname{dom} f}(x) &\subset \left[\overline{\mathrm{co}}\left(\left(\bigcup_{t\in\hat{T}}\partial_{\varepsilon}(\rho_{t,\varepsilon}f_t)(x)+\mathrm{N}^{\varepsilon}_{\operatorname{dom} f_t}(x)\right)\cup\left(\bigcup_{t\in T\setminus T_0}\mathrm{N}^{\varepsilon}_{\operatorname{dom} f_t}(x)\right)\right)\right]_{\infty}\\ &\subset \left[\overline{\mathrm{co}}\left(\left(\bigcup_{t\in\hat{T}}\partial_{2\varepsilon}(\rho_{t,\varepsilon}f_t)(x)\right)\cup\left(\bigcup_{t\in T\setminus T_0}\mathrm{N}^{2\varepsilon}_{\operatorname{dom} f_t}(x)\right)\right)\right]_{\infty},\end{aligned}$$

and the fact $\hat{T}\subset T_0$ implies, again by Theorem 4.7,

$$\mathrm{N}_{\operatorname{dom} f}(x) \subset \left[\overline{\mathrm{co}}\left(\left(\bigcup_{t\in T_0}\partial_{2\varepsilon}(\rho_{t,\varepsilon}f_t)(x)\right)\cup\left(\bigcup_{t\in T\setminus T_0}\mathrm{N}^{2\varepsilon}_{\operatorname{dom} f_t}(x)\right)\right)\right]_{\infty} = \mathrm{N}_{\operatorname{dom} f}(x).$$

□

It is worth observing that if $t\in T$ is such that $f_t\in\Gamma_0(X)$, then the corresponding term $\mathrm{N}^{\varepsilon}_{\operatorname{dom} f_t}(x)$ can be removed form the characterization of the normal cone to $\operatorname{dom} f$. More precisely, the following corollary comes easily from Theorem 4.7.

Corollary 4.9 *Consider a family $\{f_t : X\to\mathbb{R}_\infty,\ t\in T\}$ of proper convex lsc functions and $f := \sup_{t\in T} f_t$. Then, for every $x\in\operatorname{dom} f$, we have*

$$\mathrm{N}_{\operatorname{dom} f}(x) = \left[\overline{\mathrm{co}}\left(\bigcup_{t\in T}\partial_{\varepsilon}(\rho_{t,\varepsilon}f_t)(x)\right)\right]_{\infty},\quad \textit{for all } \varepsilon>0.$$

Proof. In the current case, we have that $\hat{T}=T_0=T$. Thus, it suffices to apply Theorem 4.7. □

The characterization in Theorem 4.7 takes a very simple form when the functions $f_t, t\in T$ are uniformly bounded below at x, as the following corollary shows.

Corollary 4.10 *Assume in Theorem 4.7 that, additionally, the point x satisfies*

$$\inf_{t\in T} f_t(x) > -\infty.$$

Then, for every $\varepsilon>0$, we have

$$\mathrm{N}_{\operatorname{dom} f}(x) = \left[\overline{\mathrm{co}}\left(\bigcup_{t\in T}\left(\partial_{\varepsilon}f_t(x)\bigcup\mathrm{N}^{\varepsilon}_{\operatorname{dom} f_t}(x)\right)\right)\right]_{\infty}.$$

Proof. According to Theorem 4.7, for every $\varepsilon>0$ we get

$$\mathrm{N}_{\operatorname{dom} f}(x) \subset \left[\overline{\mathrm{co}}\left(\bigcup_{t\in T}\left(\partial_{\varepsilon}(\rho_{t,\varepsilon}f_t)(x)\bigcup\mathrm{N}^{\varepsilon}_{\operatorname{dom} f_t}(x)\right)\right)\right]_{\infty},$$

where $\rho_{t,\varepsilon}\le 1$, $t\in T$, is defined in (4.4). Then, by Lemma 2.35, we have

$$\begin{aligned}\mathrm{N}_{\operatorname{dom} f}(x) &\subset \left[\overline{\mathrm{co}}\left(\bigcup_{t\in T}\partial_{\varepsilon}(\rho_{t,\varepsilon}f_t)(x)\cup\mathrm{N}^{\varepsilon}_{\operatorname{dom} f_t}(x)\right)\right]_{\infty}\\ &\subset \left[\overline{\mathrm{co}}\left(\bigcup_{t\in T}\partial_{\varepsilon}f_t(x)\cup\mathrm{N}^{\varepsilon}_{\operatorname{dom} f_t}(x)\right)\right]_{\infty},\end{aligned}$$

and we conclude because the converse inclusion is straightforward. □

Theorem 4.7 can be written in a slight general form.

Corollary 4.11 *Consider a family* $\{f_t : X \to \overline{\mathbb{R}},\ t \in T\}$ *of convex functions such that (4.2) holds, and denote* $f := \sup_{t\in T} f_t$. *Given* $x \in f^{-1}(\mathbb{R})$ *such that* f *is lsc at* x, *we choose parameters* $\varepsilon_t \in [\rho_{t,\varepsilon}, 1]$, $t \in T$, *where* $\rho_{t,\varepsilon}$ *are defined in (4.4), such that*

$$\inf_{t\in T} \varepsilon_t f_t(x) > -\infty.$$

Then, for every $\varepsilon > 0$, *we have that*

$$\mathrm{N}_{\mathrm{dom}\, f}(x) = \left[\overline{\mathrm{co}}\left(\bigcup_{t\in T}\left(\partial_\varepsilon(\varepsilon_t f_t)(x) \bigcup \mathrm{N}^{\varepsilon}_{\mathrm{dom}\, f_t}(x)\right)\right)\right]_\infty .$$

Proof. According to Theorem 4.7, for every $\varepsilon > 0$ we have that

$$\mathrm{N}_{\mathrm{dom}\, f}(x) \subset \left[\overline{\mathrm{co}}\left(\bigcup_{t\in T}\left(\partial_\varepsilon(\rho_{t,\varepsilon} f_t)(x) \bigcup \mathrm{N}^{\varepsilon}_{\mathrm{dom}\, f_t}(x)\right)\right)\right]_\infty .$$

Then, since $\rho_{t,\varepsilon} \le \varepsilon_t$, Lemma 2.35 gives

$$\mathrm{N}_{\mathrm{dom}\, f}(x) \subset \left[\overline{\mathrm{co}}\left(\bigcup_{t\in T}\left(\partial_\varepsilon(\varepsilon_t f_t)(x) \bigcup \mathrm{N}^{\varepsilon}_{\mathrm{dom}\, f_t}(x)\right)\right)\right]_\infty ,$$

and the conclusion follows since the opposite inclusion holds straightforwardly. □

We illustrate Theorem 4.7 in the linear case.

Example 4.12 *(i) Let* T *be an arbitrary index set and consider the functions*

$$f := \sup_{t\in T} f_t := \langle a_t, \cdot\rangle - b_t, \quad a_t \in X^*,\ b_t \in \mathbb{R}.$$

Take $x \in \mathrm{dom}\, f$. *Then, since* $f_t \in \Gamma_0(X)$, *by Theorem 4.7 we have for every* $\varepsilon > 0$

$$\mathrm{N}_{\mathrm{dom}\, f}(x) = \left[\overline{\mathrm{co}}\left(A_\varepsilon(x) \cup B_\varepsilon(x)\right)\right]_\infty ,$$

where

$$A_\varepsilon(x) := \bigcup_{t\in T_\varepsilon(x)} \partial_\varepsilon f_t(x) = \bigcup_{t\in T_\varepsilon(x)} \{a_t\}, \quad B_\varepsilon(x) := \bigcup_{t\in T\setminus T_\varepsilon(x)} \partial_\varepsilon(\rho_{t,\varepsilon} f_t)(x) = \bigcup_{t\in T\setminus T_\varepsilon(x)} \{\rho_{t,\varepsilon} a_t\}.$$

Hence, for the particular case when $x = 0$ *and* $f(0) = 0$, *we get*

$$\begin{aligned} A_\varepsilon(0) &= \bigcup_{t\in T} \{a_t : b_t \le \varepsilon\}, \\ B_\varepsilon(0) &= \bigcup_{t\in T} \left\{\frac{\varepsilon}{2b_t - \varepsilon} \cdot a_t : b_t > \varepsilon\right\}. \end{aligned}$$

If $\inf_{t\in T}(-b_t) > -\infty$, *then*

$$\mathrm{N}_{\mathrm{dom}\, f}(0) = \left[\overline{\mathrm{co}}\left(\cup_{t\in T}\{a_t\}\right)\right]_\infty .$$

(ii) *Suppose that* $f := \sigma_D := \sup\{\langle d^*, \cdot\rangle : d^* \in D\}$ *for some nonempty set* $D \subset X^*$. *Then for every* $x \in \operatorname{dom} f$ *and for every* $\varepsilon > 0$ *we have that*

$$\mathrm{N}_{\operatorname{dom} f}(x) = \left[\overline{\mathrm{co}}\,(A_\varepsilon(x) \cup B_\varepsilon(x))\right]_\infty,$$

where

$$\begin{aligned} A_\varepsilon(x) &= \{d^* \in D : \ \langle d^*, x\rangle \geq \sigma_D(x) - \varepsilon\}, \\ B_\varepsilon(x) &= \left\{\frac{\varepsilon d^*}{2\sigma_D(x) - \langle 2d^*, x\rangle + \varepsilon} : \langle d^*, x\rangle < \sigma_D(x) - \varepsilon\right\}. \end{aligned}$$

In particular, if $\sigma_D(-x) < +\infty$, *then*

$$\mathrm{N}_{\operatorname{dom} f}(x) = \left[\overline{\mathrm{co}}\,(D)\right]_\infty = (\operatorname{dom} \sigma_D)^-,$$

where the last equality comes from Lemma 2.13. In other words, $\mathrm{N}_{\operatorname{dom} f}(x)$ *is the same for all* x *such that* $\sigma_D(-x) < +\infty$. *Indeed, this can also be seen from the fact that for every such* x *we have*

$$(\mathrm{N}_{\operatorname{dom} f}(x))^- = \overline{\mathrm{cone}}\,(\operatorname{dom} \sigma_D - x) = \overline{\mathrm{cone}}\,(\operatorname{dom} \sigma_D).$$

4.3 Characterization of the normal cone in the compact-continuous setting

In this section we specify the results of the previous section to the compact-continuous settings. We give a family of convex function $f_t, t \in T$, such that the index set T is compact and the mappings $t \mapsto f_t(x)$, $t \in T$, are usc for all $x \in X$.

A consequence of this continuous.-compact setting is that (see [12, Lemma 4])

$$\operatorname{dom} f = \cap_{t\in T} \operatorname{dom} f_t, \tag{4.26}$$

and, for every $x \in \operatorname{dom} f$,

$$\mathbb{R}_+(\operatorname{dom} f - x) = \cap_{t\in T} \mathbb{R}_+(\operatorname{dom} f_t - x). \tag{4.27}$$

These equalities do not necessarily hold outside the compact framework.

As before, we need to use an appropriate family of parameters:

$$\overline{\rho}_{t,\varepsilon} := \begin{cases} \frac{\varepsilon}{-2f_t(x)+2f(x)+\varepsilon}, & \text{if } t \in T \setminus T(x), \\ 1, & \text{if } t \in T(x). \end{cases} \tag{4.28}$$

Let us check that these new parameters also satisfies the condition

$$\inf_{t\in T} \overline{\rho}_{t,\varepsilon} f_t(x) > -\infty. \tag{4.29}$$

If $t \in T(x)$, then $\overline{\rho}_{t,\varepsilon} f_t(x) = f_t(x) = f(x) > -\infty$. Otherwise, provided that $f(x) = 0$, for all $t \in T \setminus T(x)$ we obtain

$$\overline{\rho}_{t,\varepsilon} f_t(x) = \frac{\varepsilon f_t(x)}{-2f_t(x) + \varepsilon} > -\varepsilon.$$

More generally, when $f(x) \neq 0$, by applying the last inequality to the functions $f_t - f(x)$ we get

$$\overline{\rho}_{t,\varepsilon}(f_t - f(x))(x) = \frac{\varepsilon(f_t(x) - f(x))}{-2f_t(x) + 2f(x) + \varepsilon} > -\varepsilon > -\infty$$

and, hence,

$$\overline{\rho}_{t,\varepsilon} f_t(x) > \overline{\rho}_{t,\varepsilon} f(x) - \varepsilon,$$

showing that (4.29) holds.

In order to characterize the normal cone to the effective domain of function f in the current compact-continuous setting, we start by considering the case where the convex functions f_t's are lsc. The first characterization in the result below is given [12, Theorem 5].

Proposition 4.13 *Given proper lsc convex functions $f_t : X \to \overline{\mathbb{R}}$, $t \in T$, we denote $f := \sup_{t\in T} f_t$. Assume that T Hausdorff compact and the mappings $t \mapsto f_t(x)$ are usc for each $x \in X$. Then, for every $x \in \operatorname{dom} f$, we have that*

$$\mathrm{N}_{\operatorname{dom} f}(x) = \left[\overline{\operatorname{co}}\left(\bigcup\nolimits_{t\in T} \partial_\varepsilon(\overline{\rho}_{t,\varepsilon} f_t)(x)\right)\right]_\infty, \quad \textit{for all } \varepsilon > 0,$$

where $(\overline{\rho}_{t,\varepsilon})_{t\in T}$ are defined in (4.28). Moreover, if $\{(f_t)'(x;\cdot),\ t \in T\} \subset \Gamma_0(X)$, then

$$\mathrm{N}_{\operatorname{dom} f}(x) = \left[\overline{\operatorname{co}}\left(\bigcup\nolimits_{t\in T} \partial(\overline{\rho}_t f_t)(x)\right)\right]_\infty.$$

Proof. For the first equality we refer to [12, Theorem 5]. To prove the second statement, we suppose that $\{(f_t)'(x;\cdot),\ t \in T\} \subset \Gamma_0(X)$. Then, by [60, Corollary 2.4.15], we have that $\partial(\overline{\rho}_{t,\varepsilon} f_t)(x) \neq \emptyset$ and $(\overline{\rho}_{t,\varepsilon} f_t)'(x;\cdot) = \sigma_{\partial(\overline{\rho}_{t,\varepsilon} f_t)(x)}$ for all $t \in T$ and, so,

$$\begin{aligned}
(\mathrm{N}_{\operatorname{dom} f}(x))^- &= \operatorname{cl}(\mathbb{R}_+(\operatorname{dom} f - x)) \\
&= \operatorname{cl}(\bigcap_{t\in T} \mathbb{R}_+(\operatorname{dom} f_t - x)) \\
&= \operatorname{cl}(\bigcap_{t\in T} \mathbb{R}_+(\operatorname{dom}(\overline{\rho}_{t,\varepsilon} f_t) - x)) \\
&= \operatorname{cl}(\bigcap_{t\in T} \operatorname{dom}((\overline{\rho}_{t,\varepsilon} f_t)'(x;\cdot))) \\
&\supset \operatorname{cl}\left(\operatorname{dom}(\sup_{t\in T}(\overline{\rho}_{t,\varepsilon} f_t)'(x;\cdot))\right) \\
&= \operatorname{cl}\left(\operatorname{dom} \sigma_{\cup_{t\in T}\partial(\overline{\rho}_{t,\varepsilon} f_t)(x)}(\cdot)\right) = \left(\left[\overline{\operatorname{co}}\left(\cup_{t\in T}\partial(\overline{\rho}_{t,\varepsilon} f_t)(x)\right)\right]_\infty\right)^-,
\end{aligned}$$

that is,

$$\mathrm{N}_{\operatorname{dom} f}(x) \subset \left[\overline{\operatorname{co}}\left(\bigcup\nolimits_{t\in T} \partial(\overline{\rho}_{t,\varepsilon} f_t)(x)\right)\right]_\infty.$$

The opposite inclusion is straightforward. □

Now, we establish the general characterization of the normal cone to $\operatorname{dom} f$ in the current compact-continuous setting.

Theorem 4.14 *Consider a family of convex functions* $\{f_t : X \to \overline{\mathbb{R}},\ t \in T\}$, *and* $f = \sup_{t\in T} f_t$, *where* T *is a Hausdorff compact and the mappings* $t \mapsto f_t(x)$ *are usc for each* $x \in X$. *If condition (4.2) holds, that is,*

$$\operatorname{cl} f = \sup_{t\in T}(\operatorname{cl} f_t),$$

then, for every $x \in \operatorname{dom} f$ *such that* $f(x) = (\operatorname{cl} f)(x) \in \mathbb{R}$, *we have*

$$\mathrm{N}_{\operatorname{dom} f}(x) = \left[\overline{\operatorname{co}}\left(\left(\bigcup_{t\in T_0} \partial_\varepsilon(\rho_{t,\varepsilon} f_t)(x)\right) \bigcup \left(\bigcup_{t\in T\setminus T_0} \mathrm{N}^{\varepsilon}_{\operatorname{dom} f_t}(x)\right)\right)\right]_\infty, \quad \text{for all } \varepsilon > 0,$$

where $T_0 := \{t \in T : \operatorname{cl} f_t$ *is proper*$\}$.

Proof. We fix $x \in \operatorname{dom} f$ and $\varepsilon > 0$. Then the proof is analogous to the one given in Theorem 4.7. In fact, assuming without loss of generality that $f(x) = 0$, we first suppose that

$$\{(\operatorname{cl} f_t), t \in T\} \subset \Gamma_0(X).$$

Then, since $x \in \operatorname{dom} f \subset \operatorname{dom}(\operatorname{cl} f)$, we have that

$$\mathrm{N}_{\operatorname{dom} f}(x) = \mathrm{N}_{\operatorname{cl}(\operatorname{dom} f)}(x) = \mathrm{N}_{\operatorname{cl}(\operatorname{dom}(\operatorname{cl} f))}(x) = \mathrm{N}_{\operatorname{dom}(\operatorname{cl} f)}(x),$$

and Proposition 4.13 applied to the family $\{(\operatorname{cl} f_t), t \in T\}$ yields

$$\mathrm{N}_{\operatorname{dom} f}(x) = \mathrm{N}_{\operatorname{dom}(\operatorname{cl} f)}(x) = \left[\overline{\operatorname{co}}\left(\bigcup_{t\in T} \partial_\varepsilon(\bar{\tilde{\rho}}_{t,\varepsilon}(\operatorname{cl} f_t))(x)\right)\right]_\infty, \tag{4.30}$$

where

$$\bar{\tilde{\rho}}_{t,\varepsilon} := \begin{cases} \frac{\varepsilon}{-2(\operatorname{cl} f_t)(x)+\varepsilon}, & \text{if } t \in T \setminus \widetilde{T}(x), \\ 1, & \text{if } t \in \widetilde{T}(x), \end{cases}$$

and $\widetilde{T}(x) := \{t \in T : (\operatorname{cl} f_t)(x) = 0\} = T(x)$ (since $f(x) = \operatorname{cl} f)(x) = 0$). Moreover, for all $t \in T \setminus T(x)$, we have that

$$\bar{\tilde{\rho}}_{t,\varepsilon} = \frac{\varepsilon}{-2f_t(x) + \varepsilon + 2f_t(x) - 2(\operatorname{cl} f_t)(x)} \leq \frac{\varepsilon}{-2f_t(x) + \varepsilon} = \bar{\rho}_{t,\varepsilon}. \tag{4.31}$$

Also, for each $t \in T(x)$ we have that $(\operatorname{cl} f_t)(x) = 0 = f_t(x)$ and, so,

$$\partial_\varepsilon(\bar{\tilde{\rho}}_{t,\varepsilon}(\operatorname{cl} f_t))(x) = \partial_\varepsilon(\operatorname{cl} f_t)(x) \subset \partial_\varepsilon f_t(x).$$

At the same time, for each $t \in T \setminus T(x)$ we have that $\bar{\tilde{\rho}}_{t,\varepsilon}(\operatorname{cl} f_t)(x) > -\varepsilon \geq \bar{\tilde{\rho}}_{t,\varepsilon} f_t(x) - \varepsilon$ (because $f_t(x) \leq 0$ and $\bar{\tilde{\rho}}_{t,\varepsilon} f_t(x) \leq 0$) and, so,

$$\partial_\varepsilon(\bar{\tilde{\rho}}_{t,\varepsilon}(\operatorname{cl} f_t))(x) = \partial_\varepsilon(\operatorname{cl}(\bar{\tilde{\rho}}_{t,\varepsilon} f_t))(x) \subset \partial_{2\varepsilon}(\bar{\tilde{\rho}}_{t,\varepsilon} f_t)(x).$$

Thus, by (4.30),

$$\mathrm{N}_{\operatorname{dom} f}(x) \subset \left[\overline{\operatorname{co}}\left(\bigcup_{t\in T} \partial_{2\varepsilon}(\bar{\tilde{\rho}}_{t,\varepsilon} f_t)(x)\right)\right]_\infty.$$

Moreover, since that $\bar{\tilde{\rho}}_{t,\varepsilon} \leq \bar{\rho}_{t,\varepsilon}$ for all $t \in T$, by Lemma 2.35 we obtain

$$\mathrm{N}_{\operatorname{dom} f}(x) = \left[\overline{\operatorname{co}}\left(\bigcup_{t\in T} \partial_{2\varepsilon}(\bar{\rho}_{t,\varepsilon} f_t)(x)\right)\right]_\infty$$

(the inclusion "$\supset$" in the last relation is straightforward; see the proof in Proposition 4.13). The proof of the general case when not all the $\operatorname{cl} f_t$,'s are necessarily proper is done using similar arguments as those of the second part of the proof of Theorem 4.7. $\square$

More generally, we have the following result which relaxes the compact-continuous setting.

Theorem 4.15 *Consider a family* $\{f_t : X \to \overline{\mathbb{R}},\ t \in T\} \subset \Gamma_0(X)$ *and* $f := \sup_{t\in T} f_t$, *where* T *an arbitrary index set. Assume that, for each sequence* $(t_n)_n \subset T$, *there exists some finite set* $J \subset T$ *such that*

$$\limsup_n f_{t_n} \leq f_J := \max\{f_t, t \in J\}.$$

Then, for every $x \in \operatorname{dom} f$,

$$\mathrm{N}_{\operatorname{dom} f}(x) = \left[\overline{\operatorname{co}}\left(\bigcup_{t\in T}\partial_\varepsilon(\overline{\rho}_{t,\varepsilon} f_t)(x)\right)\right]_\infty, \text{ for all } \varepsilon > 0.$$

Proof. Take $u \in \mathrm{N}_{\operatorname{dom} f}(x)$. Then, by [34, Proposition 3], for each $L \in \mathcal{F}(x)$ there exists a countable set $\{t_1, \cdots, t_k, \cdots\} \subset T$ such that

$$u \in \mathrm{N}_{L\cap \operatorname{dom}(\sup_n f_{t_n})}(x).$$

By the current assumption we choose a finite set $J \subset T$ such that $\limsup f_{t_n} \leq f_J$; the family $\{f_{t_n},\ n \geq 1,\ f_t,\ t \in J\}$ satisfies

$$\begin{aligned}
\operatorname{dom}(\sup\{\overline{\rho}_{t_n,\varepsilon} f_{t_n},\ n \geq 1,\ \overline{\rho}_{t,\varepsilon} f_t,\ t \in J\}) &\subset \left(\bigcap_{n\geq 1} \operatorname{dom}(\overline{\rho}_{t_n,\varepsilon} f_{t_n})\right) \bigcap \left(\bigcap_{t\in J} \operatorname{dom}(\overline{\rho}_{t,\varepsilon} f_t)\right) \\
&= \left(\bigcap_{n\geq 1} \operatorname{dom}(f_{t_n})\right) \bigcap \left(\bigcap_{t\in J} \operatorname{dom}(f_t)\right) \\
&= \left(\bigcap_{n\geq 1} \operatorname{dom}(f_{t_n})\right) \bigcap \operatorname{dom}(f_J),
\end{aligned}$$

where the last equality comes from (4.26). Take $z \in \left(\bigcap_{n\geq 1} \operatorname{dom}(f_{t_n})\right) \cap \operatorname{dom}(f_J)$. Then, since $\limsup_n f_{t_n}(z) \leq f_J(z) < +\infty$, we obtain that $z \in \operatorname{dom}(\sup_n f_{t_n})$ and this implies

$$\left(\bigcap_{n\geq 1} \operatorname{dom}(f_{t_n})\right) \bigcap \operatorname{dom}(f_J) \subset \operatorname{dom}(\sup_n f_{t_n}).$$

Therefore, applying Proposition 4.13 to the countable family $\{\overline{\rho}_{t_n,\varepsilon} f_{t_n},\ n \geq 1,\ \overline{\rho}_{t,\varepsilon} f_t,\ t \in J,\ \mathrm{I}_L\}$, we have for each $\varepsilon > 0$

$$\begin{aligned}
u \in \mathrm{N}_{L\cap \operatorname{dom}(\sup_n f_{t_n})}(x) &\subset \mathrm{N}_{\operatorname{dom}(\sup\{\overline{\rho}_{t_n,\varepsilon} f_{t_n},\ n\geq 1,\ \overline{\rho}_{t,\varepsilon} f_t,\ t\in J,\ \mathrm{I}_L\})}(x) \\
&= \left[\overline{\operatorname{co}}\left(\left(\bigcup_{n\geq 1}\partial_\varepsilon(\overline{\rho}_{t_n,\varepsilon} f_{t_n})(x)\right) \bigcup \left(\bigcup_{t\in J}\partial_\varepsilon(\overline{\rho}_{t,\varepsilon} f_t)(x)\right) \bigcup L^\perp\right)\right]_\infty \\
&\subset \left[\overline{\operatorname{co}}\left(\bigcup_{t\in T}\partial_\varepsilon(\overline{\rho}_{t,\varepsilon} f_t)(x) + L^\perp\right)\right]_\infty.
\end{aligned}$$

Consequently, the desired inclusion follows by intersecting over $L \in \mathcal{F}(x)$ and applying Lemmas 2.6 and 2.8. □

Chapter 5

Subdifferential calculus for pointwise suprema

5.1 Introduction

The supremum function is a standard tool in convex analysis and optimization theory, playing a fundamental role in establishing KKT and Fritz-John optimality conditions, subdifferential calculus rules, conjugation analysis and duality theory, minimax problems, and so on. Our main goal in this chapter is to achieve a characterization for the subdifferential of the pointwise supremum function, ideally an expression that directly depends exclusively on the data functions with minimal requirements, avoiding the use of finite-dimensional sections of the associated effective domain of this supremum function.

Formally, we deal with the pointwise supremum function, $f := \sup_{t\in T} f_t$, of a family of extended real-valued convex functions defined on an lcs X, $f_t : X \to \mathbb{R}_\infty$, and indexed in an arbitrary set T. There are many contributions to this topic; we refer, for example, to Brøndsted & Rockafellar [8], Valadier [57], Brøndsted [7], Volle [58] and Ioffe [37] among many other contributions. For more recent works, we refer among many others to Hantoute, López & Zălinescu [35], López & Volle [43], Mordukhovich & Nghia [46], Correa, Hantoute & López [15] (see, also, [16], [12]) and Hantoute & López [33] (also, [34]).

A general formula of $\partial f(x)$ is given in [35, Theorem 4] under the closure condition, $\operatorname{cl} f = \sup_{t\in T} \operatorname{cl} f_t$, showing that for all $x \operatorname{dom} f$

$$\partial f(x) = \bigcap_{L\in\mathcal{F}(x),\ \varepsilon>0} \overline{\operatorname{co}} \left(\bigcup_{t\in T_\varepsilon(x)} \partial_\varepsilon f_t(x) + \mathrm{N}_{L\cap \operatorname{dom} f}(x) \right), \tag{5.1}$$

where $\mathcal{F}(x)$ is a set of finite-dimensional linear subspaces that contain the point x, and $T_\varepsilon(x) := \{t \in T : f_t(x) \geq f(x) - \varepsilon\}$ is the ε-active index set.

In the so-called compact-continuous setting (T a Hausdorff compact set and the mappings $t \mapsto f_t(x)$, $x \in X$, are usc), the characterization involves the active functions ($t \in T_0(x)$) as shown in [16, Theorem 3.8]:

$$\partial f(x) = \bigcap_{L \in \mathcal{F}(x),\ \varepsilon>0} \overline{\mathrm{co}} \left(\bigcup_{t \in T(x)} \partial_\varepsilon f_t(x) + \mathrm{N}_{L \cap \mathrm{dom}\, f}(x) \right). \tag{5.2}$$

The problem is then whether the terms using the effective domain of f can be removed. For instance, in [34, Theorem 5] where all the convex functions f_t's are assumed to be lsc and proper, the author provided the following characterization for every $x \in \mathrm{dom}\, f$,

$$\partial f(x) = \bigcap_{\varepsilon>0} \overline{\mathrm{co}} \left(\bigcup_{t \in T_\varepsilon(x)} \partial_\varepsilon f_t(x) + \{0, \varepsilon\} \bigcup_{J \in \mathcal{T}_\varepsilon(x)} \partial_\varepsilon f_J(x) \right),$$

where $\mathcal{T}_\varepsilon(x) := \{J \in \mathcal{T} : f_J(x) \geq f(x) - \varepsilon\}$, $\mathcal{T} = \{J \subset T : |J| < +\infty\}$ and $f_J := \max\{f_t : t \in J\}$. The normal cone used above is replaced with the union of subdifferentials of maximum functions, each one involving a finite set of indices. This last expression does not distinguish between the role of ε-active functions and the others, as the condition $f_J(x) \geq f(x) - \varepsilon$ may involve both kinds of functions.

At the same time, the work in [12], which deals with the compact-continuous setting, gives a more explicit representation of the subdifferential of f without using the extra maxima f_J but, instead, it uses parameters for non-active functions. There, also assuming that the data are from $\Gamma_0(X)$, it is shown that for all $x \in \mathrm{dom}\, f$

$$\partial f(x) = \bigcap_{\varepsilon>0} \overline{\mathrm{co}} \left(\bigcup_{t \in T(x)} \partial_\varepsilon f_t(x) + \{0, \varepsilon\} \bigcup_{t \in T \setminus T(x)} \partial_\varepsilon (\rho_t f_t)(x) \right).$$

Then our objective is to extend the last formula to the general framework outside the compact-continuous setting. To this aim, the characterizations of the normal cone given in the previous chapter will be of a great help. Indeed, we shall prove that, for all $x \in \mathrm{dom}\, f$,

$$\partial f(x) = \bigcap_{\varepsilon>0} \overline{\mathrm{co}} \left(\bigcup_{t \in T_\varepsilon(x)} \partial_\varepsilon f_t(x) + \left(\varepsilon \bigcup_{t \in T_0 \setminus T_\varepsilon(x)} \partial_\varepsilon (\rho_{t,\varepsilon} f_t)(x) \right) \cup \left(\bigcup_{t \in T \setminus T_0} \mathrm{N}^\varepsilon_{\mathrm{dom}\, f_t}(x) \cup \{\theta\} \right) \right),$$

where $T_0 := \{t \in T : \mathrm{cl}\, f_t \text{ is proper}\}$. The main feature of this formula is that the role of ε-active and non ε-active data appears explicitly.

This chapter is structured as follows: The new characterization of the subdifferential of the supremum is given in Section 5.2. For instance, Theorem 5.1 deals with the general framework where some appropriate parameters are assigned to non-ε-active functions. The specification of the last result to the compact-continuous setting is given in Theorem 5.3. Section 5.3 establishes a reduction process that is involved within the characterization of the subdifferential of the supremum. In Section 5.4 we present Danskin's Theorem and some extensions of it and. We end the Chapter with Section 5.5 where we apply the previous theorems to provide some optimality conditions for convex semi-infinite programming problems.

5.2 Subdifferential of the supremum function

As the following theorem shows, the subdifferential of f is constructed upon the following three blocks give for $x \in \operatorname{dom} f$ and $\varepsilon > 0$, as

$$A_\varepsilon(x) := \bigcup \{\partial_\varepsilon f_t(x) : t \in T_\varepsilon(x)\},$$

$$B_\varepsilon(x) := \bigcup \{\partial_\varepsilon (\rho_{t,\varepsilon} f_t)(x) : t \in T_0 \backslash T_\varepsilon(x)\}, \text{ with } \rho_{t,\varepsilon} \in]0,1],$$

$$C_\varepsilon(x) := \bigcup_{t \in T \backslash T_0} \mathrm{N}^\varepsilon_{\operatorname{dom} f_t}(x) \cup \{\theta\},$$

where

$$T_\varepsilon(x) := \{t \in T_0 : f_t(x) \geq f(x) - \varepsilon\},$$

with

$$T_0 := \{t \in T : \operatorname{cl} f_t \text{ is proper}\}.$$

Remember that by convention we have

$$\bigcup_{t \in \emptyset} \mathrm{N}^\varepsilon_{\operatorname{dom} f_t}(x) = \emptyset.$$

Moreover, the parameters $\rho_{t,\varepsilon}$, are define by

$$\rho_{t,\varepsilon} := \begin{cases} \frac{-\varepsilon}{2f_t(x) - 2f(x) + \varepsilon}, & \text{if } t \in T_0 \setminus T_\varepsilon(x), \\ 1, & \text{if } t \in T_\varepsilon(x). \end{cases} \tag{5.3}$$

and satisfies, as we checked before,

$$\inf_{t \in T} \rho_{t,\varepsilon} f_t(x) > -\infty.$$

The first block $A_\varepsilon(x)$ corresponds to the almost active indices, while $B_\varepsilon(x)$ involves the rest of functions but multiplied by a parameter $\rho_{t,\varepsilon} \in]0,1]$. The set $C_\varepsilon(x)$ represents somewhat those functions which are not represented in the first two blocks, and whose closure is not proper.

Theorem 5.1 *Let $f_t : X \to \overline{\mathbb{R}}$, $t \in T$, be a family of convex functions and denote $f = \sup_{t \in T} f_t$. We assume that*

$$\operatorname{cl} f = \sup_{t \in T} (\operatorname{cl} f_t).$$

Then, for every $x \in \operatorname{dom} f$, we have that

$$\partial f(x) = \bigcap_{\varepsilon > 0} \overline{\operatorname{co}} \left(A_\varepsilon(x) + (\varepsilon B_\varepsilon(x) \cup C_\varepsilon(x)) \right). \tag{5.4}$$

Proof. Fix $x \in \operatorname{dom} f$, $\varepsilon > 0$, $U \in \mathcal{N}$ and pick $L \in \mathcal{F}(x)$ such that $L^\perp \subseteq U$. We start by proving the direct inclusion "$\subset$" in the nontrivial case when $\partial f(x) \neq \emptyset$. Hence, we may assume that $f(x) = \operatorname{cl} f(x) = 0$ and, so, we can apply Theorem 4.7 to the family $\{f_t,\ t \in T; \mathrm{I}_L\}$ to obtain

$$\mathrm{N}_{L\cap \operatorname{dom} f}(x) = \left[\overline{\operatorname{co}} \left(\left(\bigcup_{t\in T_0} \partial_\varepsilon(\rho_{t,\varepsilon} f_t)(x) \right) \bigcup L^{\perp} \bigcup \left(\bigcup_{t\in T\setminus T_0} \mathrm{N}^{\varepsilon}_{\operatorname{dom} f_t}(x) \right) \right) \right]_\infty ,$$

that is, using Lemmas 2.6 and 2.7,

$$\begin{aligned} \mathrm{N}_{L\cap \operatorname{dom} f}(x) &= \left[\overline{\operatorname{co}} \left(A_\varepsilon(x) \cup (B_\varepsilon(x) \cup C_\varepsilon(x)) \cup L^{\perp} \right) \right]_\infty \\ &= \left[\overline{\operatorname{co}} \left(A_\varepsilon(x) + (\varepsilon B_\varepsilon(x) \cup C_\varepsilon(x)) + L^{\perp} \right) \right]_\infty . \end{aligned}$$

Therefore, since every $t \in T$ such that $\partial_\varepsilon f_t(x) \neq \emptyset$ belongs to $T_\varepsilon(x)$, (5.1) entails

$$\begin{aligned} \partial f(x) &\subset \overline{\operatorname{co}} \left(\left(\bigcup_{t\in T_\varepsilon(x)} \partial_\varepsilon f_t(x) \right) + \left[\overline{\operatorname{co}} \left(A_\varepsilon(x) + (\varepsilon B_\varepsilon(x) \cup C_\varepsilon(x)) + L^{\perp} \right) \right]_\infty \right) \\ &= \overline{\operatorname{co}} \left(A_\varepsilon(x) + \left[\overline{\operatorname{co}} \left(A_\varepsilon(x) + (\varepsilon B_\varepsilon(x) \cup C_\varepsilon(x)) + L^{\perp} \right) \right]_\infty \right) \\ &\subset \overline{\operatorname{co}} \left(A_\varepsilon(x) + (\varepsilon B_\varepsilon(x) \cup C_\varepsilon(x)) + L^{\perp} \right) \\ &\subset \operatorname{co} \left(A_\varepsilon(x) + (\varepsilon B_\varepsilon(x) \cup C_\varepsilon(x)) \right) + 2\, U, \end{aligned}$$

and the desired inclusion follows by taking the intersection first over $U \in \mathcal{N}$ and next over $\varepsilon > 0$.

To prove the opposite inclusion, we assume without loss of generality that $f(x) = 0$. We take x^* in the right-hand side of equality (5.4). Then, for each $\varepsilon > 0$, there are nets

$$(\alpha_{j,1}, \cdots, \alpha_{j,m_j}) \in \Delta^+_{m_j},\ m_j \geq 1,$$

and, for $i = 1, \cdots, m_j$,

$$x^*_{j,i} := u^*_{j,i} + \beta_{j,i} \varepsilon w^*_{j,i} + \gamma_{j,i} v^*_{j,i},\ \text{ where } u^*_{j,i} \in \partial_\varepsilon f_{t_{j,i}}(x), w^*_{j,i} \in \partial_\varepsilon f_{s_{j,i}}(x), v^*_{j,i} \in C_\varepsilon(x),$$

where

$$t_{j,i} \in T_\varepsilon(x),\ s_{j,i} \in T_0 \setminus T_\varepsilon(x),\ \beta_{j,i},\ \gamma_{j,i} \in \{0, 1\},$$

such that

$$x^* = \lim_j \sum_{i=1}^{m_j} \alpha_{j,i} \left(u^*_{j,i} + \varepsilon \beta_{j,i} w^*_{j,i} + \gamma_{j,i} v^*_{j,i} \right).$$

Then, for all $i,\ j$ as above and $y \in \operatorname{dom} f$ we have

$$\langle u^*_{j,i}, y - x \rangle \leq f_{t_{j,i}}(y) - f_{t_{j,i}}(x) + \varepsilon \leq f(y) + 2\varepsilon,$$

implying that $u^*_{j,i} \in \partial_{2\varepsilon} f(x)$. Similarly, since $0 < \rho_{t_{j,i}} \leq 1$ and $\inf \rho_{t,\varepsilon} f_t(x) > -\varepsilon$ (see (4.5)), we have that

$$\begin{aligned} \langle w^*_{j,i}, y - x \rangle &\leq \rho_{t_{j,i},\varepsilon} f_{t_{j,i}}(y) - \rho_{t_{j,i},\varepsilon} f_{t_{j,i}}(x) + \varepsilon \\ &\leq f^+(y) - \inf \rho_{t,\varepsilon} f_t(x) + \varepsilon \leq f^+(y) + 2\varepsilon, \end{aligned}$$

that is, $w_{j,i}^* \in \partial_{2\varepsilon}$ Thus, since each $v_{j,i}^*$ satisfies $\langle v_{j,i}^*, y-x\rangle \leq \varepsilon$, we conclude that

$$\begin{aligned}
\langle x^*, y-x\rangle &= \lim_j \sum_{i=1}^{m_j} \alpha_{j,i}\left(\langle u_{j,i}^*, y-x\rangle + \beta_{j,i}\varepsilon\langle w_{j,i}^*, y-x\rangle + \gamma_{j,i}\langle v_{j,i}^*, y-x\rangle\right)\\
&\leq \limsup_j \sum_{i=1}^{m_j} \alpha_{j,i}\left((f(y)+2\varepsilon) + \beta_{j,i}\varepsilon(f^+(y)+2\varepsilon) + \gamma_{j,i}\varepsilon\right)\\
&\leq f(y) + 3\varepsilon + \varepsilon(f^+(y)+2\varepsilon).
\end{aligned}$$

Finally, as $\varepsilon \downarrow 0$ we deduce that $x^* \in \partial f(x)$, as we wanted to prove. □

The next result extends Corollary 12 in [12] to possibly non-compact-continuous settings.

Corollary 5.2 *Assume in Theorem 5.1 that, additionally, all the functions f_t are in $\Gamma_0(X)$. Then we have*

$$\partial f(x) = \bigcap_{\varepsilon>0} \overline{\text{co}}\left[A_\varepsilon(x) + (\varepsilon B_\varepsilon(x) \cup \{0\})\right].$$

In particular, if f attains its minimum at x, then

$$\partial f(x) = \bigcap_{\varepsilon>0} \overline{\text{co}}\left[A_\varepsilon(x) \cup \varepsilon B_\varepsilon(x)\right].$$

Proof. The first statement follows immediately from Theorem 5.1 taking into account that in the current case we have $T\backslash T_0 = \emptyset$ and so, by our convention,

$$C_\varepsilon(x) = \bigcup_{t\in\ \emptyset} \mathrm{N}_{\mathrm{dom}\, f_t}^{\varepsilon}(x) \cup \{\theta\} = \{\theta\}.$$

To prove the second statement, as in the proof of Theorem 5.1 we show that

$$\mathrm{N}_{L\cap \mathrm{dom}\, f}(x) = \left[\overline{\text{co}}\left(A_\varepsilon(x) \cup \varepsilon B_\varepsilon(x) + L^\perp\right)\right]_\infty,$$

and (5.1) entails

$$\partial f(x) \subset \bigcap_{\varepsilon>0} \overline{\text{co}}\left(A_\varepsilon(x) \cup \varepsilon B_\varepsilon(x)\right).$$

Conversely, for every $\varepsilon > 0$, we know that (see the proof of Theorem 5.1)

$$\overline{\text{co}}\left(A_\varepsilon(x) \cup \varepsilon B_\varepsilon(x)\right) \subset \overline{\text{co}}\left(\partial_{2\varepsilon} f(x) \cup \varepsilon\partial_{2\varepsilon} f^+(x)\right).$$

More specifically, since f is assumed to attain its minimum at x, by [12, Lemma 1] we know that

$$\bigcap_{\varepsilon>0} \overline{\text{co}}\left(\partial_{2\varepsilon} f(x) \cup \varepsilon\partial_{2\varepsilon} f^+(x)\right) = \partial f(x),$$

and the desired result follows from the last inclusion above. □

We turn now to the compact-continuous setting. In this case, the active index set suffices to describe the subdifferential of the supremum. Given $x \in x\,\mathrm{dom}\,.f$, we denote

$$T(x) := \{t \in T_0 : f_t(x) = f(x)\},$$

together with the sets

$$\begin{aligned}\tilde{A}_\varepsilon(x) &:= \bigcup\{\partial_\varepsilon f_t(x) : t \in T(x)\},\\ \tilde{B}_\varepsilon(x) &:= \bigcup\{\partial_\varepsilon(\overline{\rho}_{t,\varepsilon} f_t)(x) : t \in T_0 \backslash T(x)\},\end{aligned}$$

where

$$\overline{\rho}_{t,\varepsilon} := \begin{cases} \frac{\varepsilon}{-2f_t(x)+2f(x)+\varepsilon}, & \text{if } t \in T_0 \setminus T(x), \\ 1, & \text{if } t \in T(x). \end{cases} \tag{5.5}$$

Theorem 5.3 *Let $f_t : X \to \overline{\mathbb{R}}$, $t \in T$, be a family of convex functions and denote $f := \sup_{t \in T} f_t$. Assume that T is Hausdorff compact and the mappings $t \mapsto f_t(x)$, $x \in X$, are usc. If condition (4.2) holds, that is,* $\mathrm{cl}\, f = \sup_{t\in T} \mathrm{cl}\, f_t$, *then for every $x \in \mathrm{dom}\, f$ we have that*

$$\partial f(x) = \bigcap_{\varepsilon>0} \overline{\mathrm{co}} \left(\tilde{A}_\varepsilon(x) + (\varepsilon \tilde{B}_\varepsilon(x) \cup C_\varepsilon(x)) \right).$$

Proof. To prove "$\subset$" we may suppose that $\partial f(x) \neq \emptyset$; for otherwise, the proof is trivial. Hence, $f(x) = \mathrm{cl}\, f(x) \in \mathbb{R}$ and the set $T(x)$ is nonempty. Moreover, by (5.2), for all $\varepsilon > 0$ and all $L \in \mathcal{F}(x)$ we know that

$$\partial f(x) \subset \overline{\mathrm{co}} \left(\cup_{t\in T(x)} \partial_\varepsilon f_t(x) + \mathrm{N}_{L \cap \mathrm{dom}\, f}(x) \right),$$

where, due to Theorem 4.14,

$$\mathrm{N}_{L\cap \mathrm{dom}\, f}(x) \;=\; \left[\overline{\mathrm{co}} \left(\left(\cup_{t \in T_0} \partial_\varepsilon(\overline{\rho}_{t,\varepsilon} f_t)(x) \right) \cup \left(\cup_{t \in T \backslash T_0} \mathrm{N}^\varepsilon_{\mathrm{dom}\, f_t}(x) \right) \cup L^\perp \right) \right]_\infty. \tag{5.6}$$

In other words, by rearranging these last terms and using Lemmas 2.6 and 2.7, we have

$$\begin{aligned} \mathrm{N}_{L\cap \mathrm{dom}\, f}(x) &= \left[\overline{\mathrm{co}} (\tilde{A}_\varepsilon(x) \cup (\tilde{B}_\varepsilon(x) \cup C_\varepsilon(x)) \cup L^\perp) \right]_\infty \\ &= \left[\overline{\mathrm{co}} (\tilde{A}_\varepsilon(x) + (\varepsilon\tilde{B}_\varepsilon(x) \cup C_\varepsilon(x)) + L^\perp) \right]_\infty. \end{aligned} \tag{5.7}$$

Additionally, since θ belongs to $C_\varepsilon(x)$ and $L^\perp$, we note that

$$\cup_{t\in T(x)} \partial_\varepsilon f_t(x) \subset \overline{\mathrm{co}} \left(\cup_{t\in T(x)} \partial_\varepsilon f_t(x) + (\varepsilon \tilde{B}_\varepsilon(x) \cup C_\varepsilon(x)) + L^\perp \right). \tag{5.8}$$

Therefore, by (5.2),

$$\begin{aligned} \partial f(x) &\subset \overline{\mathrm{co}} \left(\cup_{t\in T(x)} \partial_\varepsilon f_t(x) + \left[\overline{\mathrm{co}} \left(\cup_{t\in T(x)} \partial_\varepsilon f_t(x) + \left(\varepsilon \tilde{B}_\varepsilon(x) \cup C_\varepsilon(x) \right) + L^\perp \right) \right]_\infty \right) \\ &\subset \overline{\mathrm{co}} \left(\tilde{A}_\varepsilon(x) + \left(\varepsilon \tilde{B}_\varepsilon(x) \cup C_\varepsilon(x) \right) + L^\perp \right) \\ &\subset \mathrm{co} \left(\tilde{A}_\varepsilon(x) + (\varepsilon \tilde{B}_\varepsilon(x) \cup C_\varepsilon(x)) \right) + 2U. \end{aligned}$$

The conclusion follows by taking the intersection over $\varepsilon > 0$ and $U \in \mathcal{N}$. Finally, we are done because the inclusion "$\supset$" follows easily from Theorem 5.1. □

The proof of the following result, already given in [12, Theorem 10], is similar to the one of Corollary 5.2.

Corollary 5.4 *Let $f_t : X \to \overline{\mathbb{R}}$, $t \in T$, be a family of functions in $\Gamma_0(X)$, and denote $f := \sup_{t\in T} f_t$. Assume that T is compact Hausdorff and the mappings $t \mapsto f_t(x)$, $x \in X$, are usc. Then, under condition (4.2), that is, $\operatorname{cl} f = \sup_{t\in T}(\operatorname{cl} f_t)$, we have for every $x \in \operatorname{dom} f$*

$$\partial f(x) = \bigcap_{\varepsilon>0} \overline{\text{co}} \left(\tilde{A}_\varepsilon(x) + (\varepsilon \tilde{B}_\varepsilon(x) \cup \{0\}) \right).$$

Moreover, when f attains its minimum at x we get

$$\partial f(x) = \bigcap_{\varepsilon>0} \overline{\text{co}} \left(\tilde{A}_\varepsilon(x) \cup \varepsilon \tilde{B}_\varepsilon(x) \right).$$

The parameters $\left(\overline{\rho}_{t,\varepsilon}\right)_{t\in T} \subset]0,1]$ are used to give the exact characterization of $\partial f(x)$, by providing a control on those functions f_t which escapes at $-\infty$ at the reference point x. Hence, if we take all the parameters to be equal to 1, then we only obtain an estimate from above for the set $\partial f(x)$, which also could be of some interest. Such an upper estimate also turns into an exact characterization for $\partial f(x)$ when the functions are uniformly bounded below at x, as we show in the following corollary.

Corollary 5.5 *With the assumptions of Theorem 5.1, we suppose that*

$$\inf_{t\in T} f_t(x) > -\infty.$$

Then, for every $x \in \operatorname{dom} f$ we have that

$$\partial f(x) = \bigcap_{\varepsilon>0} \overline{\text{co}} \left(A_\varepsilon(x) + \left(\varepsilon B^1_\varepsilon(x) \cup C_\varepsilon(x) \right) \right),$$

where $B^1_\varepsilon(x) := \cup\{\partial_\varepsilon f_t(x) : t \in T_0 \backslash T_\varepsilon(x)\}$. Also, with the same condition as above, in the context of Theorem 5.3 we have that

$$\partial f(x) = \bigcap_{\varepsilon>0} \overline{\text{co}} \left(\tilde{A}_\varepsilon(x) + \left(\varepsilon \tilde{B}^1_\varepsilon(x) \cup C_\varepsilon(x) \right) \right),$$

where $\tilde{B}^1_\varepsilon(x) := \cup\{\partial_\varepsilon f_t(x) : t \in T_0 \backslash T(x)\}$.

Proof. The proof of the first statement follows by combining Theorem 5.1 and Lemma 2.35. Similarly, the second statement follows by Theorem 5.1 and the same lemma. □

Example 5.6 *Consider the function $f := \sup_{t\in T} \{\langle a_t, \cdot\rangle - b_t\}$ and take $x \in \operatorname{dom} f$. Then, according to Theorem 5.1,*

$$\partial f(x) = \textstyle\bigcap_{\varepsilon>0} \overline{\text{co}}(A_\varepsilon(x) + (\varepsilon B_\varepsilon(x) \cup \{0\})),$$

where

$$A_\varepsilon(x) := \{a_t : \langle a_t, x\rangle - b_t \geq f(x) - \varepsilon\},$$
$$B_\varepsilon(x) := \left\{\frac{\varepsilon a_t}{2b_t - 2\langle a_t, x\rangle + 2f(x) - \varepsilon} : \langle a_t, x\rangle - b_t < f(x) - \varepsilon\right\}.$$

Moreover, if x *satisfies* $\inf_{t\in T}\{\langle a_t, x\rangle - b_t\} > -\infty$, *then Corollary 5.5 yields*

$$\partial f(x) = \bigcap_{\varepsilon>0}\overline{\text{co}}\left(A_\varepsilon(x) + \left(B_\varepsilon^1(x) \cup \{0\}\right)\right),$$

where

$$B_\varepsilon^1(x) := \{a_t : \langle a_t, x\rangle - b_t < f(x) - \varepsilon\}.$$

In the following example, we consider a particular pointwise supremum function, which is the Fenchel conjugate.

Example 5.7 *Take a function* $f : X \to \mathbb{R}_\infty$ *together with its conjugate, that is,*

$$f^*(x^*) = \sup_{x\in \text{dom} f} f_x(x^*), \text{ where } f_x := \langle \cdot, x\rangle - f(x).$$

Given $x^* \in \text{dom} f^*$, *we introduce the sets*

$$A_\varepsilon(x^*) = \{x \in \text{dom} f : \langle x^*, x\rangle - f(x) \geq f^*(x^*) - \varepsilon\} = (\partial_\varepsilon f)^{-1}(x^*),$$
$$B_\varepsilon(x^*) = \left\{\frac{\varepsilon x}{2(f^*(x^*) + f(x) - \langle x, x^*\rangle) - \varepsilon} : x \in \text{dom} f, \langle x^*, x\rangle - f(x) < f^*(x^*) - \varepsilon\right\}.$$

Therefore, by Theorem 5.1,

$$\partial f^*(x^*) = \bigcap_{\varepsilon>0}\overline{\text{co}}\left((\partial_\varepsilon f)^{-1}(x^*) + (\varepsilon B_\varepsilon(x^*) \cup \{0\})\right). \tag{5.9}$$

Moreover, if $\inf_{x\in X} f_x(x) > -\infty$, *then the relation*

$$\text{Argmin}(\overline{\text{co}} f) = \partial f^*(\theta)$$

together with Corollary 5.5 gives us

$$\text{Argmin}(\overline{\text{co}} f) = \bigcap_{\varepsilon>0}\overline{\text{co}}\left(\varepsilon\text{-}\text{Argmin}(f) + \varepsilon B_\varepsilon(\theta) \cup \{0\}\right),$$

where

$$B_\varepsilon(\theta) := \left\{\frac{\varepsilon x}{2(f(x) - \inf_X f) - \varepsilon} : x \in \text{dom} f \setminus (\varepsilon\text{-}\text{Argmin}(f))\right\}.$$

The following corollary gives a way to avoid the boundedness condition used in Corollary 5.5.

Corollary 5.8 *With the assumptions of Theorem 5.1, for every* $x \in \text{dom} f$ *we have that*

$$\partial f(x) = \bigcap_{\varepsilon>0}\overline{\text{co}}\left(A_\varepsilon(x) + (\varepsilon B_\varepsilon^+(x) \cup C_\varepsilon(x))\right),$$

where

$$B_\varepsilon^+(x) := \bigcup\{\partial_\varepsilon g_t(x) : t \in T_0 \backslash T_\varepsilon(x)\}, \text{ with } g_t := \max\{f_t, f(x)\}.$$

Proof. Given $x \in \operatorname{dom} f$ and $\varepsilon > 0$, we start by supposing that $f(x) = 0$, so that $g_t = f_t^+$ for all $t \in T$, $f_t(x) \leq 0$ for all $t \in T$, and $f_t(x) < -\varepsilon$ for all $t \in T \backslash T_\varepsilon(x)$. Next we claim that

$$B_\varepsilon(x) \subset B_{2\varepsilon}^+(x).$$

Indeed, given $x^* \in B_\varepsilon(x)$, we choose $t \in T_0 \backslash T_\varepsilon(x)$ such that $x^* \in \partial_\varepsilon(\rho_{t,\varepsilon} f_t)(x)$. Then, for all $y \in \operatorname{dom}(f_t^+)$ $(= \operatorname{dom}(\rho_{t,\varepsilon} f_t))$, we have that

$$\begin{aligned} \langle x^*, y - x \rangle &\leq (\rho_{t,\varepsilon} f_t)(y) - (\rho_{t,\varepsilon} f_t)(x) + \varepsilon \\ &\leq ((f_t)^+(y) + \varepsilon + \varepsilon = f_t^+(y) - f_t^+(x) + 2\varepsilon, \end{aligned}$$

the last equality because $(f_t)^+(x) = \max\{f_t(x), 0\} = 0$. Hence,

$$x^* \in \partial_{2\varepsilon} f_t^+(x) \subset B_{2\varepsilon}^+(x),$$

and the claim holds true. Consequently, the inclusion "$\subset$" holds and we finish the proof in the current case as the opposite inclusion follows similarly like in the proof of Theorem 5.1.

In the general case, where possibly $f(x) \neq 0$, we consider the convex functions

$$\hat{f}_t := f_t(\cdot) - f(x), \quad t \in T,$$

together with the associated supremum functions $\hat{f} := \sup_{t \in T} \hat{f}_t$. Then $\hat{f} = f - f(x)$ and, so, $\hat{f}(x) = 0$. Thus, from the first paragraph we infer that

$$\partial f(x) = \partial \tilde{f}(x) = \bigcap_{\varepsilon > 0} \overline{\text{co}} \left(\hat{A}_\varepsilon(x) + (\varepsilon \hat{B}_\varepsilon^+(x) \cup \hat{C}_\varepsilon(x)) \right),$$

where

$$\hat{A}_\varepsilon(x) := \bigcup \{ \partial_\varepsilon \hat{f}_t^+(x) : \ t \in T_\varepsilon(x) \}$$

and

$$\hat{B}_\varepsilon^+(x) = \bigcup \{ \partial_\varepsilon \hat{f}_t^+(x) : \ t \in T_0 \backslash T_\varepsilon(x) \}.$$

Thus, since that

$$\hat{f}_t^+(\cdot) = \max\{f_t(\cdot) - f(x), 0\} = \max\{f_t(\cdot), f(x)\} - f(x) = g_t - f(x),$$

we get

$$\hat{A}_\varepsilon(x) = \bigcup \{ \partial_\varepsilon g_t(x) : \ t \in T_\varepsilon(x) \}$$

and

$$\hat{B}_\varepsilon^+(x) = \bigcup \{ \partial_\varepsilon g_t(x) : \ t \in T_0 \backslash T_\varepsilon(x) \}.$$

□

An immediate consequence of Theorem 5.1 is the following extension of the Brøndsted formula in [7], and the formula given in [33, Proposition 6.3] (in finite dimensions and under the continuity of the f_t's).

Corollary 5.9 *Let $f_t : X \to \overline{\mathbb{R}}$, $t \in T$, be a family of convex functions and denote $f := \sup_{t\in T} f_t$. Assume that condition (4.2) holds. Then, for every $x \in \operatorname{dom} f$ such that $T(x) = T$ we have that*

$$\partial f(x) = \bigcap_{\varepsilon>0} \overline{\text{co}} \left(\bigcup_{t\in T} \partial_\varepsilon f_t(x) + \left(\bigcup_{t\in T\setminus T_0} \text{N}^\varepsilon_{\operatorname{dom} f_t}(x) \cup \{\theta\} \right) \right).$$

In particular, if $f_t \in \Gamma_0(X)$ for all $t \in T$, then

$$\partial f(x) = \bigcap_{\varepsilon>0} \overline{\text{co}} \left(\bigcup_{t\in T} \partial_\varepsilon f_t(x) \right).$$

5.3 Reduction process and subdifferential of the supremum

In this section, we use Theorem 3.1 to provide sharp characterizations of the subdifferential of the supremum function in the finite-dimensional space $\mathbb{R}^n$, which only appeal to at most $n+1$ functions from the data.

Theorem 5.10 *Consider a family of lsc proper convex functions $f_t : \mathbb{R}^n \to \mathbb{R}_\infty$, $t \in T$, and denote $f := \sup_{t\in T} f_t$. Let $D \subset X$ be a convex set such that $\operatorname{dom} f \subset D \subset X$, and assume that T is Hausdorff compact and the mappings $t \mapsto f_t(x)$, $x \in D$, are usc. Then, for every $x \in \operatorname{dom} f$ and $\varepsilon > 0$, we have*

1. $\partial_\varepsilon f(x) = \bigcup_{\substack{\varepsilon_1+\varepsilon_2=\varepsilon \\ \varepsilon_1,\ \varepsilon_2\ \geq 0}} \bigcup_{S\subset T,\ |S|\leq n+1} \{\partial_{\varepsilon_1}(\max_{t\in S} f_t + \text{I}_D)(x) : \max_{t\in S} f_t(x) = f(x) - \varepsilon_2\}.$

2. $\partial_\varepsilon f(x) = \bigcup_{\substack{S\subset T,\ |S|\leq n+1,\ \alpha\in\Delta_{|S|} \\ \eta\in[0,\varepsilon],\ \sum_{t\in T}\alpha_t f_t(x)-f(x)+\varepsilon\geq\eta}} \partial_\eta \left(\sum_{t\in T} \alpha_t f_t + \text{I}_D\right)(x).$

Proof. 1. First, to prove the inclusion "$\subset$" we take $x^* \in \partial_\varepsilon f(x)$, so that

$$f(x) - \langle x^*, x\rangle \leq \inf_{\mathbb{R}^n}(f - \langle x^*, \cdot\rangle) + \varepsilon. \tag{5.10}$$

Then, applying Theorem 3.1 to the family of functions $\{f_t - \langle x^*, \cdot\rangle,\ t \in T\}$, there exists an index set $S \subset T$ with $|S| \leq n+1$ such that

$$\inf_{\mathbb{R}^n}(f - \langle x^*, \cdot\rangle) = \inf_{\mathbb{R}^n} \left(\max_{t\in S} f_t - \langle x^*, \cdot\rangle + \text{I}_D \right). \tag{5.11}$$

Thus, taking $\varepsilon_2 := f(x) - \max_{t\in S} f_t(x) \geq 0$,

$$\begin{aligned} \left(\max_{t\in S} f_t - \langle x^*, \cdot\rangle + \text{I}_D \right)(x) &= f(x) - \langle x^*, x\rangle - \varepsilon_2 && (5.12) \\ &\leq \inf_{\mathbb{R}^n}(f - \langle x^*, \cdot\rangle) + \varepsilon - \varepsilon_2 \\ &= \inf_{\mathbb{R}^n} \left(\max_{t\in S} f_t - \langle x^*, \cdot\rangle + \text{I}_D \right) + \varepsilon - \varepsilon_2, && (5.13) \end{aligned}$$

and we deduce that $x^* \in \partial_{\varepsilon-\varepsilon_2}(\max_{t\in S} f_t + \mathrm{I}_{\operatorname{dom} f})(x)$.

To prove the opposite inclusion, we take $x^* \in \partial_{\varepsilon-\varepsilon_2}(\max_{t\in S} f_t + \mathrm{I}_D)(x)$, where $S \subset T$, $|S| \leq n+1$, and $\varepsilon_2 = f(x) - \max_{t\in S} f_t(x)$. So, for all $z \in D$,

$$\begin{aligned}\langle x^*, z - x\rangle &\leq \left(\max_{t\in S} f_t + \mathrm{I}_D\right)(z) - \left(\max_{t\in S} f_t + \mathrm{I}_D\right)(x) + \varepsilon - \varepsilon_2 \\ &\leq f(z) - f(x) + \varepsilon_2 + \varepsilon - \varepsilon_2,\end{aligned}$$

and we infer that $x^* \in \partial_\varepsilon f(x)$.

2. If $x^* \in \partial_\varepsilon f(x)$, then $0 \in \partial_\varepsilon(f - x^*)(x)$ and Theorem 3.1 gives rise to some $S \subset T$, $|S| \leq n+1$, and $\overline{\alpha} \in \Delta_{|S|}$ such that

$$\begin{aligned}f(x) - \langle x^*, x\rangle &\leq \inf_{\mathbb{R}^n}(f - \langle x^*, \cdot\rangle) + \varepsilon = \inf_{\mathbb{R}^n}\left(\max_{t\in S} f_t + \mathrm{I}_D - \langle x^*, \cdot\rangle\right) + \varepsilon \\ &= \max_{\alpha\in\Delta_{|S|}} \inf_{\mathbb{R}^n}\left(\sum_{t\in S}\alpha_t f_t + \mathrm{I}_D - \langle x^*, \cdot\rangle\right) + \varepsilon \quad \text{(by Proposition 2.25)} \\ &= \inf_{\mathbb{R}^n}\left(\sum_{t\in S}\overline{\alpha}_t f_t + \mathrm{I}_D - \langle x^*, \cdot\rangle\right) + \varepsilon;\end{aligned}$$

in particular, we have that $\eta := \varepsilon - f(x) + \sum\limits_{t\in S}\overline{\alpha}_t f_t(x) \in [0, \varepsilon]$. Thus, since

$$\sum_{t\in S}\overline{\alpha}_t f_t(x) - \langle x^*, x\rangle \leq f(x) - \langle x^*, x\rangle,$$

we deduce that

$$x^* \in \partial_\eta\left(\sum_{t\in S}\overline{\alpha}_t f_t + \mathrm{I}_D\right)(x).$$

Now we assume that $x^* \in \partial_\eta\left(\sum\limits_{t\in S}\alpha_t f_t + \mathrm{I}_D\right)(x)$ for some $S \subset T$, $|S| \leq n+1$, $\alpha \in \Delta_{|S|}$ and $\eta \in [0, \varepsilon]$ such that $\varepsilon - f(x) + \sum_{t\in S}\alpha_t f_t(x) \geq \eta$. Then, for all $z \in \operatorname{dom} f$,

$$\begin{aligned}\langle x^*, z - x\rangle &\leq \left(\sum_{t\in S}\alpha_t f_t + \mathrm{I}_D\right)(z) - \left(\sum_{t\in S}\alpha_t f_t + \mathrm{I}_D\right)(x) + \eta \\ &\leq f(z) + \mathrm{I}_D(z) - f(x) - \mathrm{I}_D(x) + \varepsilon \\ &\leq f(z) + \mathrm{I}_{\operatorname{dom} f}(z) - f(x) + \varepsilon \\ &\leq f(z) - f(x) + \varepsilon,\end{aligned}$$

and we infer that $x^* \in \partial_\varepsilon f(x)$.

□

Let us recall that if T is finite in Theorem 5.10, say $T := \{1, \cdots, m\}$ with $m \geq 1$, then by Lemma 4.1 in section 4.2 we have proved that, for every $x \in \operatorname{dom} f$ and $\varepsilon > 0$,

$$\partial_\varepsilon f(x) \subset \bigcup \{\partial_\varepsilon (\sum_{t \in \operatorname{supp} \alpha} \alpha_t f_t)(x) : \sum_{t \in \operatorname{supp} \alpha} \alpha_t f_t(x) \geq f(x) - \varepsilon, \ \alpha \in \Delta_m, \ |\operatorname{supp} \alpha| \leq n+1\}.$$

Moreover, in this finite case, Theorem 5.10 simplifies to:

Corollary 5.11 *Consider a family of lsc proper convex functions $f_1, \cdots, f_m : \mathbb{R}^n \to \mathbb{R}_\infty$, $m > n+1$ and denote $f := \max_{1 \leq i \leq m} f_i$. Then, for every $x \in \operatorname{dom} f$ and $\varepsilon > 0$, we have*

1. $\partial_\varepsilon f(x) = \bigcup\limits_{\substack{\varepsilon_1 + \varepsilon_2 = \varepsilon \\ \varepsilon_1, \varepsilon_2 \geq 0}} \bigcup\limits_{\substack{S \subset \{1, \cdots, m\}, \\ |S| \leq n+1}} \left\{ \partial_{\varepsilon_1} \left(\max_{t \in S} f_t \right)(x) : \max_{t \in S} f_t(x) = f(x) - \varepsilon_2 \right\}.$

2. $\partial_\varepsilon f(x) = \bigcup\limits_{\substack{S \subset \{1, \cdots, m\}, |S| \leq n+1, \alpha \in \Delta_{|S|} \\ \eta \in [0, \varepsilon], \sum_{t \in S} \alpha_t f_t(x) - f(x) + \varepsilon \geq \eta}} \partial_\eta \left(\sum_{t \in S} \alpha_t f_t \right)(x).$

5.4 Danskin's Theorem

This section is dedicated to extend the Danskin Theorem [20, Theorem I (Chapter III)], by using the concept of ε-directional derivative instead of the usual directional derivative.

The classical Danskin's theorem involves a function $F(x, y)$ of two variables x and y, defined on the product of the Euclidean space $\mathbb{R}^n$ and some compact topological space $\mathcal{Y}$. The function $F(x, y)$ as well as its partial derivatives are supposed to be continuous. Then Danskin's theorem allows for an explicit characterization of the directional derivative of the function φ defined as

$$\varphi(x) = \min_{y \in \mathcal{Y}} F(x, y).$$

Namely, if $\varphi'(x; u)$ denotes the directional derivative of φ at the point $x \in X$ in the direction $u \in X$, then $\varphi'(x; u)$ is given by means of the partial directional derivatives of F given, for appropriately chosen elements $y \in \mathcal{Y}$, as

$$F_i^{'}(x, y) := \lim_{s \to 0} s^{-1}(F(x + s e_i, y) - F(x, y)), \ i = 1, \cdots, n,$$

where $\{e_1, \cdots, e_n\}$ denotes the canonical basis of $\mathbb{R}^n$. More precisely, we have

Theorem 5.12 *For every $u = (u_1, \cdots, u_n) \in \mathbb{R}^n$,*

$$\varphi'(x; u) = \min_{y \in \mathcal{Y}(x)} \sum_{i=1}^n u_i F_i^{'}(x, y),$$

where $\mathcal{Y}(x) := \{y \in \mathcal{Y} : \varphi(x) = F(x, y)\}$.

In the line of Danskin's Theorem we present below a characterization for the directional derivative of the supremum function, $f = \sup_{t \in T} f_t$, where T is any set (possible infinite).

Our characterization is given by means of the ε-directional derivatives of the f_t's and makes use of the parameters $\rho_{t,\varepsilon}$ defined in (5.3) as

$$\rho_{t,\varepsilon} := \begin{cases} \frac{-\varepsilon}{2f_t(x)-2f(x)+\varepsilon}, & \text{if } t \in T_0 \setminus T_\varepsilon(x), \\ 1, & \text{if } t \in T_\varepsilon(x), \end{cases}$$

with $T_0 := \{t \in T : \operatorname{cl} f_t \text{ is proper}\}$.

Theorem 5.13 *Consider a family of proper lsc convex functions $f_t : X \to \overline{\mathbb{R}}$, $t \in T$, and denote $f := \sup f_{t\in T}$. If $x \in X$ is such that $\partial f(x) \neq \emptyset$, then we have*

$$\operatorname{cl}(f'(x;\cdot)) = \operatorname{cl}\left(\inf_{\varepsilon>0}\left(\sup_{t\in T_\varepsilon(x)} (f_t)'_\varepsilon(x;\cdot) + \varepsilon \sup_{t\in T\setminus T_\varepsilon(x)} ((\rho_{t,\varepsilon} f_t)'_\varepsilon(x;\cdot))^+\right)\right).$$

If, in addition, $\inf_{t\in T} f_t(x) > -\infty$, then

$$\operatorname{cl}(f'(x;\cdot)) = \operatorname{cl}\left(\inf_{\varepsilon>0}\left(\sup_{t\in T_\varepsilon(x)} (f_t)'_\varepsilon(x;\cdot) + \varepsilon \sup_{t\in T\setminus T_\varepsilon(x)} ((f_t)'_\varepsilon(x;\cdot))^+\right)\right).$$

Proof. According to [60, Corollary 2.4.15], we write

$$\operatorname{cl}(f'(x;\cdot))(u) = \sup\{\langle u, x^*\rangle : x^* \in \partial f(x)\}, \text{ for all } u \in X. \tag{5.14}$$

Then, using Corollary 5.2,

$$\begin{aligned}
\operatorname{cl}(f'(x;u)) &= \sup\{\langle u, x^*\rangle : x^* \in \cap_{\varepsilon>0}\overline{\operatorname{co}}\,(A_\varepsilon(x) + (\varepsilon B_\varepsilon(x) \cup \{\theta\}))\} \\
&= \operatorname{cl}\left(\inf_{\varepsilon>0}(\sup\{\langle u, x^*\rangle : x^* \in A_\varepsilon(x) + (\varepsilon B_\varepsilon(x) \cup \{\theta\})\})\right) \text{ (by Lemma 2.23)} \\
&= \operatorname{cl}\left(\inf_{\varepsilon>0}(\sup\{\langle u, x^*\rangle : x^* \in A_\varepsilon(x)\} + \sup\{\langle u, x^*\rangle : x^* \in \varepsilon B_\varepsilon(x) \cup \{\theta\}\})\right) \\
&= \operatorname{cl}\Big(\inf_{\varepsilon>0}\big(\sup\left\{\langle u, x^*\rangle : x^* \in \cup_{t\in T_\varepsilon(x)}\partial_\varepsilon f_t(x)\right\} + \\
&\qquad \sup\left\{\langle u, x^*\rangle : x^* \in \cup_{t\in T\setminus T_\varepsilon(x)}\left(\varepsilon\partial_\varepsilon(\rho_{t,\varepsilon} f_t)(x) \cup \{\theta\}\right)\right\}\big)\Big) \\
&= \operatorname{cl}\left(\inf_{\varepsilon>0}\left(\sup_{t\in T_\varepsilon(x)}\sup_{x^*\in\partial_\varepsilon f_t(x)\}}\langle u, x^*\rangle + \sup_{t\in T\setminus T_\varepsilon(x)}\sup_{x^*\in\varepsilon\partial_\varepsilon(\rho_{t,\varepsilon}f_t)(x)\cup\{\theta\}}\langle u, x^*\rangle\right)\right) \\
&\qquad \text{(by Lemma 2.23)} \\
&= \operatorname{cl}\left(\inf_{\varepsilon>0}\left(\sup_{t\in T_\varepsilon(x)} (f_t)'_\varepsilon(x;\cdot) + \varepsilon \sup_{t\in T\setminus T_\varepsilon(x)} ((\rho_{t,\varepsilon} f_t)'_\varepsilon(x;\cdot))^+\right)\right),
\end{aligned}$$

and the first statement follows. Under the additional assumption, the proof of the second statement is done in the same way as in the paragraph above by taking as weighting parameters the value 1 instead of the $\rho_{t,\varepsilon}$'s (see Corollary 4.10). □

The following result gives the counterpart of Theorem 5.13 for the compact-continuous setting. In this case, instead of the $\rho_{t,\varepsilon}$ used above we use the the parameters $\overline{\rho}_{t,\varepsilon}$ defined in (5.3), that is,

$$\overline{\rho}_{t,\varepsilon} := \begin{cases} \frac{\varepsilon}{-2f_t(x)+2f(x)+\varepsilon}, & \text{if } t \in T_0 \setminus T(x), \\ 1, & \text{if } t \in T(x). \end{cases}$$

Theorem 5.14 *With the assumption of Theorem 5.13 suppose that T is Hausdorff compact and the mappings $t \mapsto f_t(x)$, $x \in X$, are usc. If $x \in X$ is such that $\partial f(x) \neq \emptyset$, then we have*

$$\text{cl}(f'(x;\cdot)) = \text{cl}\left(\inf_{\varepsilon>0}\left(\sup_{t\in T(x)} (f_t)'_\varepsilon(x;\cdot) + \varepsilon \sup_{t\in T\setminus T(x)} \left((\overline{\rho}_{t,\varepsilon} f_t)'_\varepsilon(x;\cdot)\right)^+\right)\right),$$

and, provided that $\inf_{t\in T} f_t(x) > -\infty$,

$$\text{cl}(f'(x;\cdot)) = \text{cl}\left(\inf_{\varepsilon>0}\left(\sup_{t\in T(x)} (f_t)'_\varepsilon(x;\cdot) + \varepsilon \sup_{t\in T\setminus T(x)} \left((f_t)'_\varepsilon(x;\cdot)\right)^+\right)\right).$$

In particular, if $f_t(x) = f(x)$ for all $t \in T$, then

$$\text{cl}(f'(x;\cdot)) = \text{cl}\left(\inf_{\varepsilon>0}\left(\sup_{t\in T} (f_t)'_\varepsilon(x;\cdot)\right)\right).$$

Proof. The proof is similar to the one of Theorem 5.13, taking into account the compact-continuous setting (using $\overline{\rho}_{t,\varepsilon}$ instead of $\rho_{t,\varepsilon}$ and applying Corollary 5.9). □

A more simple formula of the directional derivative is obtained if the supremum function is continuous somewhere.

Corollary 5.15 *Given a family of convex function $\{f_t,\ t \in T\}$ such that $f := \sup_{t\in T} f_t$ is finite and continuous at a given $x \in X$, we suppose that T is Hausdorff compact and the mappings $t \mapsto f_t(z), z \in X$ are usc. Then we obtain*

$$f'(x;\cdot) = \sup_{t\in T(x)} f'_t(x;\cdot).$$

Proof. The proof is similar to Theorem 5.14 but with the use of [16, Corollary 10]. Indeed, sine the function f is continuous at x (and so are all the f_t's), we obtain for all $u \in X$

$$\begin{aligned} f'(x;\cdot)(u) &= \sup\{\langle u, x^*\rangle : x^* \in \overline{\text{co}}\left(\cup_{t\in T(x)} \partial f_t(x)\right)\} \\ &= \sup_{t\in T(x)} \sup\{\langle u, x^*\rangle : x^* \in \partial f_t(x)\} \\ &= \sup_{t\in T(x)} f'_t(x;\cdot)(u) \quad \text{(by [60, Corollary 2.4.15])}. \end{aligned}$$

□

5.5 Optimality conditions

In the literature on semi-inifnite optimization one can find several KKT-type optimality conditions which use, for example, approximate subdifferentials of the data functions [36], exact subdifferentials of data functions but at close points [56], Farkas-Minkowski-type closedness [23] and many others. In our case we derive new optimality conditions that highlight the role played by ε-active and non ε-active functions.

We consider the following semi-infinite programming problem (SIP, in brief), written in an lcs X as

$$\begin{array}{ll} \inf & f_0(x) \\ \text{s.t.} & f_t(x) \leq 0,\ t \in T, \end{array} \tag{5.15}$$

where T is an arbitrary index set and $f_t : X \to \mathbb{R}_\infty,\ t \in \{0\} \cup T$ are convex and lsc functions.

The following establishes the KKT and Fritz-John optimality conditions for problem (5.15) at an optimal solution $\bar{x} \in X$. Remember that

$$T_0 = \{t \in T : \operatorname{cl} f_t \text{ is proper}\}$$

and, for any $\varepsilon > 0$, the parameters $\rho_{t,\varepsilon}$, $t \in \{T \cup 0\}$ are defined by

$$\rho_{t,\varepsilon} := \begin{cases} \frac{-\varepsilon}{2f_t(\bar{x})+\varepsilon}, & \text{if } t \in T_0 \setminus T_\varepsilon(\bar{x}), \\ 1, & \text{if } t \in \{0\} \cup T_\varepsilon(\bar{x}), \end{cases}$$

where $T_\varepsilon(\bar{x}) = \{t \in T_0 : f_t(\bar{x}) \geq -\varepsilon\}$.

Theorem 5.16 *For every $\varepsilon > 0$ and every $U \in \mathcal{N}$, there are associated elements $t_1, \cdots, t_m \in T_\varepsilon(\bar{x})$, $t_{m+1}, \cdots, t_{m+n} \in T_0 \setminus (T_\varepsilon(\bar{x}) \cup \{0\})$, $t_{m+n+1}, \cdots, t_{m+n+r} \in (T \cup \{0\}) \setminus T_0$ and the multipliers $(\lambda_0, \cdots, \lambda_{m+n+r}) \in \Delta_{m+n+r+1}$, $m, n, r \geq 1$ such that*

$$\theta \in \lambda_0 \partial_\varepsilon f_0(\bar{x}) + \sum_{i=1}^{m} \lambda_i \partial_\varepsilon f_{t_i}(\bar{x}) + \sum_{i=m+1}^{m+n} \varepsilon \lambda_i \partial_\varepsilon (\rho_{t_i,\varepsilon} f_{t_i})(\bar{x}) + \sum_{i=m+n+1}^{m+n+r} \lambda_i \mathrm{N}^{\varepsilon}_{\operatorname{dom} f_{t_i}}(\bar{x}) + U,$$

and, provided that $f_t \in \Gamma_0(X)$ for all $t \in T \cup \{0\}$,

$$\theta \in \lambda_0 \partial_\varepsilon f_0(\bar{x}) + \sum_{i=1}^{m} \lambda_i \partial_\varepsilon f_{t_i}(\bar{x}) + \sum_{i=m+1}^{m+n} \varepsilon \lambda_i \partial_\varepsilon (\rho_{t_i,\varepsilon} f_{t_i})(\bar{x}) + U.$$

Moreover, if $\inf_{t \in T \cup \{0\}} f_t(x) > -\infty$, then instead of $\overline{\rho}_{t,\varepsilon}$ we take the fixed value 1 in the two relations above.

Proof. Since $\bar{x}$ is an optimal point of (5.15), Lemma 3.6 ensures that $\bar{x}$ is a global minimum of the function $g = \sup\{f_0(\cdot) - f_0(\bar{x});\ f_t(\cdot),\ t \in T\}$; hence, $g(\bar{x}) = 0$ and $\theta \in \partial g(\bar{x})$. Therefore, given any $\varepsilon > 0$, Corollary 5.2 entails

$$\begin{aligned} \theta &\in \partial g(\bar{x}) \\ &= \bigcap_{\varepsilon > 0} \overline{\operatorname{co}} \Big(\bigcup_{t \in (T_\varepsilon(\bar{x}) \cup \{0\})} \partial_\varepsilon f_t(\bar{x}) + \Big(\bigcup_{t \in (T_0 \setminus (T_\varepsilon(\bar{x})) \cup \{0\})} \varepsilon \partial_\varepsilon (\rho_{t,\varepsilon} f_t)(\bar{x}) \cup \bigcup_{t \in (T \cup \{0\}) \setminus T_0} \mathrm{N}^{\varepsilon}_{\operatorname{dom} f_t}(\bar{x}) \Big) \Big). \end{aligned}$$

Therefore, for all $\varepsilon > 0$ and for all $U \in \mathcal{N}$,

$$\theta \in \operatorname{co}\Big(\bigcup_{t\in(T_\varepsilon(\bar{x})\cup\{0\})} \partial_\varepsilon f_t(\bar{x}) + \Big(\bigcup_{t\in(T_0\setminus(T_\varepsilon(\bar{x}))\cup\{0\}))} \varepsilon\partial_\varepsilon(\rho_{t,\varepsilon} f_t)(\bar{x}) \cup \bigcup_{t\in(T\cup\{0\})\setminus T_0} \mathrm{N}^\varepsilon_{\operatorname{dom} f_t}(\bar{x}) \Big)\Big) + U,$$

and so there exist elements $t_1, \cdots, t_m \in T_\varepsilon(\bar{x})$, $t_{m+1}, \cdots, t_{m+n} \in T_0\setminus (T_\varepsilon(\bar{x})) \cup \{0\})$, together with $t_{m+n+1}, \cdots, t_{m+n+r} \in (T \cup \{0\}) \setminus T_0$ and $(\lambda_0, \lambda_1, \cdots, \lambda_m, \cdots, \lambda_{m+n}, \cdots, \lambda_{m+n+r}) \in \Delta_{m+n+r+1}$, $m,\ n,\ r \geq 1$ that satisfy the first statement of the theorem.

In the case where all the f_t's are in $\Gamma_0(X)$, we get the desired conclusion by arguing as above with the use of Corollary 5.2. The same arguments are used to establish the last conclusion, thanks to Corollary 4.10. □

Next, we give the counterpart of the previous theorem for the compact-continuous setting, by arguing similarly as in the proof of Theorem 5.16 and using the parameters $\overline{\rho}_{t,\varepsilon}$ defined by

$$\overline{\rho}_{t,\varepsilon} := \begin{cases} \frac{\varepsilon}{-2f_t(x)+\varepsilon}, & \text{if } t \in T_0 \setminus T(x), \\ 1, & \text{if } t \in \{0\} \cup T(x). \end{cases}$$

Theorem 5.17 *Assume that T is a Hausdorff compact set and that the mappings $t \mapsto f_t(x)$, $x \in X$, are usc. Then, for every $\varepsilon > 0$ and every $U \in \mathcal{N}$, there are associated elements $t_1, \cdots, t_m \in T(\bar{x})$, $t_{m+1}, \cdots, t_{m+n} \in T_0\setminus (T(\bar{x}) \cup \{0\})$, $t_{m+n+1}, \cdots, t_{m+n+r} \in (T \cup \{0\}) \setminus T_0$ and $(\lambda_0, \lambda_1, \cdots, \lambda_m, \cdots, \lambda_{m+n}, \cdots, \lambda_{m+n+r}) \in \Delta_{m+n+r+1}$, $m,\ n,\ r \geq 1$ such that*

$$\theta \in \lambda_0\partial_\varepsilon f_0(\bar{x}) + \sum_{i=1}^{m} \lambda_i\partial_\varepsilon f_{t_i}(\bar{x}) + \sum_{i=m+1}^{m+n} \varepsilon\lambda_i\partial_\varepsilon(\overline{\rho}_{t_i,\varepsilon} f_{t_i})(\bar{x}) + \sum_{i=m+n+1}^{m+n+r} \lambda_i \mathrm{N}^\varepsilon_{\operatorname{dom} f_{t_i}}(\bar{x}) + U,$$

and, provided that $f_t \in \Gamma_0(X)$ for all $t \in T \cup \{0\}$,

$$\theta \in \lambda_0\partial_\varepsilon f_0(\bar{x}) + \sum_{i=1}^{m} \lambda_i\partial_\varepsilon f_{t_i}(\bar{x}) + \sum_{i=m+1}^{m+n} \varepsilon\lambda_i\partial_\varepsilon(\overline{\rho}_{t_i,\varepsilon} f_{t_i})(\bar{x}) + U.$$

Moreover, under the Slater condition and the assumption that $f_t \in \Gamma_0(X)$, for all $t \in T\cup\{0\}$, we have that $\lambda_0 > 0$.

Proof. The proof is analogous to the proof of the theorem above but with the use of Corollary 5.4 instead of Corollary 5.2. To establish the last conclusion, we proceed by contradiction and assume $\lambda_0 = 0$, so that

$$\theta \in \sum_{i=1}^{m} \lambda_i\partial_\varepsilon f_{t_i}(\bar{x}) + \sum_{i=m+1}^{m+n} \varepsilon\lambda_i\partial_\varepsilon(\overline{\rho}_{t_i,\varepsilon} f_{t_i})(\bar{x}) + U.$$

Then direct calculus lead us to

$$\theta \in \bigcap_{\varepsilon>0} \overline{\operatorname{co}} \left(\partial_\varepsilon(\sup_{t\in T} f_t)(\bar{x}), \varepsilon\partial_{2\varepsilon}(\sup_{t\in T} f_t)(\bar{x}) \right) \subset \partial(\sup_{t\in T} f_t)(\bar{x}),$$

which gives us a contradiction. □

Chapter 6

Multiobjective optimization

6.1 Introduction

Consider the multiobjective optimization problem (MOP, for short)

$$\min_{x \in S} \{f_t(x) : t \in T\}, \tag{6.1}$$

where f_t are extended real-valued convex functions, defined on $\mathbb{R}^n$ and indexed in an arbitrary set (possible infinite) T, and S is a nonempty subset of $\mathbb{R}^n$. Associated to problem (MOP) we consider the sub-problems

$$\min_{x \in S} \{f_t(x) : t \in B\}, \tag{6.2}$$

where $B \subset T$. We shall use two concepts of solutions for problem (MOP), the (classical) efficient and the weakly efficient solutions.

The model above has been proved to be very useful in different applications such as allocation problems, approximation theory, cooperative n-persons games, portfolio problems, Engineering design and so on (see, e.g., [38] and [54]).

We want to apply the reduction process of Chapter 3 to problem (MOP), in order to characterize the set of weakly efficient solutions of problem (MOP) by means of efficient and/or weakly efficient solutions of problems (6.2).

First works to set out these objectives have been realized by Lowe et al. [44] for finite index sets T, showing that the set of weakly efficient solutions of problem (MOP) can be obtained upon efficient solutions of all subproblems (6.2). In 1989, Ward [59] improved the previous result by restricting to subproblems [59] associated to subsets $B \subset T$ with cardinaly not exceeding $n+1$. He also studied some particular cases in which the number of objective functions can be reduced to n at most. These results have been done for convex functions. The case of quasi-convex functions has been recently in Plastria [49] for finite index sets T. In this case, the set of weakly efficient solutions of problems(6.1) is completely determined by the set of weakly efficient solutions of all sub-problems with at most $n+1$ objective function. In the compact-continuous setting, where T is compact and the mappings $t \mapsto f_t(x)$, $x \in X$ are continuous, Plastria & Carrizosa [50, Corollary 2.1] studied the case where the convex functions $\{f_t\}_{t\in T}$ are finite-valued on a fixed given set $\Omega \subset \mathbb{R}^n$. They showed that, if the constraint set for problem (6.1) satisfies $S \subset \Omega$, the set of weakly efficient solutions of (6.1) is completely determined by the set of weakly efficient solutions of all sub-problems (6.2) with at most $n+1$ objective functions.

Most of the results dealing with the reduction of the number of constraints in problem (6.1), including the results cited above, are based on Helly's Theorem. In our case, we shall use the results established previously in Chapter 3 for the subdifferential of the supremum function.

Roughly speaking, we shall be interested in characterizations given in the following forms,

$$\mathrm{WE}_T(S) = \bigcup_{B\subset T,\ |B|\leq n+1} \mathrm{WE}_B(S\cap D),$$

for appropriately chosen sets $D \subset \mathbb{R}^n$, and

$$x \in \mathrm{WE}_T(S) \Longleftrightarrow 0 \in \mathrm{co}\left(\bigcup_{t\in T} \partial f_t(x)\right) + \mathrm{N}_S(x).$$

This chapter is divided into several sections. Section 6.2 deals with a reduction process for problem (MOP), when T is Hausdorff compact and the geometric constraints represented by $S \subset \mathbb{R}^n$ is closed and convex. Theorem 6.4 shows that the weakly efficient solutions of the original problem is characterized by means of weakly efficient solutions of all sub-problem with at most $n+1$ objective functions. Theorem 6.4 provides an extension of [50, Theorem 2.1]. In section 6.3 we present a characterization of the weakly efficient points of (MOP) by using the subdifferential of the objective functions and the normal cone to the geometric constraints set. Our main result is Theorem 6.6, where the objective functions are extended real-valued functions defined on an lcs X.

6.2 Reduction processes in multiobjective optimization

Let $f_t : \mathbb{R}^n \to \mathbb{R}_\infty$ be a family of extended real-valued functions, $t \in T$, with T being an arbitrary set of indices, and let $S \subset \mathbb{R}^n$ be a nonempty convex set. We denote $f := \sup_{t\in T} f_t$ and consider the multiobjective optimization problem

$$\mathrm{MOP(T)} \quad \min_{x\in S}\{f_t(x) : t \in T\}. \tag{6.3}$$

There are several types of solutions for Problem (MOP(T)), but we will restrict ourselves in this chapter to efficient and weakly efficient solutions.

Definition 6.1 (Efficient solutions) *The point $\bar{x} \in S \cap \mathrm{dom}\, f$ is said to be efficient solution of problem (6.3) if there is no point $x \in S \cap \mathrm{dom}\, f$ such that (i) $f_t(x) \leq f_t(\bar{x})$, for all $t \in T$, and (ii) $f_{\bar{t}}(x) < f_{\bar{t}}(\bar{x})$, for some $\bar{t} \in T$. The set of the efficient solutions of (6.3) is denoted by $\mathrm{E}(T, S)$.*

Definition 6.2 (Weakly efficient solutions) *The point* $\bar{x} \in S \cap \operatorname{dom} f$ *is a weakly efficient solution of (6.3) if and only if there is no* $y \in S \cap \operatorname{dom} f$ *such that* $f_t(y) < f_t(\bar{x})$, *for all* $t \in T$. *The set of weakly efficient solutions of problem (6.3) is denoted by* $\mathrm{WE}_T(S)$.

Clearly, we have that $\mathrm{E}(T,S) \subset \mathrm{WE}_T(S)$ but the opposite inclusion cloud not be true. We have the following simple example (see [38, Example 11.6]).

Example 6.3 *We consider the multiobjective optimization problem given in* $\mathbb{R}^2$ *by*

$$\min_{x \in S}\{f_1(x), f_2(x)\},$$

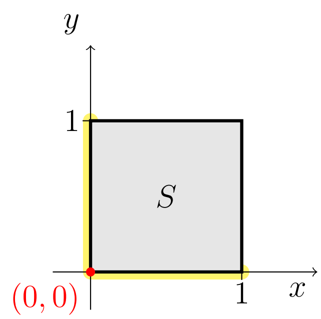


Figure 6.1: $\mathrm{E}(\{1,2\},S)$ in red, $\mathrm{WE}_{\{1,2\}}(S)$ in yellow.

where $S = \{(x_1,x_2) \in \mathbb{R}^2 : 0 \leq x_1 \leq 1,\ 0 \leq x_2 \leq 1\}$, $f_1(x_1,x_2) = x_1$ *and* $f_2(x_1,x_2) = x_2$. *The point* $(0,0)$ *is the only efficient solution of the given problem, but the set* $\{(x_1,x_2) \in S : x_1 = 0$ *or* $x_2 = 0\}$ *represents the set of weakly efficient solutions. See Figure 6.1.*

In the following theorem we characterize the set of weakly efficient solutions for constrained and unconstrained multiobjective optimization problems (6.3), where the constraint set S is convex and the objective functions are indexed in a Hausdorff compact set T.

Theorem 6.4 *Given a family of convex lsc functions* $f_t : \mathbb{R}^n \to \mathbb{R}_\infty$, $t \in T$, *and a closed convex set* $S \subset \mathbb{R}^n$, *we assume that* T *is Hausdorff compact and the mappings* $t \mapsto f_t(x)$, $x \in D$, *are continuous, where* D *is any convex set satisfying (3.7), that is,* $\operatorname{dom} f \subset D \subset \mathbb{R}^n$. *Then we have that*

$$\mathrm{WE}_T(S) = \bigcup_{B \subset T,\ |B| \leq n+1} \mathrm{WE}_B(S \cap D).$$

Proof. Since $\mathrm{WE}_T(S) \subset \operatorname{dom} f \subset D$, the current compactness-continuity assumptions give us

$$\begin{aligned}
\bar{x} \in \mathrm{WE}_T(S) &\iff \forall\, y \in S \cap \operatorname{dom} f,\ \exists\, t \in T \text{ such that } f_t(y) \geq f_t(\bar{x}) \\
&\iff \inf_{y \in S \cap \operatorname{dom} f} \max_{t \in T}(f_t(y) - f_t(\bar{x})) \geq 0 \\
&\iff \inf_{y \in S \cap \operatorname{dom} f} \sup_{t \in T}(f_t(y) - f_t(\bar{x})) \geq 0.
\end{aligned}$$

Let the functions $\varphi_t : \mathbb{R}^n \to \mathbb{R}_\infty$, $t \in T$, be defined as $\varphi_t(y) = f_t(y) - f_t(\bar{x}) + \mathrm{I}_S(y)$, and consider the associated supremum function $\varphi := \sup_{t\in T} \varphi_t$. Then, using [12, Lemma 4], we know that

$$\operatorname{dom}\varphi = \bigcap_{t\in T} \operatorname{dom}(f_t - f_t(\bar{x}) + \mathrm{I}_S) = \bigcap_{t\in T} \operatorname{dom} f_t \cap S = \operatorname{dom} f \cap S,$$

and we deduce that

$$\inf_{y\in S\cap\operatorname{dom} f} \sup_{t\in T}(f_t(y) - f_t(\bar{x})) = \inf_{y\in S\cap D} \sup_{t\in T}(f_t(y) - f_t(\bar{x})) = \inf_{y\in S} \sup_{t\in T}(f_t(y) - f_t(\bar{x})).$$

So, the equivalence above reads

$$\begin{aligned} \bar{x} \in \mathrm{WE}_T(S) \quad &\Longleftrightarrow \quad \inf_{y\in S\cap D} \sup_{t\in T}(f_t(y) - f_t(\bar{x})) \geq 0 \\ &\Longleftrightarrow \quad \inf_{y\in S} \sup_{t\in T}(f_t(y) - f_t(\bar{x})) \geq 0. \end{aligned}$$

Next we verify the hypotheses of Theorem 3.1, when applied to the family of functions $\{\varphi_t,\ t \in T\}$. Indeed, it is clear that all the φ_t's are convex and lsc, and that the mappings $t \mapsto \varphi_t(y)$ are usc for all $y \in D$. Thus, since that $\operatorname{dom}\varphi = \operatorname{dom} f \cap S \subset D$, Theorem 3.1 applies and yields

$$\begin{aligned} \inf_{y\in S} \sup_{t\in T}[f_t(y) - f_t(\bar{x})] &= \inf_{\mathbb{R}^n} \sup_{t\in T}[f_t - f_t(\bar{x}) + \mathrm{I}_S] \\ &= \max_{B\subset T,\ |B|\leq n+1} \inf_{\mathbb{R}^n} \max_{t\in B}\{f_t - f_t(\bar{x}) + \mathrm{I}_S + \mathrm{I}_D\}. \end{aligned}$$

Therefore

$$\begin{aligned} \bar{x} \in \mathrm{WE}_T(S) \quad &\Longleftrightarrow \quad \max_{B\subset T,\ |B|\leq n+1} \inf_{y\in S\cap D} \max_{t\in T}[f_t(y) - f_t(\bar{x})] \geq 0 \\ &\Longleftrightarrow \quad \bar{x} \in \bigcup_{B\subset T,\ |B|\leq n+1} \mathrm{WE}_B(S\cap D), \end{aligned}$$

and we are done. □

Below we present an extension of [50, Corollary 2.1].

Corollary 6.5 *Given a family of convex continuous functions $f_t : \mathbb{R}^n \to \mathbb{R}_\infty$, $t \in T$, we denote $f = \sup_{t\in T} f_t$. Let $S \subset \mathbb{R}^n$ be a closed convex set, and assume that T is Hausdorff compact and the mappings $t \mapsto f_t(x)$, $x \in \mathbb{R}^n$, are continuous. Then*

$$\mathrm{WE}_T(S) = \bigcup_{B\subset T,\ |B|\leq n+1} \mathrm{WE}_B(S).$$

Proof. Due to the continuity of the functions $f_t, t \in T$, it suffices to take $D := \mathbb{R}^n$ in Theorem 6.4. □

6.3 Characterization of weakly efficient solutions

The following Theorem extends [50, Theorem 2.1] to extended real-valued lsc convex functions defined on an lcs X.

Theorem 6.6 *Given a family of convex lsc proper functions $f_t : X \to \mathbb{R}_\infty$, $t \in T$, we denote $f := \sup_{t\in T} f_t$. Let S be a closed convex subset of X, and assume that T is Hausdorff compact and the mappings $t \mapsto f_t(x)$, $x \in X$, are continuous. If the function f is continuous at some point in S, then $\bar{x} \in \mathrm{WE}_T(S)$ if and only if*

$$\theta \in \overline{\mathrm{co}}\left(\bigcup_{t\in T} \partial f_t(\bar{x})\right) + \mathrm{N}_{\mathrm{dom}\, f}(\bar{x}) + \mathrm{N}_S(\bar{x}), \tag{6.4}$$

if and only if

$$\theta \in \mathrm{N}_S(\bar{x}) + \bigcap_{\varepsilon>0} \overline{\mathrm{co}}\left(\bigcup_{t\in T} \partial_\varepsilon f_t(\bar{x})\right).$$

Moreover, when $X = \mathbb{R}^n$ the same holds if we replace $\overline{\mathrm{co}}$ in (6.4) with co.

Proof. As we have observed before we have that

$$\begin{aligned} \bar{x} \in \mathrm{WE}_T(S) \iff\ & \bar{x} \text{ is an optimal solution to the Infsup problem} \\ & \min_{y\in\mathbb{R}^n} \sup_{t\in T} \left(f_t(y) - f_t(\bar{x}) + \mathrm{I}_S(y)\right) \qquad (6.5) \\ \iff\ & \theta \in \partial\left((\sup_{t\in T} f_t - f_t(x)) + \mathrm{I}_S\right)(\bar{x}). \qquad (6.6) \end{aligned}$$

Moreover, since the function $\sup_{t\in T} f_t(\cdot)$ is continuous at some point in S by the current assumption, the function $\sup_{t\in T} (f_t - f_t(\bar{x}))$ is continuous at the same point. This is a consequence of the current compactness-continuity assumption which ensure that

$$c := \inf_{t\in T} f_t(\bar{x}) > -\infty,$$

so that

$$\sup_{t\in T} (f_t - f_t(\bar{x})) \leq \sup_{t\in T} f_t - c.$$

Then, by applying Moreau-Rockafellar's Theorem, we conclude that $\bar{x} \in \mathrm{WE}_T(S)$ if and only if

$$\partial\left((\sup_{t\in T} f_t - f_t(\bar{x})) + \mathrm{I}_S\right)(\bar{x}) = \partial\left(\sup_{t\in T} f_t - f_t(\bar{x})\right)(\bar{x}) + \mathrm{N}_S(\bar{x}).$$

So, [13, Corollary 6.1.12] entails that $\bar{x} \in \mathrm{WE}_T(S)$ if and only if

$$\theta \in \overline{\mathrm{co}}\left(\bigcup_{t\in T} \partial f_t(\bar{x})\right) + \mathrm{N}_{\mathrm{dom}\, f}(\bar{x}) + \mathrm{N}_S(\bar{x}).$$

Moreover, when $X = \mathbb{R}^n$ this last conclusion is valid if we use co instead of $\overline{\text{co}}$, due to [13, Corollary 6.1.12]. Furthermore, thanks to Corollary 5.9, we have that

$$\partial \left(\sup_{t \in T} (f_t - f_t(\bar{x})) \right) (\bar{x}) = \bigcap_{\varepsilon > 0} \overline{\text{co}} \left(\bigcup_{t \in T} \partial f_t(\bar{x}) \right),$$

and so, arguing as above, we get the remaining conclusion. □

Chapter 7

Strong duality for some quadratic problems

7.1 Introduction

Let A be any real symmetric matrix of order n, and $C \subset \mathbb{R}^n$ be a pointed, closed, convex cone with non-empty interior. We study the quadratic problem

$$\mu = \min \left\{ \frac{1}{2} x^\top A x : e^\top x = 1, x \in C \right\}, \tag{7.1}$$

where $e \in \operatorname{int} C^*$ and $C^*(= -C^-)$ is the non-negative polar cone of C. Its (Lagrangian) dual problem is

$$\sup_{\lambda \in \mathbb{R}} \inf_{x \in C} \left\{ \frac{1}{2} x^\top A x + \lambda (x^\top e e^\top x - 1) \right\}.$$

We only work with homogeneous quadratic objective functions because, due to the structure of the feasible set, we can transform any quadratic function into a homogeneous one.

The standard quadratic optimization problem (StQO), introduced in Bomze [2], corresponds to the case $C = \mathbb{R}^n_+,\ e = \mathbb{1} = (1, \cdots, 1) \in \operatorname{int} \mathbb{R}^n$:

$$\min_{x \in \Delta_n} \frac{1}{2} x^\top A x,$$

where Δ_n is the n-dimensional simplex. There are many applications of this formulation such as quadratic allocation problems, portfolio optimization problems and the maximum weight clique problem among others.

One of the goals of this chapter is to establish an S-lemma. We will first state an equivalence to the fulfillment to the strong duality for (7.1) (with respect to a suitable dual problem) in terms of the convexity of $(f, g)(C) + \mathbb{R}_+(1, 0)$, without passing by a copositive representation scheme.

A second issue we will deal with is the study of the validity of strong duality for the primal problem (7.1) with respect to Lagrangian dual problems: we characterize that property via the copositivity of A on suitable subsets of $\mathbb{R}^n$.

We analyze the cases $\mu = 0$ and $\mu > 0$. Obviously:

$$\begin{aligned} \mu = 0 &\iff A \text{ is copositive but not strictly copositive on } C; \\ \mu > 0 &\iff A \text{ is strictly copositive on } C. \end{aligned}$$

The structure of this Chapter is the following: Section 7.2 is dedicated to present a review about Lagrangian duality from two different points of view. In section 7.3 we study the general quadratic problem with one constraint and provide the main result in Theorem 7.6. This shows an explicit form for the perturbed function and provide a strong duality result. Finally, in section 7.4 we treat the standard quadratic problem, here the properties proved for the general problem are obtained as a consequence. In particular, we recover in Theorem 7.8 the result of Bomze, Locatelli & Tardella in [3, Theorem 4].

7.2 Lagrangian duality

Let $C \subseteq \mathbb{R}^n$ be a non-empty set and, given two functions $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$, we consider the minimization problem called primal problem

$$\mu \doteq \min\{f(x) : g(x) = 0, x \in C\}, \tag{P}$$

and the associated Lagrangian dual problem which is given by

$$\nu \doteq \sup_{\lambda \in \mathbb{R}} \inf_{x \in C} \{f(x) + \lambda g(x)\}. \tag{D}$$

Here the set C is called *geometric set of constraints*. The feasible set is denoted by

$$K = \{x \in C : g(x) = 0\}.$$

It is possible to solve the primal problem indirectly by solving the dual problem. Associated with the problem (P), there is a family of *perturbed problems*, which vary according to certain parameter $a \in \mathbb{R}$. These problems are obtained by replacing the original feasible set with the set

$$K(a) = \{x \in C : g(x) = a\}.$$

Moreover, with this new set we define the function $\psi : \mathbb{R} \to \overline{\mathbb{R}}$, called a *perturbation function*. For each $a \in \mathbb{R}$ we associate the value

$$\psi(a) = \begin{cases} \inf\{f(x) : g(x) = a, x \in C\}, & \text{if } K(a) \neq \emptyset, \\ +\infty, & \text{if } K(a) = \emptyset. \end{cases}$$

This problem is denoted by (P_a). In particular, when $a = 0$ we recover the primal problem (P), that is, $\psi(0) = \mu$ and Flores-Bazán, Jourani & Mastroeni show in [30] that $\operatorname{dom} \psi = g(C)$.

Additionally, we associate with problem (P) the Lagrangian

$$L(\gamma, \lambda, x) = \gamma f(x) + \lambda g(x),$$

where $\gamma \geq 0$ and $\lambda \in \mathbb{R}$ are the Lagrangian multipliers. With the scheme above, the dual problem can be rewritten in the form

$$\nu = \sup_{\lambda\in\mathbb{R}} \inf_{x\in C} L(1, \lambda, x).$$

As we mentioned before, the next inequalities trivially hold

$$\gamma \inf_{x\in K} f(x) \geq \inf_{x\in K} L(\gamma, \lambda, x) \geq \inf_{x\in C} L(\gamma, \lambda, x), \quad \text{for all } \gamma \geq 0, \text{for all } \lambda \in \mathbb{R}. \tag{7.2}$$

This implies that $\nu \leq \mu$. So, we only consider the case when $\mu \in \mathbb{R}$ and $K \neq \emptyset$, since, if $\mu = -\infty \Longrightarrow \nu = -\infty$, then any element of $\mathbb{R}$ is a solution of (D). On the other hand, when K is an empty set, the problem (P) does not make sense. Thus, to have a strong duality property we need the opposite inequality in (7.2) for some $\gamma \geq 0$ and $\lambda \in \mathbb{R}$, that is, one must have

$$\begin{array}{rcll}
\gamma(f(x) - \mu) + \lambda g(x) & \geq & 0 & \text{for all } x \in C, \quad (\gamma > 0) \\
\langle(\gamma, \lambda), (f(x) - \mu, g(x))\rangle & \geq & 0 & \text{for all } x \in C \\
\langle(\gamma, \lambda), a\rangle & \geq & 0 & \text{for all } a \in F(C) - \mu(1, 0) \\
\langle(\gamma, \lambda), a\rangle & \geq & 0 & \text{for all } a \in F(C) - \mu(1, 0) + \mathbb{R}_+ \times \{0\} \\
\langle(\gamma, \lambda), a\rangle & \geq & 0 & \text{for all } a \in \mathcal{E}_\mu \\
\langle(\gamma, \lambda), a\rangle & \geq & 0 & \text{for all } a \in \text{co } \mathcal{E}_\mu \\
\langle(\gamma, \lambda), a\rangle & \geq & 0 & \text{for all } a \in \overline{\text{cone}}(\text{co } \mathcal{E}_\mu),
\end{array}$$

where $F = (f, g)$, $F(C) = \{(f(x), g(x)) : x \in C\}$ and $\mathcal{E}_\mu = F(C) - \mu(1, 0) + \mathbb{R}_+ \times \{0\}$.

From the above relationships, the first equivalence of the following theorem that relates strong Duality to the set $\overline{\text{cone}}(\text{co } \mathcal{E}_\mu)$ appears naturally. This result appears in [31, Theorem 3.2].

Theorem 7.1 *Let $\mu \in \mathbb{R}$. The following statements are equivalent*

a) Strong Duality holds.

b) $\overline{\text{cone}}\,(\text{co}\,\mathcal{E}_\mu) \cap -(\mathbb{R}_{++} \times \{0\}) = \emptyset$.

The next Theorem comes from [30, Theorem 3.1].

Theorem 7.2 *If $K(0) \neq \emptyset$, then $\nu = \psi^{**}(0)$.*

Proposition 7.3 *If $\mu = \psi(0) \in \mathbb{R}$, then the following statements hold:*

a) $\overline{\psi}(0) = \psi(0) \Longleftrightarrow \overline{\mathcal{E}_\mu} \cap -\mathbb{R}_{++} \times \{0\} = \emptyset$,

b) $\text{co}\,\psi(0) = \psi(0) \Longleftrightarrow \text{co}\,\mathcal{E}_\mu \cap -\mathbb{R}_{++} \times \{0\} = \emptyset$.

When the set $\mathcal{E}_\mu$ is convex and the optimal value of the primal problem is finite, Theorem 3.6 in [30] provides a zero duality gap.

Theorem 7.4 *If $\mathcal{E}_\mu$ is convex and $\mu \in \mathbb{R}$, then the following hold:*

a) $\nu = \overline{\psi}(0)$.

b) $\nu = \mu \Longleftrightarrow \psi(0) = \overline{\psi}(0)$.

Item *b*) in the previous theorem says that, under the convexity hypothesis, we only need to add the lower semi-continuity at 0 of the perturbed function to guarantee a zero duality gap.

In [30, Theorem 3.2] a relationship is presented between the perturbed function ψ and the set $F(C) + \mathbb{R}_+ \times \{0\}$, which says

$$F(C) + \mathbb{R}_+ \times \{0\} \subseteq \operatorname{epi}\psi \subseteq \overline{F(C) + \mathbb{R}_+ \times \{0\}}.$$

If in the previous inclusions we take the closure and convex hulls and also consider the set $\mathcal{E}_\mu$, then we obtain the following equalities

$$\begin{aligned} \overline{\operatorname{epi}\psi} &= \overline{F(C) + \mathbb{R}_+ \times \{0\}}, \\ \overline{\mathcal{E}_\mu} &= \overline{\operatorname{epi}\psi} - \mu(1,0), \\ \operatorname{co}(\mathcal{E}_\mu) &= \operatorname{co}(\operatorname{epi}\psi) - \mu(1,0). \end{aligned}$$

An alternative way to present the duality theory is exposed in Lemaréchal & Renaud [41], where problem (P) is treated geometrically, this is achieved by relating the infimum of the Lagrangian with the support function. If we keep the notation as before, problem (P) can be equivalently rewrite by

$$\inf\{r_0 : f(x) \leq r_0, g(x) = 0, x \in C\}.$$

And in its geometric version presented in the cited work,

$$\inf\{r_0 : (x; r_0, 0) \in \mathcal{G} \text{ for some } x \in C\},$$

where $\mathcal{G} = \{(x; r_0, a) \in X \times \mathbb{R} \times Y : x \in C,\ f(x) \leq r_0,\ g(x) = a\}$. To show the geometric form of the dual problem (D), we need an auxiliary function. Let the auxiliary function θ defined by $\Theta : \mathbb{R} \to \overline{\mathbb{R}}, \lambda \mapsto \Theta(\lambda) := \inf_{x\in C} L(1, \lambda, x)$.

Observe that

$$\begin{aligned} \Theta(\lambda) &= \inf_{x\in C} \left(f(x) + \lambda g(x)\right) \\ &= -\sup_{x\in C} \left(-f(x) + -\lambda g(x)\right), \end{aligned}$$

so the geometric form of this function will be $\Theta(\lambda) = -\sigma_{\mathcal{G}}(0; -1, -\lambda)$. Indeed, we have that

$$\begin{aligned}
\sigma_{\mathcal{G}}(0; -1, -\lambda) &= \sup_{(x;r_0,r)\in\mathcal{G}} \langle (0; -1, -\lambda), (x; r_0, a)\rangle \\
&= \sup_{(x;r_0,a)\in\mathcal{G}} -r_0 - \lambda^\top a \\
&= -\inf_{(x;r_0,a)\in\mathcal{G}} r_0 + \lambda^\top a \\
&= -\Theta(\lambda).
\end{aligned}$$

With the new notation, the perturbed function is expressed through

$$\psi(a) = \inf\{r_0 : (x; r_0, a) \in \mathcal{G},\ \text{para algún } x \in \mathbb{R}^n\}.$$

In [41, Proposition 2.5], Lemaréchal & Renaud present a relationship between the function Θ and the Fenchel-Legendre conjugate of the perturbed function ψ.

Proposition 7.5 *If the function Θ is not identically $-\infty$, then it holds that*

$$-\Theta(-\lambda) = \psi^*(\lambda).$$

In addition, a modification of the perturbed function is defined. Let $\tilde{\psi}(a) = \inf\{r_0 : (x; r_0, a) \in \mathcal{G}^{**},$ for some $x \in \mathbb{R}^n\}$. This function is shown to be convex and satisfy the following relations

1. $\tilde{\psi}^*(\lambda) = \psi^*(\lambda) = -\Theta(-\lambda)$.
2. $\tilde{\psi}^{**} = \psi^{**}$.

Once both schemes have been presented, it is possible to relate some results that are obtained using the perturbed function. We re-write the set $\mathcal{G}$ to obtain

$$\mathcal{G} = \{(x, f(x), g(x)) : x \in C\} + \{0\} \times \mathbb{R}_+ \times \{0\}.$$

Also we consider the set

$$C \times \operatorname{epi}\psi = \{(x, t, a) \in:\ x \in C,\ \psi(a) \leq t\}.$$

Thus, we have

$$\mathcal{G} \subseteq C \times \operatorname{epi}\psi \subseteq \overline{\mathcal{G}}.$$

7.3 General problem

Let us consider the following data,

$$\begin{aligned} f(x) &= \frac{1}{2}x^\top A x + 2r^\top x, \qquad A \in \mathcal{M}_{n\times n}(\mathbb{R}),\ A = A^\top, x \in \mathbb{R}^n. \\ C &\subseteq \mathbb{R}^n \text{ closed convex cone and pointed.} \\ e &\in \text{int } C^*. \\ P &= \{0\}. \\ g(x) &= e^\top x - 1. \\ \triangle &= \{x \in \mathbb{R}^n :\ g(x) = 0,\ x \in C\}. \end{aligned}$$

We consider the primal Problem

$$\mu = \inf_{x\in\triangle} f(x). \tag{7.3}$$

The function $f : \mathbb{R}^n \to \mathbb{R}$ is a non-homogeneous quadratic form, but as we see before in Chapter 2, due to the structure of the feasible set, we can transform the function f into a homogeneous quadratic form through the matrix

$$\tilde{A} = \frac{1}{2}A + e \cdot r^\top + r \cdot e^\top.$$

Thus,

$$\begin{aligned} \inf_{x\in\triangle} \left(\frac{1}{2}x^\top A\ x + 2 \cdot r^\top x\right) &= \inf_{x\in\triangle} x^\top \left(\frac{1}{2}A + e\cdot r^\top + r \cdot e^\top\right) x \\ &= \inf_{x\in\triangle} \frac{1}{2} x^\top \tilde{A}\ x. \end{aligned}$$

So, we just study the problem when $r = \theta$.

For this type of problem, we want to prove the *zero duality gap* and *strong duality*. In the first part we will see the properties that the problem has and then we will try to prove the desired relation.

Let $a \in \mathbb{R}$ and define the set $\triangle_a = \{x \in C : e^\top x - 1 = a\}$. From here it is easy to see that the set $\triangle_a$ is closed and convex for all $a \in \mathbb{R}$; moreover, if $a \in \mathbb{R}$ is such that $\triangle_a = \emptyset$, then the properties hold trivially. Also, when $a = 0$ we get the set of constraints for the original problem (7.3), that is, $\triangle = \triangle_0$, which is a bounded set. Since the quadratic function is continuous and the set of constraints is compact, by the Weierstrass's Theorem the problem (7.3) attains its minimum, so we can write

$$\mu = \min_{x\in\triangle_0} f(x).$$

We note that the set $\triangle_0$ is a *base* of the constraints set C, because any element in C can be represented by a suitable weight of an element of $\triangle_0$. Indeed, is $a \in \mathbb{R}$ such that $\triangle_a \neq \emptyset$,

$$\begin{aligned} x \in \triangle_a &\iff e^\top x - 1 = a, \qquad x \in C, \\ &\iff x = t \cdot y, \quad \text{for some } y \in \triangle_0 \text{ y } t = 1 + a. \end{aligned}$$

Furthermore, when we use this representation on the objective function, for each $x \in \triangle_a$ we get

$$\begin{aligned} f(x) &= \frac{1}{2}x^\top Ax = \frac{1}{2}(1+a)^2 y^\top Ay \qquad (y \in \triangle_0) \\ &= (1+a)^2 f(y) \qquad (y \in \triangle_0). \end{aligned}$$

From the above analysis, we havethat

$$\begin{aligned} \min_{x\in\triangle_a} f(x) &= (1+a)^2 \min_{y\in\triangle_0} f(y) \\ &= (1+a)^2\mu, \quad \text{for all } a \geq -1. \end{aligned} \tag{7.4}$$

Otherwise, using the perturbation function ψ, we know from Section 7.2 that its domain satisfies the equality $\operatorname{dom}\psi = g(C) + P$. So, by (7.4), for every element of its domain it is possible to compute explicitly the value of the function ψ through the following relation

$$\begin{array}{ccc} \inf_{x\in\triangle_a} f(x) & = & (1+a)^2 \inf_{y\in\triangle_0} f(y) \\ \shortparallel & & \shortparallel \\ \psi(a) & & (1+a)^2\mu. \end{array}$$

Therefore,

$$\psi(a) = \begin{cases} (1+a)^2\mu, & \text{if } a \geq -1, \\ +\infty, & \text{e.o.c.} \end{cases}$$

Then we can conclude that ψ is convex on $\mathbb{R}$ and lsc for every element in its effective domain. Hence by Theorem 7.2 and Theorem 7.3, the duality gap for problem (7.3) is zero. Based on the information obtained, we can see in Figure 7.1 the graph of the perturbation function. Figure 7.2 shows the epigraph of ψ for different optimal values.

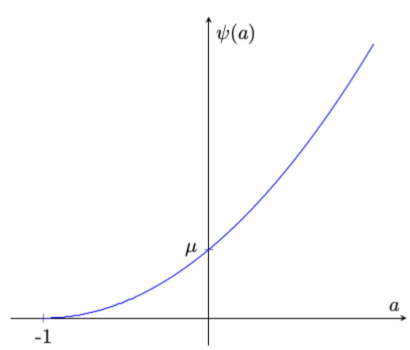
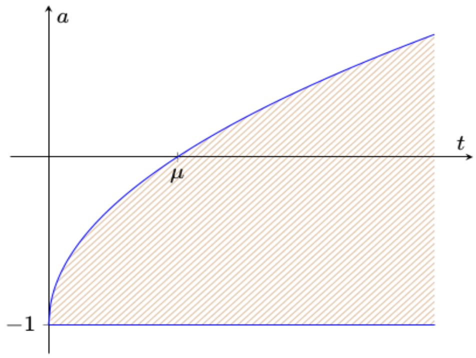


Figure 7.1: Perturbed Function, ψ.

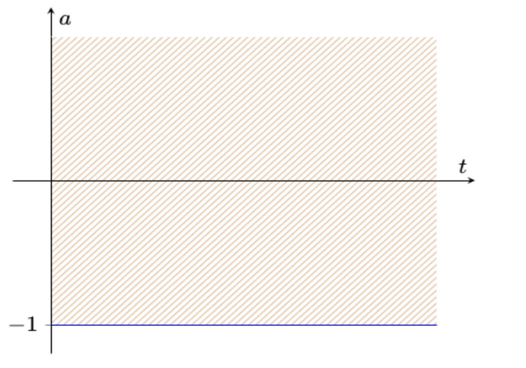
Theorem 7.6 *Let A, C and e be as before. Then,*

a) the perturbed function ψ is given by

$$\psi(a) = \begin{cases} \mu(1+a)^2, & \text{if } a \geq -1; \\ +\infty, & \text{if } a < -1. \end{cases}$$



(a) epi ψ with $\mu > 0$.



(b) epi ψ with $\mu = 0$.

Figure 7.2: Epigraph of perturbed Function, epi ψ.

Thus, ψ is convex if and only if $\mu \geq 0$.

b) $\overline{cone}\,(\mathcal{E}_\mu) \cap -(\mathbb{R}_{++} \times \{0\}) = \emptyset$

c) $F(C)$ *and* $F(C) + \mathbb{R}_+(1,0)$ *are closed, so*

$$\text{epi}\,\psi = F(C) + \mathbb{R}_+(1,0).$$

Proof. a) Seen before.

b) Now we would like to prove that problem (7.3) satisfies the strong duality. Applying Theorem 7.1 we just need to prove that the intersection between the sets is empty. By contradiction, we can assume that there is an element $(a,0) \in \overline{\text{cone}}\,(\mathcal{E}_\mu) \cap -(\mathbb{R}_{++} \times \{0\})$. So, there exists a sequence $\{(a_k,b_k)\}_{k\in\mathbb{N}} \subseteq \text{cone}\,(\mathcal{E}_\mu)$, where $(a_k,b_k) = t_k(u_k,v_k) = (t_k u_k, t_k v_k) \longrightarrow (a,0)$ with $(u_k,v_k) \in \mathcal{E}_\mu$ and $t_k > 0$ (if for all $k \in \mathbb{N}$, $t_k = 0$, we obtain a contradiction because $t_k u_x \to a < 0$). So, $(u_k+\mu, v_k) \in \text{epi}\,\psi$, that means that $\psi(v_k) \leq u_k+\mu = u_k+\psi(0)$. For item a) we know that $\psi(v_k) = (1+v_k)^2\psi(0)$. Hence, if $\psi(0) = 0$, then $0 \leq u_k$ and multiplying both sides of the inequality by t_k and then, taking the limit when $k \to +\infty$, we get a contradiction. On the other hand, if $\psi(0) > 0$, then we multiply the inequality $(1+v_k)^2\psi(0) \leq u_k + \psi(0)$ by t_k to obtain

$$\begin{aligned} & t_k(1+v_k)^2\psi(0) \leq t_k u_k + t_k\psi(0) \\ \iff\ & t_k(1+2v_k+v_k^2)\psi(0) \leq t_k u_k + t_k\psi(0) \\ \Longrightarrow\ & (2t_k v_k + t_k v_k^2)\psi(0) \leq t_k u_k \\ \Longrightarrow\ & 2t_k v_k\psi(0) \leq (2t_k v_k + t_k v_k^2)\psi(0) \leq t_k u_k. \quad (t_k v_k^2 \geq 0) \end{aligned} \tag{7.5}$$

Then, taking the limits when $k \to +\infty$ in (7.5), we come to a contradiction. So, the assumption is false and due to Theorem 7.1 strong duality holds.

c) Additionally, we show that the set $F(C) + \mathbb{R}_+ \times \{0\}$ is closed. Indeed, let $(r,a) \in \overline{F(C) + \mathbb{R}_+ \times \{0\}}$. Then there exists a sequence $\{x_k\}_{k\in\mathbb{N}} \subseteq C$, $\{q_k\}_{k\in\mathbb{N}} \subseteq \mathbb{R}_+$ such that

$$\begin{aligned} (f(x_k)+q_k, g(x_k)) &\longrightarrow (r,a), \\ f(x_k)+q_k &\longrightarrow r, \\ g(x_k) &\longrightarrow a. \end{aligned}$$

First we show that the sublevel set of function g are bounded. Let $\lambda \in \mathbb{R}$ and denote

$$\begin{aligned} S_\lambda &:= \{x \in C : g(x) \leq \lambda\}, \\ &= \{x \in C : e^\top x - 1 \leq \lambda\}. \end{aligned}$$

Since the element e belongs to int C^*, by [28, equation (2.3)] we have $e^\top x > 0$ for all $x \in C \backslash \{0\}$. Now, we assume that there exists a sequence $\{x_k\}_{k \in \mathbb{N}} \subseteq S_\lambda$ such that $\|x_k\| \to +\infty$. From this sequence it is possible to create a new one of the form $\left\{\frac{x_k}{\|x_k\|}\right\} \subseteq C$ such that $\frac{x_k}{\|x_k\|} \to v \neq 0$. Since for all $k \in \mathbb{N}$ the term x_k satisfies the inequality $e^\top x_k - 1 \leq \lambda$, we can multiply both sides by $\dfrac{1}{\|x_k\|}$ to get

$$e^\top \frac{x_k}{\|x_k\|} - \frac{1}{\|x_k\|} \leq \frac{\lambda}{\|x_k\|}. \tag{7.6}$$

Taking the limits when $k \to +\infty$ we obtain $e^\top v \leq 0$, which contradicts the fact that vector e belongs to the set int C^*. This completes the proof, that is, the sublevel sets S_λ are bounded.

Second, since C is a closed set and the sublevels of g are bounded, we can conclude that there exists a convergent subsequence $\{x_{k_l}\}_{l \in \mathbb{N}}$, with $x_{k_l} \longrightarrow \bar{x} \in C$. Then

i. $f(x_{k_l}) \longrightarrow f(\bar{x})$ (f is continuous).

ii. $q_{k_l} = [f(x_{k_l}) + q_{k_l}] - f(x_{k_l}) \longrightarrow r - f(\bar{x}) = q \geq 0$; then, $r = f(\bar{x}) + q \in f(C) + \mathbb{R}_+$.

iii. $g(x_{k_l}) \longrightarrow g(\bar{x}) = a \in g(C)$.

Therefore $(r, a) \in F(C) + \mathbb{R}_+ \times \{0\}$ and, as consequence, the set $F(C) + \mathbb{R}_+ \times \{0\}$ is closed.

Furthermore, from section 7.2 we have certain relations that establish a connection between the set $F(C) + \mathbb{R}_+ \times \{0\}$ and the perturbed function ψ. Since the perturbed function turns out to be convex, we conclude that the last set is also convex.

□

In the following Theorem we establish the explicit value of the function $\Theta(\lambda) = \inf_{x \in C} L(1, \lambda, x)$. This is a generalization of the result of Bomze, Locatelli & Tardella in [3, Theorem 4].

Theorem 7.7 *Let $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ be a symmetric matrix, C a closed convex and pointed cone, $e \in$ int C^*. We have that*

1. If $\mu > 0$, then

$$\Theta(\lambda) = \begin{cases} -\lambda, & \text{if } \lambda \geq 0, \\ -\lambda - \dfrac{1}{4\mu}\lambda^2, & \text{if } \lambda < 0. \end{cases}$$

2. If $\mu = 0$*, then*

$$\Theta(\lambda) = \begin{cases} -\lambda, & \text{if } \lambda \geq 0, \\ -\infty, & \text{if } \lambda < 0. \end{cases}$$

Therefore $\sup\{\Theta(\lambda) : \lambda \in \mathbb{R}\} = \mu$ *and, so, the duality gap is zero.*

Proof. We know that $\lambda \in \mathbb{R}$. We proceed by cases.

1. **Case 1:** $\lambda \geq 0$. The infimum of $\Theta(\lambda)$ is attained at $x = 0$, that is, $\Theta(\lambda) = -\lambda$.

   Case 2: $\lambda < 0$. The inequality $[x^\top Ax + \lambda e^\top x]^2 \geq 0$ is fundamental to the conclusion

$$\begin{aligned} [x^\top Ax + \lambda e^\top x]^2 &= 2(x^\top Ax)\left(\frac{1}{2}x^\top Ax + \lambda e^\top x\right) + (\lambda e^\top x)^2 \geq 0 \\ \iff & \quad \frac{1}{2}x^\top Ax + \lambda e^\top x \geq -\frac{(\lambda e^\top x)^2}{2x^\top Ax}. \end{aligned} \tag{7.7}$$

   Moreover,

$$\frac{(e^\top x)^2}{2x^\top Ax} \leq \frac{1}{4\mu} \quad \forall\, x \in C, x \neq 0. \tag{7.8}$$

   Indeed, let $x = t \cdot y$ where $y \in \triangle_0$. Thus,

$$\frac{(e^\top x)^2}{2x^\top Ax} = \frac{t^2(e^\top y)^2}{2t^2 y^\top Ay} = \frac{1}{2y^\top Ay}.$$

   Since $y \in \triangle_0$, the inequality $0 < \mu \leq \frac{1}{2}y^\top Ay$ is always fulfilled and this implies $\frac{1}{4\mu} \geq \frac{1}{2y^\top Ay}$. For the associated Lagrangian to problem 7.3 we have that:

$$\begin{aligned} L(1, \lambda, x) &:= \frac{1}{2}x^\top Ax + \lambda(e^\top x - 1) \\ &= \frac{1}{2}x^\top Ax + \lambda e^\top x - \lambda \\ &\geq -\frac{(\lambda e^\top x)^2}{2x^\top Ax} - \lambda \qquad \text{(by (7.7))} \\ &\geq -\frac{\lambda^2}{4\mu} - \lambda. \qquad \text{(by (7.8))} \end{aligned}$$

   Let $\bar{x}$ be an optimal solution to problem (7.3). Then $\bar{x} \in C$ and since C is a cone, we have $x_0 = -\frac{\lambda}{2\mu}\bar{x} \in C$. So, x_0 is a feasible point of problem $\inf_{x \in C} L(1, \lambda, x)$ and

$$\begin{aligned} L(1, \lambda, x_0) &= \frac{1}{2}x_0^\top Ax_0 + \lambda(e^\top x_0 - 1) \\ &= \left(-\frac{\lambda}{2\mu}\right)^2 \frac{1}{2}\bar{x}^\top A\bar{x} - \frac{\lambda^2}{2\mu} - \lambda \\ &= \frac{\lambda^2}{2\mu}\left[\frac{1}{2} - 1\right] - \lambda \\ &= -\frac{\lambda^2}{4\mu} - \lambda. \end{aligned}$$

With the arguments above it is proved that

$$\inf_{x\in C} L(1,\lambda,x) = -\frac{\lambda^2}{4\mu} - \lambda, \text{ when } \lambda < 0.$$

2. **Case 1:** $\lambda \geq 0$. The argument is the same as the previous item.

   Case 2: $\lambda < 0$. For the function Θ we have that

$$\begin{aligned}\Theta(\lambda) &= \inf_{x\in C}\left\{\frac{1}{2}x^\top Ax + \lambda(e^\top x - 1)\right\}\\ &= \inf_{t\geq 0,\ y\in\triangle_0}\left\{t^2\cdot\frac{1}{2}y^\top Ay + \lambda(t-1)\right\}\\ &= -\infty.\end{aligned}$$

□

Therefore, as a conclusion, we do not only know the explicit value of the perturbed function, ψ, but also the explicit value of the function $\Theta(\lambda) = \inf_{x\in C} L(1,\lambda,x)$.

7.4 Standard quadratic problems

The standard quadratic problem is given by:

$$\mu = \min\left\{\frac{1}{2}x^\top Ax : \mathbb{1}^\top x = 1, x \geq 0\right\}, \tag{7.9}$$

where $A \in \mathcal{M}_{n\times n}(\mathbb{R})$ is a symmetric matrix and $\mathbb{1} = (1,\cdots,1) \in \mathbb{R}^n$.

Observe that this Problem (7.9) is just a particular case of Problem (7.3), because it is enough to consider $C = \mathbb{R}^n_+$. Moreover, the closed convex pointed cone coincides with its polar set, so $\mathbb{1} = e \in \text{int } \mathbb{R}^n_+ = \mathbb{R}^n_{++}$ and the involved functions become $f(x) = \frac{1}{2}x^\top Ax$ and $g(x) = \mathbb{1}^\top x - 1$. Now, the feasible set is $K = \{x = (x_1,\cdots,x_n) \in \mathbb{R}^n_+ : g(x) = 0\}$.

On the other hand, remember that for this problem the quadratic form may or may not be homogeneous. But as we saw before, in Chapter 2, it is possible to assume that the entries of the matrix A are all strictly positive, that is, the condition $A = (a_{ij}), a_{ij} > 0$ is a consequence of the problem and not an additional hypothesis.

By virtue of these observations, we have all the results obtained in the generalized problem. In summary, the standard quadratic problem satisfies the following properties:

1. For the perturbed function ψ, we have an explicit expression plus some properties

   i) $\psi(a) = \begin{cases}(1+a)^2\mu, & \text{if } a \geq -1;\\ +\infty, & \text{if } a < -1.\end{cases}$

ii) convex on $\mathbb{R}$.

iii) continuous on $]-1, +\infty[$.

iv) lsc for all $a \geq -1$.

2. Zero duality gap.

3. Strong duality.

4. $F(C) \; + \; \mathbb{R}_+ \times \{0\}$ is closed and convex.

The following Theorem, coming from Bomze, Locatelli & Tardella in [3, Teorema 4], gives us the explicit value of the function $\Theta(\lambda) = \inf_{x \geq 0} L(1, \lambda, x)$ for the standard quadratic problems.

Theorem 7.8 *Assume that the matrix A has only positive entries. Then, for function Θ we obtain*

$$\Theta(\lambda) = \begin{cases} -\lambda, & \text{if } \lambda \geq 0, \\ -\lambda - \dfrac{1}{4\mu}\lambda^2, & \text{if } \lambda < 0. \end{cases}$$

Therefore $\sup\{\Theta(\lambda) : \lambda \in \mathbb{R}\} = \mu$ *and, thus, the duality gap is zero.*

Proof. The proof is analogous to proof of Theorem 7.8 with $C = \mathbb{R}^n_+$ and the vector $\mathbb{1} = e$. □

Chapter 8

Future work

The research developed in this PhD thesis provided solutions to many open problems, but also gave rise to some open questions that deserve to be studied in future works. Some of them are described below:

1. I would to work on numerical implementations of our reduction methods, previously presented in Chapter 3, in order to take advantage of this kind of reductions when general semi-infinite optimization problems are involved. In the literature, we can find some research works that can help us with our goal like, for example, the books of Fletcher [25] and Nocedal & Wright [47], and the paper of Cera et al. [11]. I am also considering the use of these methods in concrete problems like in locating problems, Chebyshev approximation theory and multi-objective problems.

2. I would like to extend the reduction methods obtained for SIP and semi-infinite multi-objective problems to cover quasi-convex functions, that can be considered either to model the (multi-) objective functions or the constraints. I am aware that the approach used here in the convex case could not be extended in a direct way, since there is still no consensus on a unified concept of subdifferential for such functions. But still there is a hope to obtain some things in this direction because of the rich properties of quasi-convex functions, as it can be confired by the works of Auslender & Teboulle [1], Flores-Bazán & Hadjisavvas [29], Flores-Bazán & Thiele [32], Jeyakumar et al. [39] and Rockafellar [53].

3. Due to the diversity of formulas developed for the characterization of the subdifferential for the supremum function, we would like to extend them to the non-convex setting, possibly with the use of generalized concepts like the strong and weak slopes ([22]). This would allow us to extend our results to the general setting of metric spaces and, so, to cover, a wide range of applications namely in the theory of metric critical points (see for example [18], [19] and [21]).

Bibliography

[1] Auslender, A. & Teboulle, M. Asymptotic Cones and Functions in Optimization and Variational Inequalities, Springer (2003).

[2] Bomze, I. M. *On standard quadratic optimization problems*, J. Glob. Optim., 13 (1998), pp. 369-387.

[3] Bomze, I. M., Locatelli, M. & Tardella, F. *New and old bounds for stándard quadratic optimization: dominance, equivalence and incomparability*, Math. Program. Ser. A., 115 (2008), pp. 31-64.

[4] Bonnans, J. F. & Shapiro, A. Perturbation Analysis of Optimization Problems, Springer Series in Operations Research, Springer-Verlag, New York, (2000).

[5] Borwein, J. M. *Direct theorems in semi-infinite convex programming*, Math. Program. 21, no. 3 (1981), pp. 301-318.

[6] Borwein, J. M. *The limiting Lagrangian as a consequence of Helly's theorem*, J. Optim. Theory Appl. 33, no. 4 (1981), pp. 497-513.

[7] Brøndsted, A. *On the subdifferential of the supremum of two convex functions*, Math. Scand., 31 (1972), pp. 225-230.

[8] Brøndsted, A. & Rockafellar, R. T., *On the subdifferentiability of convex functions*, Proc. Amer. Math. Soc., 16 (1965), pp. 605-611.

[9] Caro, S. & Hantoute, A. *Discretization and reduction of Infsup and SIP optimization problem*, pre-print, (2023).

[10] Caro, S. & Hantoute, A. *Highlighting the role of active and non-active functions in optimization and subdifferential calculus*, pre-print, (2023).

[11] Cera, M., Mesa, J.A., Ortega, F.A. & Plastria, F. *Locating a Central Hunter on the Plane*, J. Optim. Theory Appl., 136 (2008), pp. 155-166.

[12] Correa, R., Hantoute, A. & López, M. A. *Alternative representations of the normal cone to the domain of supremum functions and subdifferential calculus*, Set-Valued Var. Anal., 29 (2021), pp. 683-699.

[13] Correa, R., Hantoute, A. & López, M. A. Fundamentals of Convex Analysis and Optimization, A Supremum Function Approach, Springer Cham, (2023).

[14] Correa, R., Hantoute, A. & López, M. A. *Subdifferential of the supremum via compactification of the index set*, Vietnam J. Math., 38 (2020), pp. 569–588.

[15] Correa, R., Hantoute, A. & López, M. A. *Valadier-like formulas for the supremum function I*, J. Convex Anal., 25 (2018), no. 4, pp. 1253-1278.

[16] Correa, R., Hantoute, A. & López, M. A. *Valadier-like formulas for the supremum function II: the compactly indexed case*, J. Convex Anal., 26 (2019), pp. 299-324.

[17] Correa, R., Hantoute, A. & López-Cerdá, M. A. *Weaker conditions for subdifferential calculus of convex functions*, J. Funct. Anal., 271 (2016), pp. 1177-1212.

[18] Corvellec, J.-N. *A general approach to the Min-Max principle*, Z. Anal. Anwend., 16 (1997), no. 2, pp. 405-433.

[19] Corvellec, J-N. & Hantoute, A. *Homotopical Stability of isolated critical points of continuous functionals*, Set-Valued Anal., 10 (2002), pp. 143-164.

[20] Danskin, J. M. The Theory of Max-Min and its Application to Weapons Allocation Problems, Springer-Verlag Berlin, Heidelberg, (1967).

[21] Degiovanni, M. A Survey on Nonsmooth Critical Point Theory and Applications, Nonconvex Optimization and Its Applications, Springer, Boston, MA., vol 55 (2001).

[22] De Giorgi, E., Marino, A. & Tosques, M. *Problemi di evoluzione in spazi metrici e curve di massima pendenza* Atti della Accademia Nazionale dei Lincei. Classe di Scienze Fisiche, Matematiche e Naturali. Rendiconti 68.3 (1980), pp. 180-187.

[23] Dinh, N., Goberna, M.A. & López, M.A. *From linear to convex systems: consistency, Farkas' lemma and applications*, J. Convex Anal. 13 (2006), no. 1, pp. 113-133.

[24] Drezner, Z. *On minimax optimization problems*, Math. Program., 22 (1982), no. 2, pp. 227-230.

[25] Fletcher, R. Practical Methods of Optimization (2nd Edition), Wiley (2000).

[26] Flores-Bazán, F. & Cárcamo, G. *Strong duality and KKT conditions in nonconvex optimization with a single equality constraint and geometric constraint*, Math. Program., Ser. B 168 (2018), pp. 369-400.

[27] Flores-Bazán, F., Cárcamo, G. & Caro, S. *Extensions of the Standard Quadratic Optimization Problem: Strong Duality, Optimality, Hidden Convexity and S-lemma*, Appl. Math. Optim. 81 (2020), no. 2, pp. 383-408.

[28] Flores-Bazán, F., Flores-Bazán, F. & Vera, C. *Gordan-type alternative theorems and vector optimization*, Recent Developments in Vector Optimization, Q. H Ansari and J.-C. Yao (Editors), Springer-Verlag, Berlin, (2012), Vol. 1.

[29] Flores-Bazán, F. & Hadjisavvas, N. *Zero-Scale asymptotic functions and quasiconvex Optimization*, J. Convex Anal., 26 (2019), No. 4, pp. 1253-1274.

[30] Flores-Bazán, F., Jourani, A. & Mastroeni, G. *On the convexity of the value function for a class of nonconvex variational problems: existence and optimality conditions*, SIAM J. Control Optim., 52 (2014), pp. 3673-3693.

[31] Flores-Bazán, F. & Mastroeni, G. *Strong Duality in cone constrained nonconvex optimization*, SIAM J. Optim., 23 (2013), no. 1 , pp. 153–169.

[32] Flores-Bazán, F. & Thiele, F. *On the lower semicontinuity of the value function and existence of solutions in quasiconvex optimization*, J. Optim. Theory Appl., 195 (2022), pp. 390-417.

[33] Hantoute, A. & López, M. A., *A complete characterization of the subdifferential set of the supremum of an arbitrary family of convex functions.* J. Convex Anal., 15 (2008), pp. 831-858.

[34] Hantoute, A. & López-Cerdá, M. A. *New Tour on the Subdifferential of Supremum via Finite Sums and Suprema*, J. Optim. Theory Appl., 193 (2022), pp. 81-106.

[35] Hantoute, A., López-Cerdá, M. A. & Zălinescu, C. *Subdifferential calculus rules in convex analysis: A unifying approach via pointwise supremum functions*, SIAM J. Optim., 19 (2008), pp. 863-882.

[36] Hiriart-Urruty, J.-B. & Phelps, R. R. *Subdifferential calculus using epsilon-subdifferentials*, J. Funct. Anal., 118 (1993), pp. 154-166.

[37] Ioffe, A. D. A note on subdifferentials of pointwise supremo, TOP 20 (2012), pp. 456–466.

[38] Jahn, J. Vector Optimization Theory, Applications, and Extensions, Second Edition, Springer (2011).

[39] Jeyakumar, V. , Oettli, W. & Natividad, M. *A solvability Theorem for a Class of quasiconvex mappings with applications to Optimization*, J. Math. Anal. Appl., 179 (1993), pp. 537-546 .

[40] Klee, V. Jr. *The critical set of a convex body*, Amer. J. Math. 75 (1953), pp. 178-188.

[41] Lemaréchal, C. & Reanud, A. *A geometric study of duality gaps, with applications*, Math. Program. Ser. A., 90 (2001), pp. 399–427.

[42] Levin, V. L. *An application of Helly's theorem in convex programming, problems of best approximation and related questions*. Mat. Sb., Nov. Ser., 79 (1969), pp. 250-263. English transl.: Math. USSR, Sb., 8, pp. 235-247.

[43] López, M. A. & Volle, M. *A formula for the set of optimal solutions of relaxed minimization problems. Applications to subdifferential calculus*, J. Convex Anal., 17 (2010), pp. 1057-1075.

[44] Lowe, T., Thiesse, J., Ward, J. & Wendell, R. *On Efficient Solutions to Multiple Objective Mathematical Programs*, Manag. Sci., 30 (1984), no. 11, pp. 1346-1349, .

[45] Mawhin, J. & Willem, M. Critical Point Theory and Hamiltonian Systems, Springer (1989).

[46] Mordukhovich, B. & Nghia, T. *Subdifferentials of nonconvex supremum functions and their applications to semi-infinite programs with Lipschitzian data*, SIAM J. Optim., 23 (2013), no. 1, pp. 406-431.

[47] Nocedal, J. & Wright, S. J. Numerical Optimization, Springer (2006).

[48] Peypouquet, J. Convex Optimization in Normed Spaces: Theory, Methods and Examples, springer briefs in optimization (2015).

[49] Plastria, F. *On the Structure of the Weakly Efficient Set for Quasiconvex Vector Minimization*, J. Optim. Theory Appl., 184 (2020), pp. 547-564.

[50] Plastria, F. & Carrizosa, E. *Geometrical characterization of weakly efficient points*, J. Optim. Theory Appl., 90 (1996), no. 1, pp. 217-223.

[51] Polik, I. & Terlaky, T. A survey of the S-lemma, SIAM J. Optim., 49 (2007), no. 3, pp. 371–418.

[52] Rockafellar, R. T. Convex Analysis, Princeton Landmarks in Mathematics and Physics, (1970).

[53] Rockafellar, R. T. *Helly's theorem and minima of convex functions*, Duke Math. J., 32 (1965), pp. 381-397.

[54] Sawaragi,Y, Nakayama, H. & Tanino, T. Theory of Multiobjective Optimization, (1985).

[55] Shapiro, A. *Semi-infinite programming, duality, discretization and optimality conditions*, Optimization, 58 (2009), no. 2, pp. 133-161.

[56] Thibault, L. *Sequential convex subdifferential calculus and sequential Lagrange multipliers*, SIAM J. Control Optim., 35 (1997), pp. 1434-1444.

[57] VALADIER, M. *Sous-différentiels d'une borne supérieure et d'une somme continue de fonctions convexes*, C. R. Acad. Sci. Paris Sér. A-B 268 (1969), pp. A39-A42 .

[58] VOLLE, M. *Sous-différentiel d'une enveloppe supérieure de fonctions convexes*, C. R. Acad. Sci. Paris Sr. I Math. 317 (1993), pp. 845-849.

[59] WARD, J. *Structure of Efficient Sets for convex Objectives*, Math. Oper. Res., 14 (1989), no. 2, pp. 249-257.

[60] ZĂLINESCU, C. Convex Analysis in General Vector Spaces, World Scientific Publishing Company, River Edge, (2002).