OPTIMALITY CONDITIONS AND DUALITY RELATIONS IN NONCONVEX OPTIMIZATION

PhD Dissertation

Samet BİLA

Eskişehir 2022

OPTIMALITY CONDITIONS AND DUALITY RELATIONS IN NONCONVEX OPTIMIZATION

Samet BİLA

PhD DISSERTATION

Graduate School of Sciences
Department of Mathematics
Supervisor: Prof. Dr. Refail KASIMBEYLİ

Eskişehir
Eskişehir Technical University
Graduate School of Sciences
August 2022

FINAL APPROVAL FOR THESIS

This thesis titled "Optimality Conditions and Duality Relations in Nonconvex Optimization" has been prepared and submitted by Samet BILA in partial fulfillment of the requirements in "Eskişehir Technical University Directive on Graduate Education and Examination" for the Degree of Doctor of Philosophy (PhD) in Mathematics has been examined and approved on 12/08/2022.

Committee Members	**Title, Name and Surname**	**Signature**
Member :	Prof. Dr. Serkan Ali DÜZCE	....
Member :	Prof. Dr. Gürkan ÖZTÜRK	....
Member :	Asst. Prof. Dr. İlknur ATASEVER GÜVENÇ	....
Member :	Asst. Prof. Dr. Mustafa SOYERTEM	....
Member :	Asst. Prof. Dr. Ali Hakan TOR	....

Prof. Dr. Murat TANIŞLI
Director of Graduate School of Sciences

12/08/2022

SUPERVISOR APPROVAL

PhD student Samet BILA, whom I supervise, has completed this thesis titled OPTIMALITY CONDITIONS AND DUALITY RELATIONS IN NONCONVEX OPTIMIZATION. According to my inspections, the work is scientifically and ethically appropriate for the student to the thesis defense exam.

Supervisor
Prof. Dr. Refail KASIMBEYLI

ABSTRACT

OPTIMALITY CONDITIONS AND DUALITY RELATIONS IN NONCONVEX OPTIMIZATION

Samet BİLA
Department of Mathematics
Eskisehir Technical University, Graduate School of Sciences, August, 2022

Supervisor : Prof. Dr. Refail KASIMBEYLI

This thesis aims to establish conditions for the objective and the constraint functions such that the zero duality gap condition is satisfied without convexity assumptions. In the literature, the conditions given on the objective and the constraint functions for guaranteeing the zero duality gap condition are still an unsolved problem. We investigate the zero duality gap condition for a class of nonconvex optimization problems. In this study, a positively homogeneous and lower semicontinuous objective and constraint functions are defined on a conic set and the zero duality gap condition is obtained by showing the perturbation function is weakly subdifferentiable at the origin. Also, the weak subdifferentiability theorem is proven for the Lipschitz functions. This theorem is based on the separation theorem in the space where epigraph is defined. To prove this theorem augmented dual cones are extended to a higher dimension since epigraph is defined in $\mathbb{R}^{n+1}$. Additionally, we investigate the sum rule for the weak subdifferential. It is shown that the weak subdifferentials of some classes of the Clarke directionally differentiable functions and the tangentially convex functions, satisfy this property in the form of equality. A relationship between the weak subdifferential of the indicator function and the augmented normal cone to a nonconvex set is revealed.

Keywords: The weak subdifferential, Sum rule, Max relation, Nonconvex analysis, Zero duality gap.

ÖZET

KONVEKS OLMAYAN OPTİMİZASYONDA OPTİMALLİK KOŞULLARI VE DUALLİK İLİŞKİLERİ

Samet BİLA
Matematik Anabilim Dalı
Eskişehir Teknik Üniversitesi, Fen Bilimleri Enstitüsü, August, 2022

Danışman: Prof. Dr. Refail KASIMBEYLI

Bu tezin amacı, amaç ve kısıt fonksiyonları üzerine koşullar koyarak sıfır ikil boşluk koşulunu konvekslik koşulu olmadan inşa etmektir. Literatürde, amaç ve kısıt fonksiyonlarının üzerine koşullar konularak sıfır ikil boşluk koşulunu garantilemek hâlâ açık bir problemdir. Bu çalışmada, konveks olmayan optimizasyon problem sınıfı için koşullar konularak sıfır ikil boşluk koşulu elde edildi. Bu konveks olmayan optimizasyon problem sınıfında konik bir küme üzerinde tanımlı pozitif homojen ve alttan yarı sürekli amaç ve kısıt fonksiyonları ele alındı. Bu şartlar altında sıfır ikil boşluk pertürbasyon fonksiyonunun orijinde zayıf subdifferensiyellenebilirliği gösterilerek elde edildi. Ayrıca, Lipschitz fonksiyonları için zayıf subdiferansiyel teoremi ispatlandı. Bu teorem epigrafın tanımlı olduğu uzayda ayırma teoremine dayanmaktadır. Bu teoremi ispatlamak için epigraf $\mathbb{R}^{n+1}$ uzayında tanımlı olduğundan genişletilmiş dual koniler bir üst boyuta taşındı. Ek olarak, zayıf subdiferansiyel için toplam kuralı irdelendi. Clarke ve teğetsel yönlü türevlenenebilir fonksiyonların bazı sınıfları için bu toplam kuralının eşitlik halinde sağlandığı gösterildi. Bu çalışmada ayrıca, indikatör fonsiyonunun zayıf subdiferansiyeli ile konveks olmayan bir kümenin genişletilmiş normal konisi arasındaki ilişki gösterildi.

Anahtar Sözcükler: Zayıf subdiferansiyel, Toplam kuralı, Maksimum ilişkisi, Konveks olmayan analiz, Sıfır ikil boşluk.

ACKNOWLEDGEMENTS

I would like to use this opportunity to express my sincere thanks to my Supervisor Prof. Dr. Refail KASIMBEYLI. This thesis would not have been possible without his timely guidence, willingness to share his expertise and mathematical acumen.

I would like to also thank my parents for their lifelong support.

Samet BİLA

12/08/2022

STATEMENT OF COMPLIENCE WITH ETHICAL PRINCIPLES AND RULES

I hereby truthfully declare that this thesis is an original work prepared by me; that I have behaved in accordance with the scientific ethical principles and rules throughout the stages of preparation, data collection, analysis and presentation of my work; that I have cited the sources of all the data and information that could be obtained within the scope of this study, and included these sources in the references section; and that this study has been scanned for plagiarism with “scientific plagiarism detection program” used by Eskişehir Technical University, and that “it does not have any plagiarism” whatsoever. I also declare that, if a case contrary to my declaration is detected in my work at any time, I hereby express my consent to all the ethical and legal consequences that are involved.

Samet BİLA

CONTENTS

LIST OF FIGURES

GLOSSARY OF SYMBOLS AND ABBREVIATIONS

$\langle \cdot, \cdot \rangle$	: Inner product
$\| \cdot \|$	: Norm
$\mathbb{R}^n$	: n dimensional Euclidean space
$\mathbb{R}^n_+$	: The nonnegative orthant
$intS$	: Interior of S
clS	: Closure of S
$\partial f(\cdot)$	: The subdifferential of f
$\partial^o f(\cdot)$	: The Clarke subdifferential of f
$\partial^w f(\cdot)$	: The weak subdifferential of f
$B(x, \varepsilon)$	: A closed ball centered at x with radius ε
$N_C(\cdot)$	: Dual cone of set C
$N^a_C(\cdot)$	: Augmented dual cone of set C
$f^{'}(\cdot, \cdot)$	: The directional derivative of f
$f^o(\cdot, \cdot)$	: The Clarke directional derivative of f
δ_C	: Indicator function of set C
$epif$	: Epigraph of function f
$\mathscr{N}_\varepsilon(\overline{x})$	: ε neighborhood of $\overline{x}$

1. INTRODUCTION

Separation theorems using supporting hyperplane for convex sets at the boundary points are a basic mathematical tool that occupies a central role in convex analysis. It has numerous applications in many areas such as optimization, operations research, computer science, business, and economics. However, nonconvex sets do not hold such property at their boundary points. Thus, it is not possible to separate a point in the domain from the graph by using separating hyperplanes. Ekeland and Temam [17] considered the pointwise supremum of continuous affine functions and show some important properties that these functions possess. In this work, we take into account the pointwise supremum of continuous superlinear functions, the definition given earlier, and we derive similar properties as is done in the convex case. Therefore our functions may lead a way for examining the key ideas from convex analysis that is carried out in the nonconvex case.

The idea of supporting cones was extensively used to establish nonconvex separation theorems and to develop new tools for analyzing nonconvex optimization problems. In this thesis, a special type of functions is introduced, called superlinear functions. These are extensions of affine functions, with an norm function defined in reflexive Banach spaces. The main purpose of having such a family of superlinear functions is to build a theory to be able to separate a point from epigraph taken from the domain of a function. This separation is mainly about the cone separation theorem.

In this thesis the following optimization problem is considered: assume that S is a nonempty subset of $\mathbb{R}^n$, let $f : S \to \overline{\mathbb{R}}$, and $g = (g_1, \dots, g_m) : S \to \mathbb{R}^m$ then we consider the minimization problem

$$(P) \quad \begin{cases} \inf_{x \in \mathbb{R}^n} f(x) \end{cases}$$

where

$$\{x \in S : g_j(x) \leq 0, j = 1, \dots, m\} \neq \emptyset.$$

It is well know that duality theory is a fundamental branch of optimization. Dual problems of optimization take many researchers' interest. Dual problems play a crucial role in optimization to obtain optimality conditions. The beginning of this research dates back to Gale, Kuhn, and Tucker [18]. However, the dual problems is not as efficient if the zero duality gap is not satisfied. Perturbation function also has a significant role in duality and a well-organized scheme of primal and dual problems is constructed by Rockafellar and Wets [49, 11.40, 11.41]. The meticulous work done by Aubin and Ekeland [17] is related to nonconvex duality and duality for nonconvex problems of optimization. Aubin and Ekeland [17] constructed their work mainly by using the theory of abstract convexity. Duality theory in terms of general augmented Lagrangian was extensively investigated by Rockafellar and Wets [49]. Their work requires level-boundedness in x and locally uniformly in u condition. However, the work done by Rockafellar and Wets [49] lacks relating the zero duality gap condition directly to the objective and the constraint functions.

The perturbation function plays an important role in the study of the dual problem of optimization. For a given problem (P) the corresponding perturbation function is defined on $\mathbb{R}^m$ as

$$h(y) = \inf\{f(x) | x \in S, g(x) \leqq y\}.$$

In this work, a conic set, positively homogeneous and lower semicontinuous objective function, and positively homogeneous and lower semicontinuous constraint functions are considered. Then it is shown that the perturbation function h is weakly subdifferentiable at the origin. Thus the zero duality gap condition is satisfied under the given conditions.

From the perspective of optimization, the subdifferential notion of a convex function offers many convenient properties of the derivative. One can recall that at boundary points a convex set possesses a supporting hyperplane that leads to a subdifferential notion, denoted by ∂f, which constructs the core of the convex analysis. This notion was first introduced by R.T. Rockafellar, in his thesis in 1963 [44] for a convex function f. Then, F.H. Clarke in his thesis in 1973 [11] extended this definition to Lipschitz continuous function by introducing the Clarke subdifferential $\partial_o f$. Another extension of the subgradient notion to nonconvex functions, called the weak subdifferential $\partial^w f$, was introduced by R. Kasimbeyli in his thesis in 1992

[19]. The weak subdifferential is the generalization of the subdifferential notion of convex analysis to nonconvex functions. The idea of this notion is based on the idea of using supporting cones to the epigraph of a given function, which replaces in some sense the supporting hyperplanes used in convex analysis and was applied to investigate duality relations, optimality conditions, and solution methods in nonconvex mathematical programming.

Derivative is a powerful tool which was found widespread applications in optimization. One says that the function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is differentiable at $\overline{x}$, if there exists a gradient vector $\nabla f(\overline{x}) \in \mathbb{R}^n$ such that

$$\lim_{t \to 0} \frac{f(\overline{x} + th) - f(\overline{x})}{t} = \langle \nabla f(\overline{x}), h \rangle. \tag{1.1}$$

The existence of a gradient $\nabla f(\overline{x})$ at a point $\overline{x}$, is equivalent to the existence of a tangent plane $y = f(\overline{x}) + \langle \nabla f(\overline{x}), x - \overline{x} \rangle$ to the graph of f at $(\overline{x}, f(\overline{x}))$. In the case, if f is a convex and differentiable function, then the graph of this affine function becomes a supporting hyperplane to $epi f$ at $(\overline{x}, f(\overline{x}))$ and for every $x \in dom f$ one has

$$f(x) - f(\overline{x}) \geq \langle \nabla f(\overline{x}), x - \overline{x} \rangle. \tag{1.2}$$

The necessity of investigating the situations, when the differentiability condition can be considered too restrictive, has led to the development of the theory of generalized differentiation and hence to the development of the theory of nonsmooth analysis (see e.g. [1,11,16,21,28,37–40,43,48]). At the first step of such a generalization, the limit $t \to 0$ in (1.1), was relaxed by changing it to the one-sided form $t \downarrow 0$, which will be called in this thesis, a directional derivative $f'(\overline{x}; h)$ of f at $\overline{x}$ in a direction $h \in \mathbb{R}^n$:

$$f'(\overline{x}; h) = \lim_{t \downarrow 0} \frac{f(\overline{x} + th) - f(\overline{x})}{t},$$

if the limit exists. In this case, we will say that f is directionally differentiable function at $\overline{x}$ in the direction h.

The relation (1.2) was extended to a nondifferentiable case by Rockafellar, in his dissertation [44], where he introduced the subdifferential $\partial f(\overline{x})$ as a set of subgradient vectors $v \in \mathbb{R}^n$:

$$\partial f(\overline{x}) = \{v \in \mathbb{R}^n : \langle v, x - \overline{x}\rangle \leq f(x) - f(\overline{x}) \text{ for all } x\}. \tag{1.3}$$

This notion successfully extends the affine support relations given above for the gradient vector $\nabla f(\overline{x})$. The subdifferentiability of f at $\overline{x}$ means the existence of a vector $v \in \mathbb{R}^n$ such that the hyperplane

$$H(v, -1) = \{(x, y) : \langle (v, -1), (x - \overline{x}, y - f(\overline{x}))\rangle = 0\}$$

with normal vector $(v, -1)$, is a supporting hyperplane to the epigraph of f at $(\overline{x}, f(\overline{x}))$:

$$epi f \subset H^{-}(v, -1) = \{(x, y) : \langle (v, -1), (x - \overline{x}, y - f(\overline{x}))\rangle \leq 0\}. \tag{1.4}$$

The global nature of this relation has lead to the following necessary and sufficient condition for unconstrained global optimality of $\overline{x}$:

$$0 \in \partial f(\overline{x}) \Leftrightarrow f(x) \geq f(\overline{x}), \forall x \in \mathbb{R}^n. \tag{1.5}$$

Rules that relate ∂f to generalized directional derivatives of f, or allow ∂f to be expressed or estimated in terms of the subdifferentials of other functions (for example, when $f = f_1 + f_2$, etc.), comprise the subdifferential calculus. Such rules are used especially in analyzing optimality conditions (for example, like (1.5)). Rockafellar defined ∂f for such functions:

$$\partial f(\overline{x}) = \{v \in \mathbb{R}^n : f'(\overline{x}; h) \geq \langle v, h\rangle \text{ for all } h \in \mathbb{R}^n\}, \tag{1.6}$$

and showed how to characterize ∂f in terms of one-sided directional derivatives $f'(\overline{x}; h)$:

$$f'(\overline{x}; h) = \max\{\langle v, h\rangle : v \in \partial f(\overline{x})\} \text{ for all } h \in \mathbb{R}^n, \tag{1.7}$$

and proved that under mild restrictions the sum rule

$$\partial(f_1 + f_2)(\overline{x}) = \partial f_1(\overline{x}) + \partial f_2(\overline{x}), \tag{1.8}$$

is valid. This branch of convex analysis was developed further by Moreau, Rockafellar and others in the 1960's and applied to many kinds of optimization problems (cf. [9,22,23,45,46]).

A major contribution to the nonsmooth and nonconvex analysis was made by F.H. Clarke, who introduced a generalized directional derivative concept f° and showed how the definition of subdifferential ∂f could be extended to arbitrary lower semicontinuous, locally Lipschitz functions defined on Banach spaces [11].

Clarke defined the subdifferential $\partial^\circ f(\overline{x})$ as a set of subgradient vectors $v \in \mathbb{R}^n$ with

$$\partial^\circ f(\overline{x}) = \{v \in \mathbb{R}^n : f^\circ(\overline{x}; h) \geq \langle v, h \rangle, \forall h \in \mathbb{R}^n\}$$

and proved that if f attains a local minimum or maximum at $\overline{x}$ then $0 \in \partial^\circ f(\overline{x})$ [13, Proposition 2.3.2, p. 38]. Clarke introduced the regularity notion which plays an important role in nonsmooth analysis and allows to characterize the Clarke directional derivative f° in terms of the directional derivative f'.

Definition 1.1. [13, Definition 2.3.4, p. 39] f is said to be regular at x provided that the directional derivative $f'(x; h)$ exists and $f^\circ(x; h) = f'(x; h)$ for all h.

The following theorem for regularity was given in [13].

Theorem 1.2. [13, Proposition 2.3.6 (b), p. 40] *Let f be Lipschitz near x. If f is convex, then f is regular at x.*

The sum rule (1.8) for the subdifferential ∂f (of the convex analysis) was established for the Clarke subdifferential $\partial^\circ f$ in general case, only in the form of one sided inclusion:

$$\partial^\circ(f_1 + f_2)(x) \subset \partial^\circ f_1(x) + \partial^\circ f_2(x)$$

(see [13, Proposition 2.3.3, page 38]).

Note that, neither the graph of the directional derivative, nor the graph of the Clarke's directional derivative does not guarantee the existence of the supporting surface to the epigraph of a given function from below, similar to the relation given in (1.2) for the gradient vector in the convex case. Moreover, the Clarke subgradients do not provide an affine support relation similar to (1.4), in general.

The questions, "under which conditions does the affine support relation given in (1.4), or the global optimality condition (1.5), or the sum rule (1.8) can be extended to a nonconvex case", are very interesting and important from the perspective of developing optimality conditions in nonconvex optimization. But as it was pointed out by Rockafellar and Wets in [49, Commentaries to Ch.8, page 344], "applications not fitting within the domain of convexity had to wait for other ideas, however." The selection of a fitting supporting surface depends on the function under consideration. One of the such supporting surfaces fitting the extension of the affine support relation to a nonconvex case was suggested by Kasimbeyli in his thesis in 1992 [19] for lower Lipschitz functions. Kasimbeyli used special superlinear functions and by weakening the inequality used in the definition (1.3) of subgradients, introduced the notion of the weak subdifferential and extended it in this way to a nonconvex case.

This notion is based on the idea of using supporting cones to the epigraph of a function under consideration, which extends the affine support relation (1.4) to a cone support relation, which was applied to investigate duality relations, optimality conditions, and solution methods in nonconvex mathematical programming (see e.g. [3,5,6,10,14,15,20,32–34,51,52]). The idea of supporting cones was also used to establish nonconvex separation theorems and to develop new tools for analyzing nonconvex optimization problems [24–31].

Furthermore, the conditions for establishing the sum rule for the weak subdifferential which allows representing the weak subdifferential of the sum of two functions as a sum of the weak subdifferentials of these functions are proved. This representation is proved by using the so-called "sup relation" which expresses the classical directional derivative as a support function to the weak subdifferential (see [34, Theorem 1], or [33, Theorem 4.5]). Note that, the "sup relation" extends the "max relation" (1.7) of the convex analysis, which expresses the directional derivative as a support function to the subdifferential (1.6). Under some conditions that the "sup relation" for the weak subdifferential can be represented as a "max relation" for two classes of functions is shown: Clarke directionally differentiable functions and tangentially convex functions. This representation is used to prove the sum rule for the weak subdifferential. The sum rule is then applied to investigate relations between the weak subdifferential of the indicator function of a nonconvex set and the augmented normal cone to this set and to establish some additional properties of nonconvex sets.

2. THE WEAK SUBDIFFERENTIAL

In this chapter, we give the definitions we will be using throughout this thesis such as, the weak subdifferential and augmented normal cone and we investigate some properties of these notions.

Let $\mathbb{X}$ and $\mathbb{Y}$ be reflexive Banach space and their dual spaces are $\mathbb{X}^*$ and $\mathbb{Y}^*$ respectively. Consider the extended real line $\overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty\}$. Throughout this thesis, $\mathbb{R}^n$ denotes the Euclidean space with norm $\|x\| = \sqrt{\langle x, x\rangle}$ where $\langle \cdot, \cdot \rangle$ is the scalar product and $0_{\mathbb{R}^n}$ stands for the zero of $\mathbb{R}^n$. The unit sphere of $\mathbb{R}^n$ is denoted by $\mathbb{U}$ defined as:

$$\mathbb{U} = \{y \in \mathbb{R}^n : \|y\| = 1\}.$$

The unit ball of $\mathbb{R}^n$ is denoted by $\mathbb{B}$ defined as:

$$\mathbb{B} = \{y \in \mathbb{R}^n : \|y\| \leq 1\}.$$

Let $A \subset \mathbb{R}^n$ be a nonempty set. The *indicator function* δ_A of the set A, is defined as $\delta_A(x) = 0$ if $x \in A$, and $\delta_A(x) = +\infty$ if $x \notin A$. A function g is positively homogeneous whenever $g(\alpha x) = \alpha g(x)$, for each $\alpha > 0$ and for each $x \in \mathbb{R}^n$.

Definition 2.1. [7, page 55] The conjugate of the indicator function of a nonempty set $C \subset \mathbb{R}^n$, namely $\delta_C^* : \mathbb{R}^n \to \overline{\mathbb{R}}$, is called the support function of C.

Throughout this thesis, the set of nonnegative real numbers is denoted by $\mathbb{R}_+$. The *closure*, the *boundary*, the *interior* and the *convex hull* of a set S are denoted with the notations $\mathsf{cl}(S)$, $\mathsf{bd}(S)$, $\mathsf{int}(S)$, and $\mathsf{co}(S)$, respectively.

Definition 2.2. [7, page 44] The core of a set S is the set of points x in S such that for any direction d in $\mathbb{R}^n$, $x + td$ lies in S, for all small real numbers t.

The core of a set is different from the interior of the set. We provide the following example.

Example 2.3. Consider the set A in $\mathbb{R}^2$:

$$A = \{(x_1, x_2) | x_1 = 0 \text{ or } |x_1| \geq x_2^2\}$$

The set A is given in Figure 2.2. $\overline{\mathbf{x}} = (0,0) = \mathbf{0}$ belongs to the core(A) for any direction $\mathbf{d} \in \mathbb{R}^2$, $\overline{\mathbf{x}} + t\mathbf{d} = t\mathbf{d} \in A$ for all small real numbers t. However, there does not exist any neighborhood of $\overline{\mathbf{x}} = \mathbf{0}$ such that $\mathscr{N}(\mathbf{0})$ lies in A.

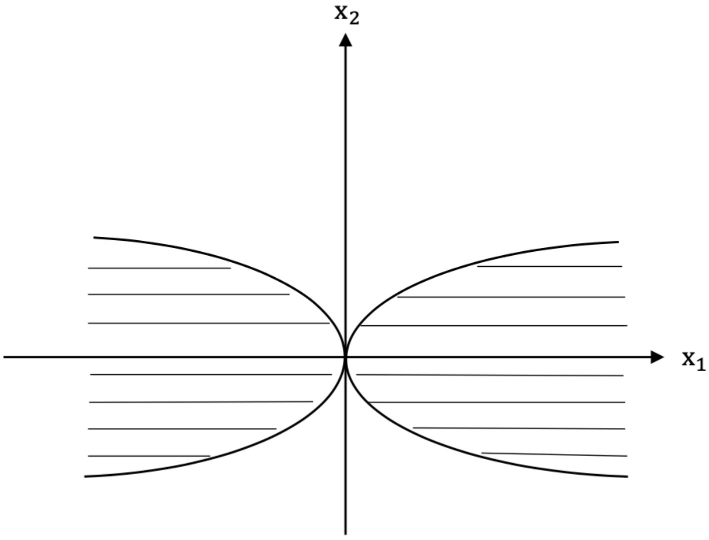

Figure 2.1. *The geometric interpretation of core given in Definition 2.2*

Definition 2.4. [33, Definition 4.8] A set $S \subseteq \mathbb{R}^n$ is called star-shaped with respect to a given point $\overline{x} \in S$ if it contains the line segment between each of its points and $\overline{x} \in S$ that is if $x \in S$, then

$$\lambda x + (1-\lambda)\overline{x} \in S \quad \text{for all } \lambda \in [0,1]. \tag{2.9}$$

Definition 2.5. A function $f : X \to \mathbb{R}$ is called lower locally Lipschitz at $\overline{x} \in X$, if there exists a positive number L and a neighborhood $\mathscr{N}(\overline{x})$ of $\overline{x}$ such that

$$f(x) - f(\overline{x}) \geq -L\|x - \overline{x}\| \text{ for all } x \in \mathscr{N}(\overline{x}). \tag{2.10}$$

f is called lower Lipschitz at $\overline{x}$ with the Lipschitz constant L if the inequality (2.10) holds for all $x \in X$.

Definition 2.6. [34, Definition 5] Let $S \subseteq \mathbb{R}^n$ and $\bar{x} \in S$. The normal cone to S at $\bar{x}$ defined as:

$$N_S(\bar{x}) = \{x^* \in \mathbb{R}^n : \langle x^*, x - \bar{x}\rangle \leq 0 \ \text{ for all } \ x \in S\}.$$

Assume that $S \subseteq \mathbb{R}^n$. Let $\bar{x} \in S$ with $S \setminus \{\bar{x}\} \neq \emptyset$. We define the augmented normal cone to S at $\bar{x}$ as follows:

$$N_S^a(\bar{x}) = \{(x^*, c) \in \mathbb{R}^n \times \mathbb{R}_+ : \langle x^*, x - \bar{x}\rangle + c\|x - \bar{x}\| \leq 0 \quad \forall x \in S\}.$$

Definition 2.7. [3, Definition 3.1] A pair $(x^*, c) \in \mathbb{R}^n \times \mathbb{R}_+$ is called a weak subgradient of f at $\overline{x}$ on S if

$$\langle x^*, x - \overline{x}\rangle - c\|x - \overline{x}\| \leq f(x) - f(\overline{x}), \quad \forall x \in S. \tag{2.11}$$

The set of all weak subgradients of f at $\overline{x}$ is called the weak subdifferential of f at $\overline{x}$ and is denoted $\partial_S^w f(\overline{x})$:

$$\partial_S^w f(\overline{x}) = \{(x^*, c) \in \mathbb{R}^n \times \mathbb{R}_+ : \text{ (2.11) is satisfied}\}. \tag{2.12}$$

If $\partial_S^w f(\overline{x}) \neq \emptyset$, then f is called the weakly subdifferentiable at $\overline{x}$. If $S = \mathbb{R}^n$, we omit the subscript S in $\partial_S^w f(\overline{x})$, that is: $\partial^w f(\overline{x}) = \partial_{\mathbb{R}^n}^w f(\overline{x})$.

It is clear that if f is subdifferentiable at $\bar{x}$ then f is also weakly subdifferentiable at $\bar{x}$; that is, if $x^* \in \partial f(\bar{x})$, then by definition $(x^*, c) \in \mathbb{R}^n \times \mathbb{R}_+$ for every $c \geq 0$. Now, we express the geometric interpretation of the weak subgradient of a function. The definition (2.11) implies that the pair $(x^*, c) \in X^* \times \mathbb{R}^+$ is a subgradient of f at $\bar{x} \in X$ if there is a continuous concave function

$$g(x) = \langle x^*, x - \bar{x}\rangle + f(\bar{x}) - c\|x - \bar{x}\|$$

such that $g(x) \leq f(x)$ for all $x \in X$ and $g(\bar{x}) = f(\bar{x})$. The set $\text{hypo}(g) = \{(x, a) \in X \times \mathbb{R} \mid g(x) \geq a\}$ is a closed cone in $X \times \mathbb{R}$ with its vertex at $(\bar{x}, f(\bar{x}))$. Indeed, we have

$$\begin{aligned}\text{hypo}(g) - (\bar{x}, f(\bar{x})) &= \{(x - \bar{x}, a - f(\bar{x})) \in X \times \mathbb{R} \mid -c\|x - \bar{x}\| + \langle x^*, x - \bar{x}\rangle \geq a - f(\bar{x})\} \\ &= \{(u, b) \in X \times \mathbb{R} | -c\|u\| + \langle x^*, u\rangle \geq b\}.\end{aligned}$$

Thus, it follows from (2.11) and (2) that the cone hypo(g) is a supporting cone of the set

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

at the point $(\bar{x}, f(\bar{x}))$ in the sense that $epi(f) \subset (X \times \mathbb{R}) \setminus \text{hypo}(g)$ and $cl(epi(f)) \cap graph(g) \neq \emptyset$ where $graph(g) = \{(x, a) \in X \times \mathbb{R} \mid g(x) = a\}$. The geometric interpretation of the weak subdifferential can be seen in Figure 2.7.

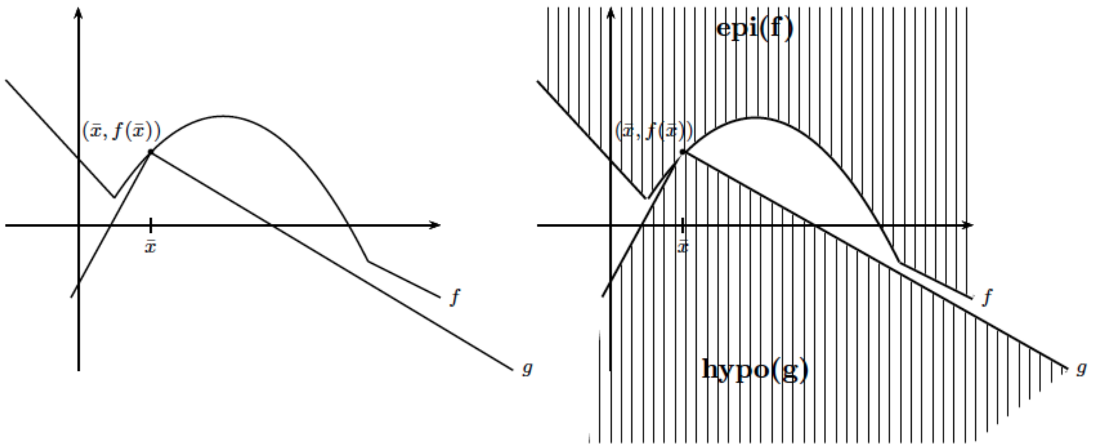

Figure 2.2. *The geometric interpretation of the weak subdifferential given in Definition 2.7*

In convex duality theory the conjugate, the biconjugate, the weak conjugate and the weak biconjugate are defined as follows:

Definition 2.8. [3, Definition 2.1] (a) Let X and X^* be two topological vector spaces placed in duality by the bilinear pairing denoted by $\langle \cdot, \cdot \rangle$. Let f be a function of X into $\overline{\mathbb{R}}$. Then the functions

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

and

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

are conjugate and biconjugate functions of f, respectively.

(b) A function $f^w : X \times X^* \times \mathbb{R}_+ \to \overline{\mathbb{R}}$ defined by

$$f^w(\bar{x}, x^*, c) = \sup_{x \in X} \{ -c\|x - \bar{x}\| + c\|\bar{x}\| + \langle x, x^* \rangle - f(x)\}$$

is called the weak conjugate of f.

A function f^{ww} on X defined by

$$f^{ww}(x) = \sup_{(\bar{x}, x^*, c) \in X \times X^* \times \mathbb{R}_+} \{ -c\|x - \bar{x}\| + c\|\bar{x}\| + \langle x, x^* \rangle - f^w(\bar{x}, x^*, c)\}$$

is called the weak biconjugate of f.

The geometric interpretatiton of the weak conjugate function f^w is that the graph of the function $g(x) = -c\|x - \bar{x}\| + c\|\bar{x}\| + \langle x, x^*\rangle$ is a closed cone with vertex at the point $(\bar{x}, c\|\bar{x}\| + \langle x, x^*\rangle) \in X \times \mathbb{R}$. In fact,

$$\begin{aligned} graph(g) - (\bar{x}, c\|\bar{x}\| + \langle x, x^*\rangle) =& \{(x - \bar{x}, \alpha - c\|\bar{x}\| - \langle \bar{x}, x^*\rangle) \in X \times \mathbb{R} | \\ & -\|x - \bar{x}\| + c\|\bar{x}\| + \langle x, x^*\rangle = \alpha\} \\ =& \{(y, \alpha) \in X \times \mathbb{R} \mid -c\|y\| + \langle y, x^*\rangle = \alpha\} \end{aligned}$$

The number $f^w(\bar{x}, x^*, c)$ is the supremum of the values β for which the cone $graph(g-\beta)$, intersects $epi(f)$. Thus, the cone $graph(g - f^w(\bar{x}, x^*, c))$ with vertex $(\bar{x}, c\|\bar{x}\| + \langle \bar{x}, x^*\rangle - f^w(\bar{x}, x^*, c))$ is a support cone of $epi(f)$ in the sense that $epi(f) \subseteq epi(g - f^w(\bar{x}, x^*, c))$, and $cl(epi(f)) \cap graph(g - f^w(\bar{x}, x^*, c)) \neq \emptyset$. The $graph(g - f^w(\bar{x}, x^*, c))$ intersects vertical axis (i.e $x = 0$) at $(0, -f^w(\bar{x}, x^*, c))$.

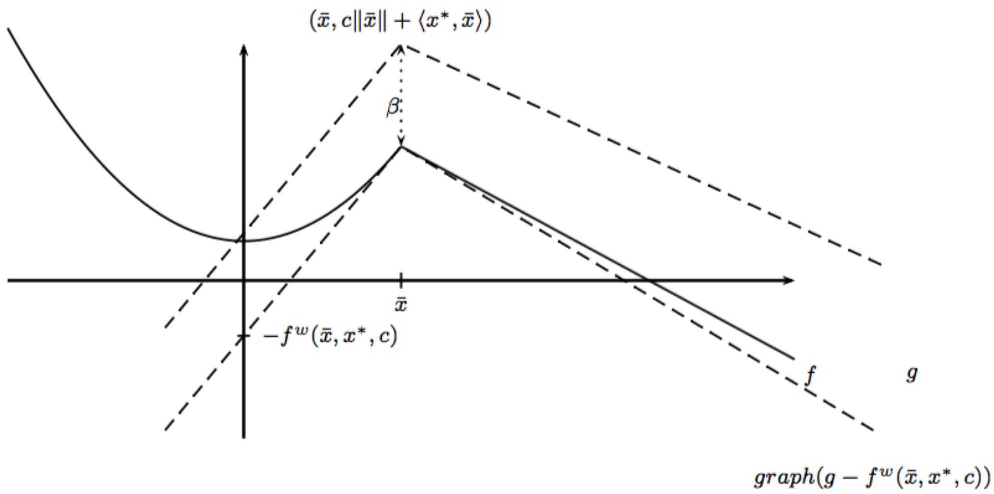

Figure 2.3. *The geometric interpretation of the weak conjugate given in Definition 2.8*

It is clear from (2.7) that the pair (x^*, c) is a weak subgradient of f at $\overline{x} \in S$ if the continuous function $\varphi(x)$ defined by

$$\varphi(x) = \langle x^*, x - \overline{x}\rangle + f(\overline{x}) - c\,\|x - \overline{x}\|\,, \tag{2.13}$$

satisfies $\varphi(x) \leq f(x)$ for all $x \in S$ and $\varphi(\overline{x}) = f(\overline{x})$. Hence, it follows that, if f is weakly subdifferentiable at $\overline{x}$ and $(x^*, c) \in \partial_S^w f(\overline{x})$, then the graph of function

φ defined in (2.13) becomes a supporting surface to the epigraph of f at the point $(\overline{x}, f(\overline{x}))$:

$$epi(f) \subset epi(\varphi) \text{ and } epi(f) \cap \mathsf{cl}(graph(\varphi)) \neq \emptyset.$$

The geometric interpretation of the weak conjugate is provided in Figure 2.8.

Remark 1. It is clear that if f is subdifferentiable at $\overline{x}$ (in the sense of convex analysis) then f is also weakly subdifferentiable at $\overline{x}$, that is if $v \in \partial f(\overline{x})$, then by definition, $(v, c) \in \partial^w f(\overline{x})$ for every $c \geq 0$. The famous optimality condition of convex analysis states that if the function f is a convex, subdifferentiable at $\overline{x}$, then f attains a global minimum at $\overline{x}$ if and only if $0 \in \partial f(\overline{x})$. Similarly, the weak subdifferential implies that, the global optimality condition (1.5) can easily be extended to a nonconvex case as:

$$(0_{\mathbb{R}^n}, 0_{\mathbb{R}}) \in \partial^w f(\overline{x}) \Leftrightarrow f(x) \geq f(\overline{x}), \forall x \in \mathbb{R}^n.$$

The following examples are given to demonstrate the definition of the weak subdifferentiable. We will revisit some of these examples in the following Chapter.

Example 2.9. Define $f : \mathbb{R} \to \mathbb{R}$ by

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

We are interested in the weak subdifferentiable of f at $\overline{x} = 0$. From Figure 2.4 it is obvious that there does not exists a supporting hyperplane that support epi(f) at $x = 0$. However, there does exists a supporting conic surface support epi(f) at $x = 0$.

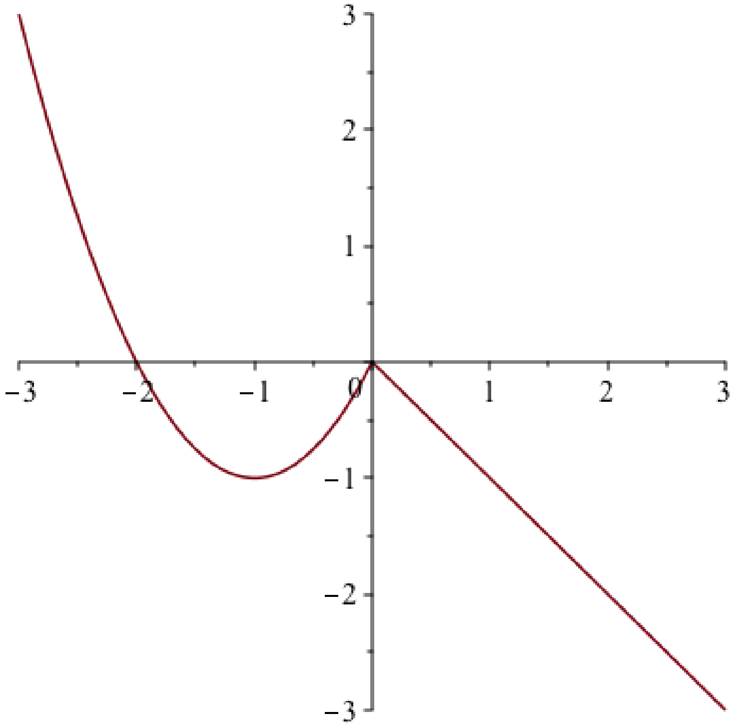

Figure 2.4. *The graph of $f(x)$ given in Example 2.9*

First we consider the case when $x \leq 0$. Then the weak subdifferential definition implies that:

$$\begin{aligned}
&\langle v, x-0\rangle - c|x-0| \leq f(x) - f(0),\\
&vx - c|x| \leq (x+1)^2 - 1 - 0,\\
&(v+c)x \leq x(x+2).
\end{aligned}$$

Then $\partial^w f(0)$ when $x \leq 0$ is equivalent to:

$$v + c \geq 2.$$

Now we consider the case $x > 0$. Then again the weak subdifferential definition yields to:

$$\begin{aligned}
&\langle v, x-0\rangle - c|x-0| \leq f(x) - f(0),\\
&vx - cx \leq -x,\\
&x(v-c) \leq -x.
\end{aligned}$$

Then $\partial^w f(0)$ when $x > 0$ calculated as:

$$v - c \leq -1.$$

Hence,

$$\partial^w f(0) = \{(v, c) \in \mathbb{R} \times \mathbb{R}^+ : -c + 2 \leq v \leq c - 1\}.$$

Now present the second example of the weak subdifferential.

Example 2.10. Let $h : \mathbb{R} \to \mathbb{R}$ defined as:

$$h(x) = \begin{cases} (x+1)^2 & \text{if } x \leq 1, \\ (x-2)^2 + 3 & \text{if } x > 1. \end{cases}$$

We want to obtain the weak subdifferentiable of f at $\overline{x} = 1$. Now we consider the case $x \leq 1$. The weak subdifferential definition yields to:

$$\begin{aligned} &\langle v, x - \overline{x} \rangle - c\|x - \overline{x}\| \leq h(x) - h(\overline{x}), \\ &\langle v, x - 1 \rangle - c|x - 1| \leq (x+1)^2 - f(1), \\ &v(x-1) - c(1-x) \leq (x+1)^2 - 4, \\ &(x-1)(v+c) \leq x^2 + 2x - 3. \end{aligned}$$

Then $\partial^w h(1)$ when $x \leq 1$ obtained as:

$$(v + c) \geq 4.$$

Now consider the case $x > 1$.

Then again by using the definition, we get:

$$\begin{aligned} &\langle v, x - \overline{x} \rangle - c\|x - \overline{x}\| \leq h(x) - h(\overline{x}), \\ &\langle v, x - 1 \rangle - c|x - 1| \leq h(x) - h(1), \\ &v(x-1) - c(x-1) \leq (x-2)^2 + 3 - 4, \\ &(x-1)(v-c) \leq x^2 - 4x + 3. \end{aligned}$$

Then $\partial^w h(1)$ when $x > 1$ calculated as:

$$(v - c) \leq -2.$$

Finally,

$$\partial^w h(1) = \{(v,c) \in \mathbb{R} \times \mathbb{R}^+ : -c + 4 \leq v \leq c - 2\}.$$

Now another example of the weak subdifferential is presented.

Example 2.11. Define $g : \mathbb{R} \to \mathbb{R}$ by

$$g(x) = \begin{cases} |x+2| & \text{if } x \leq 0, \\ 4x + 2 & \text{if } 0 < x \leq 1, \\ 2|x-2| + 4 & \text{if } x > 1. \end{cases}$$

We want to get the weak subdifferentiable of g at $\overline{x} = 1$. We consider the case $x \leq 1$. Then the weak subdifferential definition implies that:

$$\begin{aligned}
&\langle v, x - \overline{x} \rangle - c\|x - \overline{x}\| \leq g(x) - g(\overline{x}), \\
&\langle v, x - 1 \rangle - c|x - 1| \leq 4x + 2 - g(1), \\
&v(x-1) - c(1-x) \leq 4x + 2 - 6, \\
&(x-1)(v+c) \leq 4(x-1).
\end{aligned}$$

Then $\partial^w g(1)$ when $x \leq 1$ obtained as:

$$(v + c) \geq 4.$$

Now if we consider the second case when $1 < x < 2$.
It follows that:

$$\begin{aligned}
&\langle v, x - \overline{x} \rangle - c\|x - \overline{x}\| \leq g(x) - g(\overline{x}), \\
&\langle v, x - 1 \rangle - c|x - 1| \leq g(x) - g(1), \\
&v(x-1) - c(x-1) \leq 2|x-2| + 4 - 6, \\
&(x-1)(v-c) \leq 2(2-x) - 2 = -2(x-1).
\end{aligned}$$

Then $\partial^w h(1)$ when $1 < x < 2$ calculated as:

$$(v - c) \leq -2.$$

The last case is when $x \geq 2$ calculated as:

$$
\begin{aligned}
&\langle v, x-\overline{x}\rangle - c\|x-\overline{x}\| \leq g(x) - g(\overline{x}),\\
&\langle v, x-1\rangle - c|x-1| \leq g(x) - g(1),\\
&v(x-1) - c(x-1) \leq 2|x-2| + 4 - 6,\\
&(x-1)(v-c) \leq 2(x-2) - 2 = 2(x-3).
\end{aligned}
$$

Then $\partial^w h(1)$ when $x \geq 2$ calculated as:

$$(v-c) \leq 2.$$

Thus we obtain:

$$\partial^w g(1) = \{(v,c) \in \mathbb{R} \times \mathbb{R}_+ : -c+4 \leq v \leq c-2\}.$$

The following lemma is given in [34].

Lemma 2.12. *Let $f : S \to \mathbb{R}$ be a given function. Assume that the weak subdifferential $\partial^w_S f(\overline{x}) \neq \emptyset$. Then the weak subdifferential $\partial^w_S f(\overline{x})$ is convex and closed.*

2.1. Properties of the Weak Subdifferential

In this subsection we investigate the properties of the weak subdifferential and augmented normal cone. The properties are known to hold for the classical subdifferential and normal cone.

Proposition 2.13. *(Scalar Multiples) $\partial^w(\alpha f(\bar{x})) = \alpha \partial^w f(\bar{x})$ where $\alpha > 0$.*

Proof.

$$
\begin{aligned}
&\iff (x^*, \beta) \in \partial^w(\alpha f(\bar{x})),\\
&\iff \langle x^*, x-\bar{x}\rangle - \beta\|x-\bar{x}\| \leq (\alpha f)(x) - (\alpha f)(\bar{x}) \qquad \text{for all } x \in \mathbb{R}^n,\\
&\iff \langle x^*, x-\bar{x}\rangle - \beta\|x-\bar{x}\| \leq \alpha(f(x) - f(\bar{x})) \qquad \text{for all } x \in \mathbb{R}^n,\\
&\iff \langle \frac{x^*}{\alpha}, x-\bar{x}\rangle - \frac{\beta}{\alpha}\|x-\bar{x}\| \leq f(x) - f(\bar{x}) \qquad \text{for all } x \in \mathbb{R}^n,\\
&\iff (\frac{x^*}{\alpha}, \frac{\beta}{\alpha}) \in \partial^w(f(\bar{x})),\\
&\iff \frac{1}{\alpha}(x^*, \beta) \in \partial^w(f(\bar{x})),\\
&\iff (x^*, \beta) \in \alpha\partial^w(f(\bar{x})).
\end{aligned}
$$

The proof is completed. □

Proposition 2.14. $\partial^w(f(x)+c) = \partial^w(f(x))$

Proof.

$$\begin{aligned}
&\iff (x^*, \alpha) \in \partial^w(f(x)+c), \\
&\iff \langle x^*, x-\bar{x}\rangle - \alpha\|x-\bar{x}\| \leq f(x)+c-(f(\bar{x})+c) \qquad \text{for all } x \in \mathbb{R}^n, \\
&\iff \langle x^*, x-\bar{x}\rangle - \alpha\|x-\bar{x}\| \leq f(x)-f(\bar{x}) \qquad \text{for all } x \in \mathbb{R}^n, \\
&\iff (x^*, \alpha) \in \partial^w(f(\bar{x})).
\end{aligned}$$

which completes the proof. □

Proposition 2.15. *Let $a \in \mathbb{R}^n$, then $\partial^w(f(x) - \langle a, x\rangle) = \partial^w(f(x)) - \{(a, 0)\}$ for all $x \in \mathbb{R}^n$.*

Proof.

$$\begin{aligned}
&\iff (x^*, \alpha) \in \partial^w(f(x) - \langle a, x\rangle), \\
&\iff \langle x^*, x-\bar{x}\rangle - \alpha\|x-\bar{x}\| \leq f(x) - \langle a, x\rangle - f(\bar{x}) + \langle a, \bar{x}\rangle \qquad \text{for all } x \in \mathbb{R}^n, \\
&\iff \langle x^*, x-\bar{x}\rangle + \langle a, x-\bar{x}\rangle - \alpha\|x-\bar{x}\| \leq f(x)-f(\bar{x}) \qquad \text{for all } x \in \mathbb{R}^n, \\
&\iff \langle x^*+a, x-\bar{x}\rangle - \alpha\|x-\bar{x}\| \leq f(x)-f(\bar{x}) \qquad \text{for all } x \in \mathbb{R}^n, \\
&\iff (x^*+a, \alpha) \in \partial^w(\ f(\bar{x})), \\
&\iff (a, 0) + (x^*, \alpha) \in \partial^w(\ f(\bar{x})), \\
&\iff (x^*, \alpha) \in \partial^w(f(x)) - \{(a, 0)\}.
\end{aligned}$$

the proof is completed. □

The augmented normal cone satisfies the following property.

Proposition 2.16. *For any $(\bar{x}, \bar{r}) \in epi(f)$,*

$$N^a_{epif}(\bar{x}, \bar{r}) \subset N^a_{epif}(\bar{x}, f(\bar{x})).$$

Proof. Let $((x^*, \alpha), c) \in N^a_{epif}(\bar{x}, \bar{r})$. It implies that

$$\langle (x^*, \alpha), (x, r) - (\bar{x}, \bar{r})\rangle + c\|(x, r) - (\bar{x}, \bar{r})\| \leq 0, \quad \forall (x, r) \in epif.$$

Since $(\bar{x}, \bar{r}) \in epi(f)$, we can choose $\bar{r} = f(\bar{x})$ and this yields to,

$$\langle (x^*, \alpha), (x, r) - (\bar{x}, f(\bar{x}))\rangle + c\|(x, r) - (\bar{x}, f(\bar{x}))\| \leq 0, \quad \forall (x, r) \in epif.$$

It implies that $((x^*, \alpha), c) \in N^a_{epif}(\bar{x}, f(\bar{x}))$.
The proof is completed. $\square$

$\partial^w f$ can be considered as a set-valued mapping from $\mathbb{R}^n$ to $\mathbb{R}^n \times \mathbb{R}^+$. The next proposition proves the upper semicontinuity property for the weak subdifferential mapping when it is considered as a set-valued mapping. This property shows that the weak subdifferential mapping has a closed graph.

Proposition 2.17. *(Upper semicontinuity of the weak subdifferential) Let f be a weakly differentiable and lower semicontinuous function. Then $\partial^w f$ is an upper semicontinuous set-valued mapping from $\mathbb{R}^n$ to $\mathbb{R}^n \times \mathbb{R}^+$.*

Proof. Assume that (x_n^*, c_n) belongs to $\partial^w f(\bar{x})$, then it implies that

$$\langle x_n^*, x - x_n\rangle - c_n\|x - x_n\| \leq f(x) - f(x_n) \text{ for all } x \in X. \tag{2.14}$$

Since f is lower semicontinuous at $\overline{x}$, then

$$\liminf_{x_n \to \overline{x}} f(x_n) \geq f(\overline{x}) \text{ for all } x_n \to \overline{x},$$

or equivalently,

$$-\liminf_{x_n \to \overline{x}} f(x_n) \leq -f(\overline{x}) \text{ for all } x_n \to \overline{x}. \tag{2.15}$$

Taking the limit of (2.14) as $n \to \infty$ and taking into account (2.15) yield to

$$\langle \overline{x}^*, x - \overline{x}\rangle - \overline{c}\|x - \overline{x}\| \leq f(x) - \liminf f(x_n) \leq f(x) - f(\overline{x}) \text{ for all } x \in X.$$

It implies that $(\bar{x}^*, \bar{c})$ belongs to $\partial^w f$. Hence $\partial^w f$ is an upper semicontinuous set-valued mapping. $\square$

The next proposition presents a special type of monotonicity for the weak subdifferential mapping. This property weakens the usual monotonicity property earlier given in the literature for the classical subdifferential and hence extends the class of functions having this property. This is similar to the definition of the weak subdifferential which weakens the usual definition of the classical subdifferential and by this way extends the class of weakly subdifferentiable functions. Due to this extension the weak subdifferential makes it possible to investigate a large class of nonconvex problems.

Proposition 2.18. *(Monotonicity property for the weak subdifferential) Let f be a weakly differentiable function, then*

$$\langle x_1^* - x_2^*, x_1 - x_2 \rangle + (c_1 + c_2)\|x_1 - x_2\| \geq 0,$$

for all $x_1, x_2 \in \mathbb{R}^n, (x_1^*, c_1) \in \partial^w f(x_1), (x_2^*, c_2) \in \partial^w f(x_2)$.

Proof. Let $(x_1^*, c_1) \in \partial^w f(x_1), (x_2^*, c_2) \in \partial^w f(x_2)$. Then, by the definition of the weak subdifferential we deduce:

$$f(x_2) - f(x_1) \geq \langle x_1^*, x_2 - x_1 \rangle - c_1\|x_2 - x_1\|,$$

and

$$f(x_1) - f(x_2) \geq \langle x_2^*, x_1 - x_2 \rangle - c_2\|x_1 - x_2\|.$$

By adding these inequalities side-by-side, we obtain:

$$0 \geq -\langle x_1^*, x_1 - x_2 \rangle - c_1\|x_1 - x_2\| + \langle x_2^*, x_1 - x_2 \rangle - c_2\|x_1 - x_2\|$$

or

$$0 \geq -\langle x_1^* - x_2^*, x_1 - x_2 \rangle - (c_1 + c_2)\|x_1 - x_2\|$$

which leads to the required relation. □

The following two propositions present new properties for the weak conjugate function. The weak conjugate and biconjugate functions were first introduced by Kasimbeyli in [3] and investigated in detail in [3,4,14]. These functions play an important role in developing and investigating duality relations and strong zero duality gap conditions in nonconvex optimization.

Proposition 2.19. $(\lambda f)^w(x_0, x^*, c) = \lambda f^w(x_0, \frac{x^*}{\lambda}, \frac{c}{\lambda})$ *where λ is a positive real number.*

Proof. Let $g = \lambda f$, then by definition of the weak conjugate we get

$$\begin{aligned} g^w(x_0, x^*, c) &= \sup_{x\in X}\{-c\|x - x_0\| + c\|x_0\| + \langle x^*, x\rangle - g(x)\}, \\ &= \sup_{x\in X}\{-c\|x - x_0\| + c\|x_0\| + \langle x^*, x\rangle - \lambda f(x)\}, \\ &= \lambda \sup_{x\in X}\{-\frac{c}{\lambda}\|x - x_0\| + \frac{c}{\lambda}\|x_0\| + \langle \frac{x^*}{\lambda}, x\rangle - f(x)\}, \\ &= \lambda f^w(x_0, \frac{x^*}{\lambda}, \frac{c}{\lambda}). \end{aligned}$$

□

Proposition 2.20. *For every $\alpha \in \mathbb{R}$, we have $(f+\alpha)^w(x_0, x^*, c) = f^w(x_0, x^*, c) - \alpha$*

Proof. Let $g = f + \alpha$, then by definition of the weak conjugate of function,

$$\begin{aligned} g^w(x_0, x^*, c) &= \sup_{x\in X}\{-c\|x - x_0\| + c\|x_0\| + \langle x^*, x\rangle - g(x)\}, \\ &= \sup_{x\in X}\{-c\|x - x_0\| + c\|x_0\| + \langle x^*, x\rangle - f(x) - \alpha\}, \\ &= \sup_{x\in X}\{-c\|x - x_0\| + c\|x_0\| + \langle x^*, x\rangle - f(x)\} - \alpha, \\ &= f^w(x_0, x^*, c) - \alpha. \end{aligned}$$

□

2.2. Augmented Dual Cones in $\mathbb{R}^n \times \mathbb{R}$

In this subsection, we calculate the augmented dual cones in $\mathbb{R}^n \times \mathbb{R}$. This calculation helps us with the proof of the separation theorem presented in the following subsection.

Let $\mathbb{Y}$ be a reflexive Banach space with dual space $\mathbb{Y}^*$. The unit sphere and the unit ball of $\mathbb{Y}$ are denoted by $\mathbb{U} = \{y \in \mathbb{Y} : \|y\| = 1\}$ and $\mathbb{B} = \{y \in \mathbb{Y} : \|y\| \leq 1\}$, respectively. A nonempty subset $\mathbb{C}$ of $\mathbb{Y}$ is called a cone if $y \in \mathbb{C}$, $\lambda \geq 0 \implies \lambda y \in \mathbb{C}$. $\mathbb{C}$ is called pointed cone if $\mathbb{C} \cap -\mathbb{C} = \{0_{\mathbb{Y}}\}$. $cone(S) = \{\lambda s : \lambda \geq 0 \text{ and } s \in S\}$ denotes the cone generated by a set S.

The set $\mathbb{C}_{\mathbb{U}} = \mathbb{C} \cap \mathbb{U} = \{y \in \mathbb{C} : \|y\| = 1\}$ is called a norm-base of cone $\mathbb{C}$. A nonempty subset S of $\mathbb{R}^n$ is called cone-shaped at $\overline{x} \in S$, if $cl(cone(S - \overline{x})) \neq \mathbb{R}^n$.

The dual cone of $\mathbb{C}^*$ of $\mathbb{C}$ and its quasi-interior are defined by

$$\mathbb{C}^* = \{y^* \in \mathbb{Y}^* : \langle y^*, y\rangle \geq 0 \text{ for all } y \in \mathbb{C}\}$$

and

$$\mathbb{C}^\# = \{y^* \in \mathbb{Y}^* : \langle y^*, y\rangle > 0 \text{ for all } y \in \mathbb{C} \setminus \{0_\mathbb{Y}\}\}, \text{ respectively.}$$

The following definitions of augmented dual cone of $\mathbb{C}$ are introduced in [29],

$$\mathbb{C}^{a*} = \{(y^*, \alpha) \in \mathbb{C}^\# \times \mathbb{R}_+ : \langle y^*, y\rangle - \alpha\|y\| \geq 0 \text{ for all } y \in \mathbb{C}\}.$$

and

$$\mathbb{C}^\# = \{(y^*, \alpha) \in \mathbb{C}^\# \times \mathbb{R}_+ : \langle y^*, y\rangle - \alpha\|y\| \geq 0 \text{ for all } y \in \mathbb{C} \setminus \{0_\mathbb{Y}\}\}.$$

Similarly, we can rewrite all the definitions for $\mathbb{R}^{n+1}$ as follows: The set

$$\mathbb{C}_\mathbb{U} = \mathbb{C} \cap \mathbb{U} = \{(y, a) \in \mathbb{C} : \|(y, a)\| = 1\}$$

is called a norm-base of cone $\mathbb{C}$. The dual cone $\mathbb{C}^*$ of $\mathbb{C}$ and its quasi-interior are defined by

$$\mathbb{C}^* = \{(y^*, a^*) \in \mathbb{Y}^* : \langle (y^*, a^*), (y, a)\rangle \geq 0 \text{ for all } (y, a) \in \mathbb{C}\}$$

and

$$\mathbb{C}^\# = \{(y^*, a^*) \in \mathbb{Y}^* : \langle (y^*, a^*), (y, a)\rangle > 0 \text{ for all } (y, a) \in \mathbb{C} \setminus \{0_\mathbb{Y}\}\}, \text{ respectively.}$$

The following augmented dual definitions are introduced in [29],

$$\mathbb{C}^{a*} = \{((y^*, a^*), \alpha) \in \mathbb{C}^\# \times \mathbb{R}_+ : \langle (y^*, a^*), (y, a)\rangle - \alpha\|(y, a)\| \geq 0 \text{ for all } (y, a) \in \mathbb{C}\},$$

and

$$\mathbb{C}^{a\#} = \{((y^*, a^*), \alpha) \in \mathbb{C}^\# \times \mathbb{R}_+ : \langle (y^*, a^*), (y, a)\rangle - \alpha\|(y, a)\| > 0 \text{ for all } (y, a) \in \mathbb{C} \setminus \{0_\mathbb{Y}\}\}.$$

The augmented dual cones of $\mathbb{C} = \mathbb{R}^{n+1}_+$ is calculated as follows:

Example 2.21. If $\mathbb{C} = \mathbb{R}_+$ then, the dual cone $\mathbb{C}^*$ of $\mathbb{C}$ and its quasi-interior, respectively are defined by

$$\begin{aligned}\mathbb{C}^* &= \{y^* \in \mathbb{Y}^* : \langle y^*, y\rangle \geq 0 \text{ for all } y \in \mathbb{C}\},\\ &= \{y^* \in \mathbb{Y}^* : \langle y^*, y\rangle \geq 0 \text{ for all } y \geq 0\},\\ &= \{y^* \in \mathbb{Y}^* : y^* \geq 0\}.\end{aligned}$$

and

$$\begin{aligned}\mathbb{C}^\# &= \{y^* \in \mathbb{Y}^* : \langle y^*, y\rangle \geq 0 \text{ for all } y \in \mathbb{C} \setminus \{0\}\},\\ &= \{y^* \in \mathbb{Y}^* : \langle y^*, y\rangle > 0 \text{ for all } y > 0\},\\ &= \{y^* \in \mathbb{Y}^* : y^* > 0\}.\end{aligned}$$

Similarly augmented dual cones of $\mathbb{C}$ can be derived as:

$$\begin{aligned}\mathbb{C}^{a*} &= \{(y^*, \alpha) \in \mathbb{Y}^* \times \mathbb{R}_+ : \langle y^*, y\rangle - \alpha\|y\| \geq 0 \text{ for all } y \in \mathbb{C}\},\\ &= \{(y^*, \alpha) \in \mathbb{Y}^* \times \mathbb{R}_+ : y(y^* - \alpha) \geq 0 \text{ for all } y \geq 0\},\\ &= \{(y^*, \alpha) \in \mathbb{Y}^* \times \mathbb{R}_+ : y^* \geq \alpha \geq 0\}.\end{aligned}$$

and

$$\begin{aligned}\mathbb{C}^{a\#} &= \{(y^*, \alpha) \in \mathbb{Y}^* \times \mathbb{R}_+ : \langle y^*, y\rangle - \alpha\|y\| > 0 \text{ for all } y \in \mathbb{C} \setminus \{0_{\mathbb{Y}}\}\},\\ &= \{(y^*, \alpha) \in \mathbb{Y}^* \times \mathbb{R}_+ : y(y^* - \alpha) > 0 \text{ for all } y > 0\},\\ &= \{(y^*, \alpha) \in \mathbb{Y}^* \times \mathbb{R}_+ : y^* > \alpha \geq 0\}.\end{aligned}$$

We calculate the augmented dual cones of $\mathbb{C} = \mathbb{R}^{n+1}_+$ as follows:

$$\begin{aligned}\mathbb{C}^{a*} = \{&(y^*, x^*, \alpha) \in \mathbb{Y}^* \times \mathbb{R}_+ \times \mathbb{R}_+ : \langle (y^*, x^*), (y, x)\rangle - \alpha\|(y, x)\| \geq 0,\\ &\text{for all } (y, x) \in \mathbb{R}^n_+ \times \mathbb{R}_+\},\\ &= \{(y^*, x^*, \alpha) \in \mathbb{Y}^* \times \mathbb{R}_+ \times \mathbb{R}_+ : \langle y^*, y\rangle + \langle x^*, x\rangle - \alpha\|y\| - \alpha\|x\| \geq 0,\\ &\forall (y, x) \in \mathbb{R}^n_+ \times \mathbb{R}_+\},\\ &= \{(y^*, x^*, \alpha) \in \mathbb{Y}^* \times \mathbb{R}_+ \times \mathbb{R}_+ : (\langle y^*, y\rangle - \alpha\|y\|) + (\langle x^*, x\rangle - \alpha\|x\|) \geq 0,\\ &\forall (y, x) \in \mathbb{R}^n_+ \times \mathbb{R}_+\},\\ &= \{(y^*, x^*, \alpha) \in \mathbb{Y}^* \times \mathbb{R}_+ \times \mathbb{R}_+ : y^*_i \geq \alpha \geq 0, i = 1, \ldots, n \text{ and } x^* \geq \alpha \geq 0\}.\end{aligned}$$

Similarly,

$$
\begin{aligned}
\mathbb{C}^{a\#} \quad &= \{(y^*,x^*,\alpha) \in \mathbb{Y}^* \times \mathbb{R}_+ \times \mathbb{R}_+ : \langle (y^*,x^*),(y,x)\rangle - \alpha\|(y,x)\| > 0, \\
&\quad \text{for all } (y,x) \in \mathbb{R}^n_+ \times \mathbb{R}_+\}, \\
&= \{(y^*,x^*,\alpha) \in \mathbb{Y}^* \times \mathbb{R}_+ \times \mathbb{R}_+ : \langle y^*,y\rangle + \langle x^*,x\rangle - \alpha\|y\| - \alpha\|x\| > 0, \\
&\quad \text{for all } (y,x) \in \mathbb{R}^n_+ \times \mathbb{R}_+\}, \\
&= \{(y^*,x^*,\alpha) \in \mathbb{Y}^* \times \mathbb{R}_+ \times \mathbb{R}_+ : (\langle y^*,y\rangle - \alpha\|y\|) + (\langle x^*,x\rangle - \alpha\|x\|) > 0, \\
&\quad \text{for all } (y,x) \in \mathbb{R}^n_+ \times \mathbb{R}_+\}, \\
&= \{(y^*,x^*,\alpha) \in \mathbb{Y}^* \times \mathbb{R}_+ \times \mathbb{R}_+ : y_i^* > \alpha \geq 0, i = 1,\ldots,n \text{ and } x^* > \alpha \geq 0\}.
\end{aligned}
$$

2.3. The Weak Subdifferential Theorem for Lipschitz Functions

Let V be a real normed space and its dual V^*. One of the main problems related to the theory of abstract convexity is to derive conditions guaranteeing that the weak subdifferentiability of a function $f : V \to \mathbb{R}$ at a given point is not empty. Instead of dealing with linear functions, we consider $g_{(x_i^*,c_i,x_i,\alpha_i)} : V^* \times \mathbb{R}^+ \times V \times \mathbb{R} \to \mathbb{R}$ type of functions and we call these functions superlinear. We show that the pointwise supremum of superlinear functions at a given point is weakly subdifferentiable under some conditions.

The following is the definition of superlinear functions.

Definition 2.22. The set of functions $f : \mathbb{V} \subseteq \mathbb{R}^n \to \overline{\mathbb{R}}$ which are pointwise supremum of a family of functions $g_{(x_i^*,c_i,x_i,\alpha_i)}(x) = \langle x_i^*, x - x_i\rangle - c_i\|x - x_i\| + \alpha_i$, $i \in I$ is denoted by $S(V)$ and we call these superlinear functions. $S_0(V)$ is the subset of $S(V)$ which consists of lower Lipschitz functions.

We question whether the superlinear functions are continuous. In the following theorem, we show that superlinear functions are lower semicontinuous.

Theorem 2.23. *Assume that $f \in S_0(V)$, then f is a lower semicontinuous function from V into $\overline{\mathbb{R}}$ at any point.*

Proof. Let $x, y \in dom(g)$, by the definition of the supremum,

$$\forall \varepsilon > 0, \text{ there exists } j_0 \in I : g_{(x_{j_0}^*, x_{j_0}, c_{j_0}, \alpha_{j_0})}(x) > \sup_i g_{(x_i^*,c_i,x_i,\alpha_i)}(x) - \varepsilon \tag{2.16}$$

Multiplying (2.16) by -1 yields to,

$$-\sup_i g_{(x_i^*,c_i,x_i,\alpha_i)}(x)+\varepsilon > -g_{(x_{j_0}^*,x_{j_0},c_{j_0},\alpha_{j_0})}(x) \tag{2.17}$$

and we have

$$\sup_{i\in I} g_{(x_i^*,c_i,x_i,\alpha_i)}(y) \geq g_{(x_{j_0}^*,x_{j_0},c_{j_0},\alpha_{j_0})}(y), \quad \forall i \tag{2.18}$$

From (2.17) and (2.18) and considering g_{j_0} is a lower Lipschitz with the Lipschitz constant c_{j_0} we obtain that

$$\begin{aligned} f(y)-f(x) &= \sup_{i\in I} g_{(x_i^*,c_i,x_i,\alpha_i)}(y) - \sup_{i\in I} g_{(x_i^*,c_i,x_i,\alpha_i)}(x) \\ &\geq g_{(x_{j_0}^*,x_{j_0},c_{j_0},\alpha_{j_0})}(y) - g_{(x_{j_0}^*,x_{j_0},c_{j_0},\alpha_{j_0})}(x) - \varepsilon \\ &\geq -c_{j_0}\|y-x\| - \varepsilon \end{aligned}$$

Taking the liminf as $y \to x$ leads to

$$\liminf_{y\to x} f(y) \geq f(x) - \varepsilon.$$

Since ε is arbitrary; this implies that f is lower semicontinuous. $\square$

Next an illustrative example for Theorem 2.23 is presented.

Example 2.24. Consider the following functions :

$$\begin{aligned} g_1(x) =& \langle -\frac{1}{4}, x+2\rangle - \frac{5}{4}|x+2| + 6, \\ g_2(x) =& \langle -\frac{3}{2}, x+1\rangle - \frac{5}{2}|x+1| + 7, \\ g_3(x) =& \langle -5, x+\frac{1}{12}\rangle - 6|x+\frac{1}{12}| + \frac{91}{12}. \\ g_4(x) =& \langle -24, x+\frac{1}{10}\rangle - 25|x+\frac{1}{10}| + \frac{79}{10}. \\ g_5(x) =& \langle -249, x+\frac{1}{100}\rangle - 250|x+\frac{1}{100}| + \frac{799}{100}. \end{aligned}$$

The graph of these functions is provided in Figure 2.5. It is obvious that g_1, g_2, g_3, g_4 and g_5 belong to S(V). Let us construct the family of broken lines in a way that all the functions pass through the point $(-8, 0)$ and $(0, 3)$.

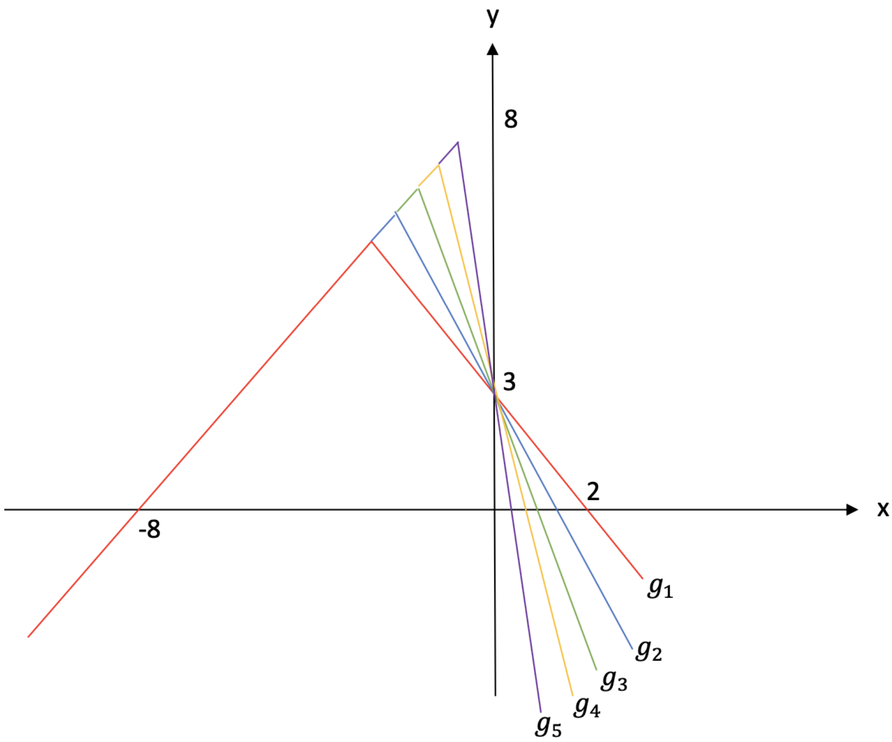

Figure 2.5. *The graphs of $g_i, i = 1, \ldots, 5$.*

We obtain the general form of the family of these superlinear functions in the following way. Assume that the function f takes the highest value at x_i where $f_i(x) = \langle x_i^*, x - x_i \rangle - c_i|x - x_i| + \alpha_i$. If $x < x_i$ then $|x - x_i| = -(x - x_i)$. Then we have $g_i(x) = \langle x_i^*, x - x_i \rangle + c_i(x - x_i) + \alpha_i = (x_i^* + c_i)x + (-x_i^* x_i - c_i x_i + \alpha_i)$. The equation of the left side of the highest point of this superlinear function is $x + 8$ this implies that if $x \leq x_i$ then $x_i^* + c_i = 1$ and $-x^* x_i - c_i x_i + \alpha_i = 8$ and if $x \geq x_0$ the line must pass through the point $(0, 3)$. By considering $x \geq x_i$ on the right side of the highest point results $g_i(x) = \langle x_i^*, x - x_i \rangle - c_i(x - x_i) + \alpha_i = (x_i^* - c_i)x + (-x_i^* x_i + c_i x_i + \alpha_i)$. Finally, we have $-x_i^* x_i + c_i x_i + \alpha_i = 3$. We can write the conditions in the following way, $x_i^* + c_i = 1, -x_i^* x_i - c_i x_i + \alpha_i = 8$, and $-x_i^* x_i + c_i x_i + \alpha_i = 3$. If we consider the equations $-x_i^* x_i - c_i x_i + \alpha_i = 8$ and $-x_i^* x_i + c_i x_i + \alpha_i = 3$, we obtain that $c_i x_i = -\frac{5}{2}$. Thus, the family, named I, consists of the $(x_i^*, x_i, c_i, \alpha_i)'$ s such that $x_i^* + c_i = 1, -x_i^* x_i - c_i x_i + \alpha_i = 8$ and $c_i x_i = -\frac{5}{2}$.

If we consider functions g_1 and g_5. For g_1, we have $x_1^* = -\frac{1}{4}$, $x_1 = -2$, $c_1 = \frac{5}{4}$ and $\alpha_1 = 6$. Clearly, $x_1^* + c_1 = -\frac{1}{4} + \frac{5}{4} = 1$, $-x_1^* x_1 - c_1 x_1 + \alpha_1 = -\frac{1}{4}(-2) - (\frac{5}{4})(-2) + 6 = 8$ and $c_1 x_1 = \frac{5}{4}(-2) = -\frac{5}{2}$. In the same way for g_5, we have $x_5^* = -249$, $x_5 = -\frac{1}{100}$,

$c_5 = 250$ and $\alpha_5 = \dfrac{799}{100}$. Clearly, $x_5^* + c_5 = -249 + 250 = 1$, $-x_5^* x_5 - c_5 x_5 + \alpha_5 = -(-249)(-\dfrac{1}{100}) - (250)(-\dfrac{1}{100}) + \dfrac{799}{100} = 8$ and $c_4 x_4 = 250(-\dfrac{1}{100}) = -\dfrac{5}{2}$.

If we consider the pointwise supremum of this family, we get the graph in Figure 2.6.

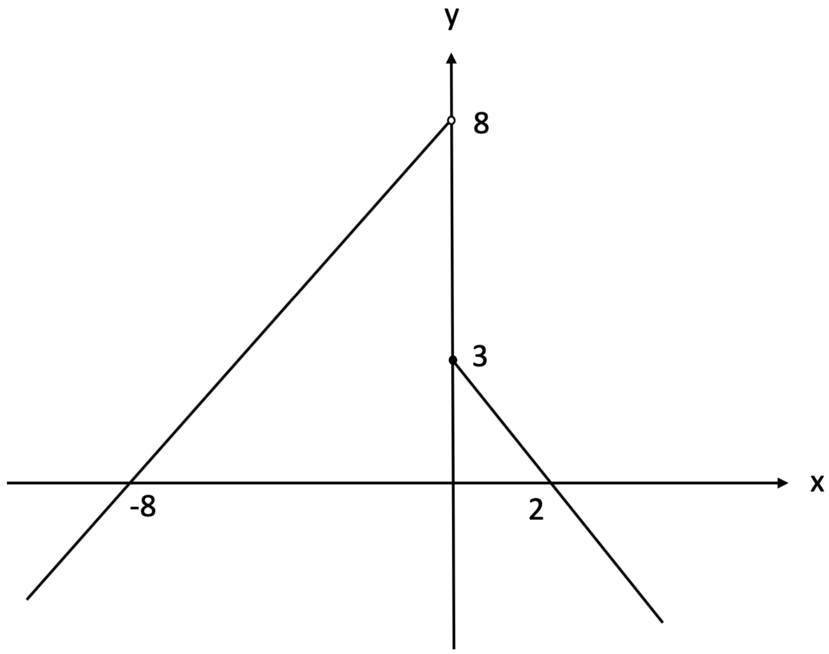

Figure 2.6. *The pointwise supremum of a family of superlinear functions* $g_i(x), i \in I$.

Clearly this function is lower semicontinuous. However, there exists a function that can be represented as a pointwise supremum of a family of superlinear functions that is not lower Lipschitz. Consider the function $f(x) = -\sqrt{|x|}$. The function f is lower semicontinuous and it can be represented as a pointwise supremum of the family of functions $g_{(x_i^*, c_i, x_i, \alpha_i)}(x) = g_{(0,c_i,0,0)}(x) = -c_i|x|$, which means that, $-\sqrt{|x|} = \sup_{i \in I}\{-c_i|x|\}$. However, at $x = 0$, the function $-\sqrt{|x|}$ is not lower Lipschitz. Assume the contrary, that f is lower Lipschitz at $x = 0$. Then there exists an $L > 0$ such that $f(x) - f(0) \geq -L|x - 0|$, equivalently $-\sqrt{|x|} \geq -L|x|$. Multiplying the last inequality yields that $\sqrt{|x|} \leq L|x|$. Then it yields that $|x| \leq L^2 x^2$. This inequality does not hold for every $L > 0$. Thus, $-\sqrt{|x|}$ is not a lower Lipschitz.

The following separation property definition given in [29].

Definition 2.25. [29, Definition 4.1] Assume that $\mathbb{C}$ and $\mathbb{K}$ be closed cones of a normed space $(\mathbb{Y}, \|\cdot\|)$ with norm bases $\mathbb{C}_{\mathbb{U}}$ and $\mathbb{K}_{\mathbb{U}}$, respectively. Suppose that $\mathbb{K}_{\mathbb{U}}^{\partial} = \mathbb{K}_{\mathbb{U}} \cap bd(\mathbb{K})$, and let $\tilde{\mathbb{C}}$ and $\tilde{\mathbb{K}}^{\partial}$ be the closures of the sets $co(\mathbb{C}_{\mathbb{U}})$ and $co(\mathbb{K}_{\mathbb{U}}^{\partial} \cup \{0_{\mathbb{Y}}\})$, respectively. The cones $\tilde{\mathbb{C}}$ and $\tilde{\mathbb{K}}^{\partial}$ are said to have the separation property with respect to the norm $\|\cdot\|$ if

$$\tilde{\mathbb{C}} \cap \tilde{\mathbb{K}}^{\partial} = \emptyset$$

The following Lemma is proved in [29] and we show that the lemma holds for the $\mathbb{R}^{n+1}$ case.

Lemma 2.26. *Suppose that $\mathbb{C}$ and $\mathbb{K}$ be nonempty cones in $\mathbb{Y}$. Assume that $\mathbb{C}^{a*} \neq \emptyset$. Then for each $((x^*, a^*), \alpha) \in \mathbb{C}^{a*}$ with $\alpha > 0$, the sublevel set $S((x^*, a^*), \alpha)$ defined by*

$$S((x^*, a^*), \alpha) = \{(x, a) \in \mathbb{Y} : \langle (x^*, a^*), (x, a) \rangle + \alpha \|(x, a)\| \leq 0\}$$

is a pointed closed cone that contains $-\mathbb{C}$.

Proof. The proof can be done by following the similar steps of the proof of Lemma 3.2 given in [29]. $\square$

The following Theorem is presented in [29] and we show that when the cones $\mathbb{C}$ and $\mathbb{K}$ belong to $\mathbb{R}^{n+1}$, the theorem remains true.

Theorem 2.27. *Assume that $\mathbb{C}$ and $\mathbb{K}$ be closed cones in a reflexive Banach space $(\mathbb{Y}, \|\cdot\|)$. Suppose that the cones $-\mathbb{C}$ and $\mathbb{K}$ satisfy the separation property given in definition 2.25,*

$$-\tilde{\mathbb{C}} \cap \tilde{\mathbb{K}}^{\partial} = \emptyset \tag{2.19}$$

Then, $\mathbb{C}^{a\#} \neq \emptyset$ and there exists a pair $((x^*, a^*), \alpha) \in \mathbb{C}^{a\#}$ such that the corresponding sublevel set $S((x^*, a^*), \alpha)$ of the strongly monotonically increasing sublevel function

$$g(x, a) = \langle (x^*, a^*), (x, a) \rangle + \alpha \|(x, a)\|$$

separates the cones $-\mathbb{C}$ and $bd(\mathbb{K})$ in the following sense

$$\langle (x^*, a^*), (\hat{x}, \hat{a}) \rangle + \alpha \|(\hat{x}, \hat{a})\| < 0 \leq \langle (x^*, a^*), (x, a) \rangle + \alpha \|(x, a)\| \tag{2.20}$$

for all $(\hat{x}, \hat{a}) \in -\mathbb{C} \setminus \{0_{\mathbb{Y}}\}$ *and* $(x, a) \in bd(\mathbb{K})$. *Then the cone* $-\mathbb{C}$ *is pointed. Conversely, if there exists a pair* $((x^*, a^*), \alpha) \in \mathbb{C}^{a\#}$ *such that the corresponding sublevel set* $S((x^*, a^*), \alpha)$ *of the strongly monotonically increasing sublinear function* $g(x, a) = \langle (x^*, a^*), (x, a) \rangle + \alpha \|(x, a)\|$ *separates the cones* $-\mathbb{C}$ *and* $bd(\mathbb{K})$ *in the sense of (2.20) and if either the cone* $\mathbb{C}$ *is closed and convex or* $(\mathbb{Y}, \|\cdot\|)$ *is a finite dimensional space, then the cones* $-\mathbb{C}$ *and* $\mathbb{K}$ *satisfy the separation property in (2.19).*

Proof. We omit the proof since it can be done similar to the proof of [29, Theorem 4.3]. □

Now, we present a separation relation for a arbitrary closed cone $\mathbb{K} \subset \mathbb{R}^{n+1}$. The steps of the proof is similar to [31, Lemma 3.1].

Lemma 2.28. *Assume that* $\mathbb{K}$ *is a closed cone in* $\mathbb{R}^n \times \mathbb{R}$ *and* $(\hat{y}, \hat{a}) \notin \mathbb{K}$. *Then there exists a vector* $(y, a) \in \mathbb{R}^{n+1} \setminus \{0_{\mathbb{R}^{n+1}}\}$ *and a real number* $\alpha \geq 0$ *such that separates the set*

$$\langle (y^*, a^*), (\hat{y}, \hat{a}) \rangle + \alpha \|(\hat{y}, \hat{a})\| < 0 \leq \langle (y^*, a^*), (y, a) \rangle + \alpha \|(y, a)\| \quad \forall (y, a) \in \mathbb{K}$$

Proof. In this proof, the idea is based on [33, Lemma 3.1]. Let $\|(\hat{y}, \hat{a})\| = 1$ and $\alpha = 1 - \frac{\varepsilon^2}{2}$. $\mathbb{K}$ is closed and $(\hat{y}, \hat{a}) \notin \mathbb{K}$ thus there exists an $\varepsilon \in (0, 1)$ such that

$$N_\varepsilon(\hat{y}, \hat{a}) = \{(y, a) \in \mathbb{R}^n \times \mathbb{R} : \|(y - \hat{y}, a - \hat{a})\| \leq \varepsilon\}$$

Assume that

$$\mathbb{C} = cone(N_\varepsilon(\hat{y}, \hat{a}))$$

and

$$\mathbb{C}_{\mathbb{U}} = \{(y, a) \in \mathbb{U} : \|(y - \hat{y}, a - \hat{a})\| \leq \varepsilon\}$$

$$
\begin{aligned}
(y,a)\in\mathbb{C}_{\mathbb{U}} \quad &\Longleftrightarrow \quad \|(y-\hat{y},a-\hat{a})\|^2\leq\varepsilon^2\\
&\Longleftrightarrow \quad \|y-\hat{y}\|^2+(a-\hat{a})^2\leq\varepsilon^2\\
&\Longleftrightarrow \quad \|y\|^2-2\langle y,\hat{y}\rangle+\|\hat{y}\|^2+a^2-2a\hat{a}+\hat{a}^2\leq\varepsilon^2\\
&\Longleftrightarrow \quad 2-2(\langle y,\hat{y}\rangle+a\hat{a})\leq\varepsilon^2\\
&\Longleftrightarrow \quad 1-\frac{\varepsilon^2}{2}\leq\langle(y,\hat{y}),(y,a)\rangle \quad \forall(y,a)\in\mathbb{U}\cap\mathbb{C}
\end{aligned}
$$

Thus we obtain that $\mathbb{C}_{\mathbb{U}}=\{(y,a)\in\mathbb{U}:\langle(y,\hat{y}),(y,a)\rangle\geq\alpha\}$.

The rest of the proof follows similarly. □

For a given set $S\subseteq\mathbb{R}^{n+1}$ and a given point $(\bar{y},\bar{a})\in A$, we present the separation theorem.

Theorem 2.29. *[31, Theorem 3.2] Assume that $S\subseteq\mathbb{R}^{n+1}$ be cone-shaped at $(\bar{y},\bar{a})\in S$. Then a closed pointed cone $\mathbb{C}\subset\mathbb{R}^{n+1}$ exists satisfying*

$$(S-\{(\bar{y},\bar{a})\}\cap\mathbb{C}\setminus\{0_{\mathbb{R}^{n+1}}\}=\emptyset$$

and there exists $((y^*,a^*),\alpha)\in(-\mathbb{C})^{a\#}$ satisfying

$$\langle(y^*,a^*),(y,a)-(\bar{y},\bar{a})\rangle+\alpha\|(y,a)-(\bar{y},\bar{a})\|\geq 0 \text{ for all } (y,a)\in S$$

Proof. The proof is similar to the proof of Theorem 3.2 in [31]. □

Theorem 2.30. *Let $f:\mathbb{R}^n\to\mathbb{R}$. Assume that the function $f\in S_0(V)$ is finite and continuous. Suppose also that $f(0)=0$. Then $epi(f)\cap N_\varepsilon(0,\overline{a})=\emptyset$.*

Proof. Since f is lower Lipschitz at $\overline{x}=0$, $f(x)-f(0)\geq -L\|x-0\|, \quad \forall x\in V$. Under the assumption $f(0)=0$, it implies that $f(x)\geq -L\|x\|, \quad \forall x\in V$.
Assume that $\bar{a}=-1$. Note that $epi(f)$ is a closed set. We want to show that there exists a positive real number ε such that $epi(f)\cap N_\varepsilon(0,\overline{a})=\emptyset$. $N_\varepsilon(0,\overline{a})$ is calculated as follows:

$$
\begin{aligned}
N_\varepsilon(0,\overline{a}) \quad &=\{(x,a)\in\mathbb{R}^n\times\mathbb{R}:\|(x,a)-(0,\overline{a})\|\leq\varepsilon\}\\
&=\{(x,a):\|x\|+|a-\overline{a}|\leq\varepsilon\}.
\end{aligned}
$$

$\forall (x_n, a_n) \in \text{epi}(f)$, there exists ε_n such that $f(x_n) \leq a_n$ and $(x_n, a_n) \notin N_\varepsilon(0, \overline{a})$. Assume that $(x_n, a_n) \in N_\varepsilon(0, \overline{a})$. It implies that $\|x_n\| + |a_n - \overline{a}| \leq \varepsilon_n$. We know that $(\alpha, x) : \alpha \leq -L\|x\|$ defines a cone. By the assumption $(x_n, a_n) \in N_\varepsilon(0, \overline{a})$ it implies that $-\varepsilon < x_n < \varepsilon$ and $\overline{a} - \varepsilon < a_n < \overline{a} + \varepsilon$. Thus, $f(x_n) \leq a_n < \overline{a} + \varepsilon$. We have $-L\|x_n\| \leq f(x_n)$. Since $\overline{a} + \varepsilon$ belongs to the defined cone it implies that $\overline{a} + \varepsilon \leq -L\|x_n\|$. Finally we have, $\overline{a} + \varepsilon \leq -L\|x_n\| \leq f(x_n) \leq a_n < \overline{a} + \varepsilon$ which is a contradiction. Thus $\text{epi}(f) \cap N_\varepsilon(0, \overline{a}) = \emptyset$. The proof is completed. □

We present the next theorem that investigates the finiteness of function f.

Theorem 2.31. *Assume that $f \in S_0(V)$. If f takes the value $-\infty$, then it cannot take any finite value.*

Proof. Suppose that there exists $\overline{x} \in V$ with $f(\overline{x})$ is finite. Then by the Teorem 2.29, the point $(\overline{x}, \overline{a})$ can be separated from $\text{epi}(f)$,

$$\langle x^*, \overline{x} \rangle - c\|\overline{x}\| + \alpha \overline{a} < 0 \leq \langle x^*, x \rangle - c\|x\| + \alpha a \quad \forall (x, a) \in \text{epi}(f) \tag{2.21}$$

Let $x = \overline{x}$ and $a = f(\overline{x})$ then it yields to $\alpha(f(\overline{x}) - \overline{a}) > 0$. Thus $\alpha > 0$.

Then (2.21) implies that

$$\langle x^*, \overline{x} - x \rangle - c\|\overline{x}\| + c\|x\| + \alpha \overline{a} \leq \alpha f(x) \quad \forall x \in dom(f).$$

Considering the inequality $-c\|x - \overline{x}\| \leq c\|x\| - c\|\overline{x}\|$, it follows that,

$$\langle x^*, \overline{x} - x \rangle - c\|x - \overline{x}\| + \alpha \overline{a} \leq \langle x^*, \overline{x} - x \rangle - c\|\overline{x}\| + c\|x\| + \alpha \overline{a} \leq \alpha f(x) \quad \forall x \in dom(f).$$

Thus,

$$\langle \frac{x^*}{\alpha}, \overline{x} - x \rangle - \frac{c}{\alpha}\|x - \overline{x}\| + \overline{a} \leq f(x) \quad \forall x \in dom(f).$$

which is a contradiction. □

Corollary 2.32. *[3, Corollary 3.1] Let f be bounded from below and proper function from X into $\mathbb{R} \cup \{+\infty\}$ and Lipschitz at the point $\overline{x}$. Then f is weakly subdifferentiable at $\overline{x}$.*

Remark 2. Theorem 4.4 and Corollary 2.32 show that any lower Lipschitz function is weakly subdifferentiable but it is proved for x^* component of the weak subgradient (x^*, c) when $x^* = \mathbf{0}$. The following Theorem shows that any lower Lipschitz function is weakly subdifferentiable with x^* component of the weak subgradient is different from $\mathbf{0}$.

Theorem 2.33. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a Lipschitz function. Then f is weakly subdifferentiable at $\bar{x} \in int(dom(f))$, that is $\partial^w f(\bar{x}) \neq \emptyset$ and there exists $(x^*, c) \in \partial^w f(\bar{x})$ with $x^* \neq 0$ and $c > 0$.*

Proof. The proof is based on the nonlinear cone separation theorem in [29]. Assume that $\overline{x} \in int(dom(f))$. Since $(\bar{x}, f(\bar{x}))$ belongs to the boundary of $\text{epi}(f) \subset \mathbb{R}^n \times \mathbb{R}$, we can separate it from $int(\text{epi}(f))$ by a closed pointed cone. By Theorem 2.29 there exists $((x^*, a^*), c) \in (-\mathbb{C})^{a\#}$ such that

$$\langle (x^*, a^*), (x, a) - (\bar{x}, \bar{a}) \rangle + c\|(x, a) - (\bar{x}, \bar{a})\| \geq 0 \quad \forall (x, a) \in \text{epi}(f),$$
$$\langle (x^*, a^*), (x - \bar{x}, a - \bar{a}) \rangle + c\|(x - \bar{x}, a - \bar{a})\| \geq 0 \quad \forall (x, a) \in \text{epi}(f),$$
$$\langle x^*, x - \bar{x} \rangle + a^*(a - f(\bar{x})) + c\|x - \bar{x}\| + c|a - f(\bar{x})| \geq 0 \quad \forall (x, a) \in \text{epi}(f),$$
$$\langle x^*, x - \bar{x} \rangle + a^*(f(x) - f(\bar{x})) + c\|x - \bar{x}\| + c|f(x) - f(\bar{x})| \geq 0 \quad \forall x \in dom(f).$$

By the assumption that f is Lipschitz, it implies that,

$$c|f(x) - f(\bar{x})| \leq cL\|x - \bar{x}\|.$$

Then we have,

$$\langle x^*, x - \bar{x} \rangle + a^*(f(x) - f(\bar{x})) + c\|x - \bar{x}\| + cL\|x - \bar{x}\|,$$
$$\geq \langle x^*, x - \bar{x} \rangle + a^*(f(x) - f(\bar{x})) + c\|x - \bar{x}\| + c|f(x) - f(\bar{x})| \geq 0 \quad \forall x \in dom(f).$$

Hence we obtain,

$$\langle x^*, x - \bar{x} \rangle + a^*(f(x) - f(\bar{x})) + c\|x - \bar{x}\| + cL\|x - \bar{x}\| \geq 0 \quad \forall x \in dom(f).$$

It follows that,

$$\langle x^*, x - \bar{x} \rangle + (c + cL)\|x - \bar{x}\| \geq -a^*(f(x) - f(\bar{x})) \quad \forall x \in dom(f).$$

Then we have,

$$\langle -\frac{x^*}{a^*}, x - \bar{x} \rangle - (\frac{c + cL}{a^*})\|x - \bar{x}\| \leq f(x) - f(\bar{x}) \quad \forall x \in dom(f).$$

Hence f is weakly subdifferentiable at $\bar{x}$; that is, $(-\frac{x^*}{a^*}, \frac{c + cL}{a^*}) \in \partial^w f(\bar{x})$. The proof is completed. $\square$

3. THE SUM RULE FOR THE WEAK SUBDIFFERENTIAL

In this section, we present the sum rule in the nonconvex case for two classes of functions. The class of functions are: The Clarke directionally differentiable functions and tangentially convex functions. The following theorem explains some classes of weakly subdifferentiable functions.

Theorem 3.1. [33, Lemma 2.7] *Let $f : S \to (-\infty, +\infty]$ be a given function. If f is a positively homogeneous function bounded from below on some neighborhood of $0_{\mathbb{R}^n}$, then f is weakly subdifferentiable at $0_{\mathbb{R}^n}$.*

3.1. The "max relation" for the Clarke Directionally Differentiable Functions

Kasimbeyli and Mammadov [34] established conditions under which a (nonconvex) directionally differentiable function is weakly subdifferentiable and its directional derivative in any direction can be represented as a support function of its weak subdifferential. We will call this property a "sup relation". In this section, we prove that this property is satisfied as a "max relation" for some class of Clarke directionally differentiable functions. We begin by recalling the definitions of the Clarke's directional derivative and the Clarke subdifferential [12] (see also [13]).

Definition 3.2. Let $\mathbb{X}$ be a Banach space, $f : \mathbb{X} \to \mathbb{R}$ be a locally Lipschitz function and $\overline{x} \in \mathbb{X}$ and $h \in \mathbb{X}$ be given elements. Then, the Clarke directional derivative $f^{\circ}(\overline{x}; h)$ of f at $\overline{x}$ in the direction h, is defined by

$$f^{\circ}(\overline{x}; h) = \limsup_{t \downarrow 0, y \to \overline{x}} \frac{f(y + th) - f(y)}{t}. \tag{3.22}$$

In this case, we will say that f is Clarke directionally differentiable function at $\overline{x}$ in the direction h.

It was proved that the Clarke directional derivative $f^{\circ}(\overline{x}; \cdot)$ is convex (see [47]), upper semicontinuous, positively homogeneous, sublinear and Lipschitz function (see [13, Proposition 2.1.1, p. 25]).

The following example is introduced in [13] but we show the detailed calculation of the Clarke directional derivative and the directional derivate.

Example 3.3. Let's consider the directional derivative of $f(x) = -|x|$ at $x = 0$;

$$f(x;h) = \lim_{t\downarrow 0} \frac{f(x+td) - f(x)}{t}$$
$$f(0;h) = \lim_{t\downarrow 0} \frac{f(0+td) - f(0)}{t} = \lim_{t\downarrow 0} \frac{-|0+td| + |0|}{t}$$
$$f(0;h) = -|d|$$

Now, we calculate the Clarke directional derivative of $f(x) = -|x|$ at $x = 0$;

$$f^o(x;d) = \limsup_{y\to x, t\downarrow 0} \frac{f(y+td) - f(y)}{t}$$
$$f^o(0;d) = \limsup_{y\to 0, t\downarrow 0} \frac{-|y+td| + |y|}{t}$$

Since we have two directions in $\mathbb{R}$, we have two cases. Now consider the case $d = 1$:
If $y > 0$ then,

$$\frac{f(y+td) - f(y)}{t} = \frac{-|y+t| + |y|}{t} = \frac{-y-t+y}{t} = -1$$

If $y < 0$ and $|y| \leq t$, then consider the sequences $y = \frac{-1}{n^2}$ and $t = \frac{1}{n}$. Thus we have:

$$\begin{aligned}\frac{f(y+td) - f(y)}{t} &= \frac{-|-1/n^2 + 1/n| + |-1/n^2|}{1/n} \\ &= \frac{1/n^2 - 1/n + 1/n^2}{1/n} \\ &= \frac{(2-n)/n^2}{1/n} = -1\end{aligned}$$

If $y < 0$ and $|y| > t$, then consider the sequences $y = \frac{-1}{n}$ and $t = \frac{1}{n^2}$. It follows that:

$$\begin{aligned}\frac{f(y+td) - f(y)}{t} &= \frac{-|-1/n + 1/n^2| + |-1/n|}{1/n^2} = \frac{-1/n + 1/n^2 + 1/n}{1/n^2} \\ &= \frac{1/n^2}{1/n^2} = 1\end{aligned}$$

Thus $f^o(0;1) = 1$. Now consider the second case that $d = -1$.

If $y < 0$ then we have

$$\frac{f(y+td) - f(y)}{t} = \frac{-|y-t| + |y|}{t} = \frac{-(t-y) - y}{t} = -1$$

If $y > 0$ and $y \leq t$, then if we consider the sequences $y = \frac{1}{n^2}$ and $t = \frac{1}{n}$. It yields to:

$$\begin{aligned}\frac{f(y+td)-f(y)}{t} &= \frac{-|1/n^2-1/n|+|1/n^2|}{1/n}\\ &= \frac{-(-1/n^2+1/n)+1/n^2}{1/n}\\ &= \frac{(2-n)/n^2}{1/n} = -1\end{aligned}$$

If $y > 0$ and $y > t$, then we may choose $y = \frac{1}{n}$ and $t = \frac{1}{n^2}$. It follows that

$$\begin{aligned}\frac{f(y+td)-f(y)}{t} &= \frac{-|1/n-1/n^2|+|1/n|}{1/n^2}\\ &= \frac{-1/n+1/n^2+1/n}{1/n^2}\\ &= \frac{1/n^2}{1/n^2} = 1\end{aligned}$$

Thus $f^o(0;-1) = 1$.

From $f^o(0;1) = 1$ and $f^o(0;-1) = 1$, we conclude that $f^o(x;d) = |d|$.

The following two theorems were also proved in [13].

Theorem 3.4. [13, Proposition 2.1.2, p. 27] *Let f be Lipschitz of rank M near $\overline{x}$. Then*

(a) $\partial^\circ f(\overline{x})$ is a nonempty, convex, weak$-$compact subset of $\mathbb{X}^*$ and $\|x^*\|_* \leq M$ for every $x^* \in \partial^\circ f(\overline{x})$.*

(b) For every $h \in \mathbb{X}$, one has

$$f^\circ(\overline{x};h) = \max\{\langle x^*, h\rangle : x^* \in \partial^\circ f(\overline{x})\}. \tag{3.23}$$

Theorem 3.5. [13, Proposition 2.2.7, p. 36] *When f is convex on an open convex set S and Lipschitz near $\overline{x}$, then $\partial^\circ f(\overline{x})$ coincides with the subdifferential $\partial f(\overline{x})$ of f at $\overline{x}$ in the sense of convex analysis, and $f^\circ(\overline{x};h)$ coincides with the directional derivative $f'(\overline{x};h)$ for each h.*

A simple example demonstrates that the convexity is an essential condition in Theorem 3.5. For example if $f(x) = -|x|$, then $f^\circ(0;h) = |h|$ and $\partial^\circ f(0) = [-1,1]$, while $f'(0;h) = -|h|$ and $\partial f(0) = \emptyset$.

Now we formulate the following assumption, which will play an important role in this thesis.

Assumption 3.6. *Let $S \subseteq \mathbb{R}^n$ be star-shaped at $\overline{x} \in S$, and $f : S \to \mathbb{R}$ be a Clarke directionally differentiable function at $\overline{x}$ in every direction $x - \overline{x}$ with arbitrary $x \in S$. Asssume that the following condition holds:*

$$f(x) - f(\overline{x}) \geq f^\circ(\overline{x}; x - \overline{x}) \text{ for all } x \in S. \tag{3.24}$$

The following theorem establishes conditions, under which the given function f and its Clarke derivative are both weakly subdifferentiable and their weak subdifferential sets are equal. The idea of the proof of this theorem is taken from [34, Theorem 1].

Theorem 3.7. *Let $S \subseteq \mathbb{R}^n$ be star-shaped at $\overline{x} \in S$ and let Assumption 3.6 be satisfied for function $f : S \to \mathbb{R}$. Then $f^\circ(\overline{x}; \cdot)$ is weakly subdifferentiable at $0_{\mathbb{R}^n}$, f is weakly subdifferentiable at $\overline{x}$ and*

$$\partial^w_{S-\overline{x}} f^\circ(\overline{x}; 0) = \partial^w_S f(\overline{x}). \tag{3.25}$$

Proof. Suppose that the Clarke directional derivative $f^\circ(\overline{x}; \cdot)$ exists. Since $f^\circ(\overline{x}; \cdot)$ is a convex and positively homogeneous function, it is weakly subdifferentiable at $0_{\mathbb{R}^n}$ (see Theorem 3.1, or Remark 1). If $(x^*, c) \in \partial^w_S f^\circ(\overline{x}; 0)$, by definition of the weak subgradient we have:

$$f^\circ(\overline{x}; x - \overline{x}) \geq \langle x^*, x - \overline{x} \rangle - c\|x - \overline{x}\| \text{ for all } x \in S.$$

By Assumption 3.6, this implies that $(x^*, c) \in \partial^w_S f(\overline{x})$.

Now let $(x^*, c) \in \partial^w_S f(\overline{x})$. By definition of the Clarke derivative, we have:

$$f^\circ(\overline{x}; x - \overline{x}) = \limsup_{t\downarrow 0, y \to \overline{x}} \frac{f(y + t(x - \overline{x})) - f(y)}{t} \geq \limsup_{t\downarrow 0, y \to \overline{x}} \frac{t(\langle x^*, x - \overline{x} \rangle - c\|x - \overline{x}\|)}{t}$$

for all $x \in S$, which implies that $(x^*, c) \in \partial^w_{S-\overline{x}} f^\circ(\overline{x}; 0)$. □

Corollary 3.8. *Let f be Lipschitz of rank M near $\overline{x}$ and let the condition* (3.24) *is satisfied. Then $f^\circ(\overline{x}; \cdot)$ is subdifferentiable at $0_{\mathbb{R}^n}$, f is subdifferentiable at $\overline{x}$ and*

$$\partial f^\circ(\overline{x}; 0) = \partial f(\overline{x}). \tag{3.26}$$

Proof. The subdifferentiability of the Clarke directional derivative at $\overline{x}$ follows from Theorem 3.4. The remaining part of the proof can be obtained from the proof of

Theorem 3.7 by changing the weak subdifferential to the classical subdifferential of the convex analysis. □

The following theorem establishes conditions under which the Clarke directional derivative of a given function, can be represented as a support function of the weak subdifferential of this function.

Theorem 3.9. *Let $S \subseteq \mathbb{R}^n$ be star-shaped at $\overline{x} \in S$ and let Assumption 3.6 be satisfied for function $f : S \to \mathbb{R}$. Assume additionally that f is Lipschitz of rank M near $\overline{x}$. Then f is weakly subdifferentiable at $\overline{x}$, and there exists a positive number L such that*

$$f^\circ(\overline{x}; h) = \max\{\langle x^*, h\rangle - c\,\|h\| : \ (x^*, c) \in \partial_S^w f(\overline{x}), \|x^*\| + c \leq L\}, \ \forall h \in S - \overline{x}. \quad (3.27)$$

Proof. Let $\overline{x} \in S$. Then, under the hypotheses of the theorem, f is Clarke subdifferentiable at $\overline{x}$. It follows from Definition 3.2 that the set $\partial^\circ f(\overline{x})$ is actually a subdifferential set $\partial f^\circ(\overline{x}; 0)$ of the Clarke directional derivative at $\overline{x}$. By Theorem 3.4, the subdifferential $\partial f^\circ(\overline{x}; 0)$ is convex and compact. On the other hand, as it was pointed out in Remark 1, if $x^* \in \partial f^\circ(\overline{x}; 0)$, then for arbitrary constant $c_1 > 0$, we have $(x^*, c) \in \partial_S^w f^\circ(\overline{x}; 0)$ for all $c \leq c_1$. Then, the relation (3.23) of Theorem 3.4 can be written in the form:

$$f^\circ(\overline{x}; h) = \max\{\langle x^*, h\rangle : x^* \in \partial f^\circ(\overline{x})\}$$

for all $h \in S - \overline{x}$, or in view of Theorem 3.7, in the form:

$$f^\circ(\overline{x}; h) = \max\{\langle x^*, h\rangle - c\|h\| : (x^*, c) \in \partial_S^w f(\overline{x}), c \leq c_1\} \quad (3.28)$$

for all $h \in S - \overline{x}$. By Theorem 3.4, $\|x^*\| \leq M$ for all $x^* \in \partial f^\circ(\overline{x})$. Now, by taking this into account in (3.28) and by setting $M + c_1 = L$ we obtain the relation (3.27). □

Remark 3. By denoting

$$\partial_{S_L}^w f(\overline{x}) = \{(x^*, c) \in \partial_S^w f(\overline{x}) : \|x^*\| + c \leq L\}, \quad (3.29)$$

the relation (3.27) can be written in the form

$$f^\circ(\overline{x}; h) = \max\{\langle x^*, h\rangle - c\,\|h\| : \ (x^*, c) \in \partial_{S_L}^w f(\overline{x})\}, \ \forall h \in S - \overline{x}, \quad (3.30)$$

where the maximum is taken over the convex (see Lemma 2.12) and compact set $\partial^w_{S_L} f(\overline{x})$.

3.2. The "max relation" for the Tangentially Convex Functions

In this section, we show that the "max relation" given in (3.30) for Clarke directionally differentiable functions, remains true for a special class of directionally differentiable functions called tangentially convex functions. These functions were first presented by Pshenichnyi [41,42] and they were named as "tangentially convex" by Lemarechal [35]. Martinez-Legaz obtained KKT type theorem for this class of functions [36].

Definition 3.10. A function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is called tangentially convex at $x \in f^{-1}(\mathbb{R})$ if the directional derivative $f'(x; \cdot)$ exists, is finite and convex. The tangential subdifferential of $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ at $x \in f^{-1}(\mathbb{R})$ is the set

$$\partial_T f(x) := \{x^* \in \mathbb{R}^n : \langle x^*, h\rangle \leq f'(x; h), \forall h \in \mathbb{R}^n\}.$$

It obviously follows from this definition that, if f is tangentially convex at x, then $\partial_T f(x) \neq \emptyset$ and $f'(x; \cdot)$ is the support functional of $\partial_T f(x)$, that is, for every $d \in \mathbb{R}^n$ one has

$$f'(x; h) = \max_{x^* \in \partial_T f(x)} \langle x^*, h\rangle. \tag{3.31}$$

Assumption 3.11. *Suppose that $S \subseteq \mathbb{R}^n$ be star-shaped at $\overline{x} \in S$, and let $f : S \to \mathbb{R}$ be a given function. Assume that f has directional derivative at $\overline{x}$ in every direction $h \in \mathbb{R}^n$, the directional derivative $f'(\overline{x}; \cdot)$ of f at $\overline{x}$ is lower semi-continuous on $\mathbb{K} = \textsf{cone}(S - \{\overline{x}\})$ and the following conditions is satisfied:*

$$f(x) - f(\overline{x}) \geq f'(\overline{x}; x - \overline{x}) \quad \textit{for all } x \in S \setminus \{\overline{x}\}, \tag{3.32}$$

and

$$\inf\{f'(\overline{x}; h) : \ h \in \mathbb{K} \cap \mathbb{U}\} > -\infty. \tag{3.33}$$

The following theorem gives the "sup relation", which establishes conditions under which the directional derivative of a given (nonconvex) function, can be represented as a support function of the weak subdifferential of this function.

Theorem 3.12. [34, Theorem 1] *Let Assumption 3.11 hold. Then f is weakly subdifferentiable at $\overline{x}$ on S; that is, $\partial_S^w f(\overline{x}) \neq \emptyset$ and*

$$f'(\overline{x}; h) = \sup\{\langle x^*, h\rangle - c\,\|h\| : \ (x^*, c) \in \partial_S^w f(\overline{x})\}, \ \forall h \in \mathbb{K}. \tag{3.34}$$

The following theorem which was proved in [15], shows that the "sup relation" given above by Theorem 3.12, can be formulated as a "max relation" for tangentially convex functions.

Theorem 3.13. [15, Theorem 2.8] *Let Assumption 3.11 hold. Assume in addition that the function $f : S \to \mathbb{R}$ is tangentially convex at $\overline{x} \in S$, that is the directional derivative $f'(\overline{x}; \cdot)$ of f at $\overline{x}$, is a convex function on $\mathbb{R}^n$. Then f is weakly subdifferentiable at $\overline{x}$, $f'(\overline{x}; \cdot)$ is weakly subdifferentiable at $0_{\mathbb{R}^n}$, $\partial_S^w f(\overline{x}) = \partial_{S-\overline{x}}^w f'(\overline{x}; \mathbf{0})$ and there exists a positive number L such that*

$$f'(\overline{x}; h) = \max\{\langle x^*, h\rangle - c\,\|h\| : \ (x^*, c) \in \partial_S^w f(\overline{x}), \|x^*\| + c \leq L\}, \ \forall h \in \mathbb{K}. \tag{3.35}$$

Proof. The proof is given in [15, Theorem 2.8]. Note also that, all the assertions can be obtained by similar ways used in proofs of Theorems 3.7 and 3.9. □

Remark 4. It is clear that if the directional derivative of a function is convex it does not have to imply the convexity of the function itself. Thus, Theorem 3.13 is extremely important since it presents a representation of the directional derivative of a nonconvex function in the form of "max relation" by using a compact subset of the weak subdifferential. By using the notation $\partial_{S_L}^w f(\overline{x})$ given in Remark 3 (see (3.29)), we can use the similar to (3.30) representation for directionally differentiable functions:

$$f'(\overline{x}; h) = \max\{\langle x^*, h\rangle - c\,\|h\| : \ (x^*, c) \in \partial_{S_L}^w f(\overline{x})\}, \ \forall h \in \mathbb{K}. \tag{3.36}$$

Next, we present examples to illustrate Assumptions 3.6 and 3.11.

Example 3.14. Define $f_1 : \mathbb{R} \to \mathbb{R}$ by

$$f_1(x) = \begin{cases} x^2 \sin^2 \frac{1}{x} & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

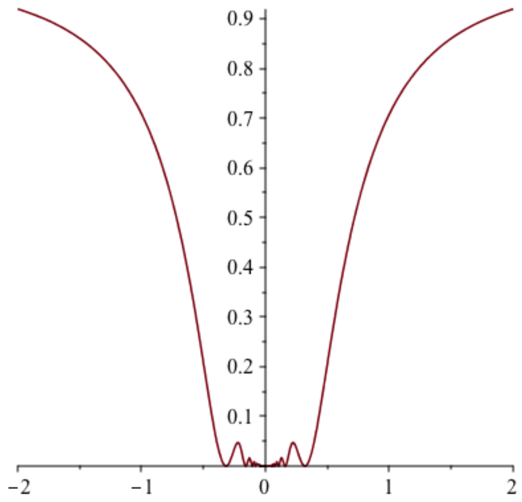

Figure 3.1. *The graph of* $f_1(x)$

Then for $\overline{x} = 0$ we have:

$$f_1'(0;h) = \lim_{t\downarrow 0} \frac{f_1(0+th) - f_1(0)}{t} = \lim_{t\downarrow 0} \frac{t^2h^2\sin^2(\frac{1}{th})}{t} = 0.$$

Thus $f_1'(0;h) = 0$ for all $h \in \mathbb{R}$, but the derivative mapping ∇f is discontinuous at $\overline{x} = 0$. We have $f_1(x) - f_1(0) \geq f_1'(0; x-0)$ for all $x \in \mathbb{R}$, and f_1 is a nonconvex function, see Figure 3.1.

This function also demonstrates the case when the given function is directionally differentiable, but is not Clarke directionally differentiable. For the Clarke directional derivative of f_1 at $\overline{x} = 0$ we have:

$$\begin{aligned} f_1^\circ(0;h) &= \limsup_{t\downarrow 0, u\to 0} \frac{f_1(u+th) - f_1(u)}{t} = \limsup_{t\downarrow 0, u\to 0} \frac{(u+th)^2\sin^2\frac{1}{u+th} - u^2\sin^2\frac{1}{u}}{t} \\ &= \limsup_{t\downarrow 0, u\to 0} \frac{u^2(\sin^2\frac{1}{u+th} - \sin^2\frac{1}{u})}{t}. \end{aligned} \tag{3.37}$$

It is easy to show that the above lim sup approaches infinity. Indeed, choose sequences $u_n = \frac{1}{\alpha+2\pi n}$ and $t_n = \frac{1}{\beta+2\pi n^4}$ with $0 < \alpha, \beta < \pi/2$. Then clearly, $u_n \to 0$ and $t_n \to 0$ as $n \to \infty$. Then,

$$\begin{aligned} &\limsup_{n\to\infty} \frac{(1/(\alpha+2\pi n))^2(\sin^2\frac{1}{1/(\alpha+2\pi n)+h/(\beta+2\pi n^4)} - \sin^2\frac{1}{1/(\alpha+2\pi n)})}{1/(\beta+2\pi n^4)} \\ &\qquad = \limsup_{n\to\infty} \frac{\sin^2\frac{(\alpha+2\pi n)(\beta+2\pi n^4)}{h(\alpha+2\pi n)+\beta+2\pi n^4} - \sin^2(\alpha+2\pi n)}{(\alpha+2\pi n)^2/(\beta+2\pi n^4)}. \end{aligned} \tag{3.38}$$

Now since the sequence $\frac{(\alpha+2\pi n)(\beta+2\pi n^4)}{h(\alpha+2\pi n)+\beta+2\pi n^4}$ approaches to infinity at a rate n, we can choose a subsequence $\gamma + 2\pi n$ with $\alpha < \gamma < \pi/2$. Then for this subsequence, the numerator of the fraction on the right side of (3.38) will tend to the positive number $\sin^2\gamma - \sin^2\alpha$, and the denominator will approach to zero. This shows that the lim sup in (3.37) approaches to infinity, and hence f_1 is not Clarke directionally differentiable at $\overline{x} = 0$. □

The next two examples demonstrate different cases related to Assumptions 3.11 and 3.6, as well as to Theorems 3.9 and 3.13.

Example 3.15. Define $f_2 : \mathbb{R} \to \mathbb{R}$ by

$$f_2(x) = \begin{cases} (x+1)^2 & \text{if } x \leq 1, \\ (x-2)^2 + 3 & \text{if } x > 1. \end{cases}$$

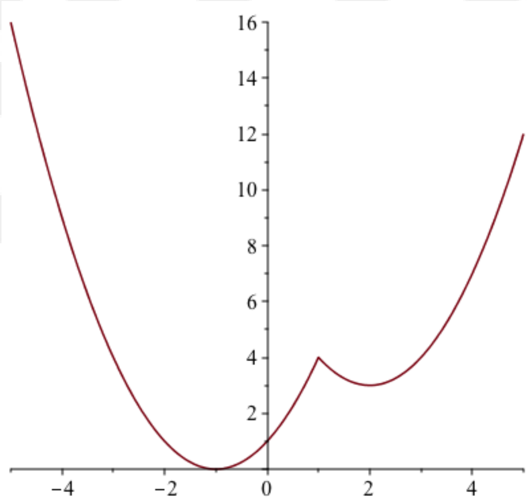

Figure 3.2. *The graph of $f_2(x)$*

Then for $\overline{x} = 1$ we have:

$$f_2'(1; h) = \begin{cases} 4h & \text{if } h \leq 0, \\ -2h & \text{if } h > 0. \end{cases}$$

Clearly, neither f_2 nor its directional derivative $f_2'(1; \cdot)$ are not convex. Nevertheless it satisfies the condition $f_2(x) - f_2(1) \geq f_2'(1; x)$ for all $x \in \mathbb{R}$ and moreover the "max relation" (3.35) of Theorem 3.13 remains true, see Figure 3.2. It is easy to show that

$$\partial^w f_2'(1; 0) = \partial^w f_2(1) = \{(v, c) \in \mathbb{R} \times \mathbb{R}_+ : -c + 4 \leq v \leq c - 2\}$$

and

$$f_2'(1;h) = \sup\{vh - c|h| : (v,c) \in \partial^w f_2(1)\} = \sup\{vh - c|h| : -c+4 \leq v \leq c-2\}.$$

Then, obviously we can write the last equality as follows:

$$f_2'(1;h) = \max\{vh - c|h| : -c+4 \leq v \leq c-2, c \leq 4\},$$

which implies that, the "max relation" of Theorem 3.13 is satisfied. Note that function f_2 is locally Lipschitz and Clarke directionally differentiable at x_1 and

$$f_2^\circ(1;h) = \begin{cases} -2h & \text{if } h \leq 0, \\ 4h & \text{if } h > 0 \end{cases}$$

which shows that neither the condition (3.24) of Assumption 3.6, nor the relation (3.26) and Theorem 3.9 are satisfied at $\overline{x} = 1$. However all these conditions and the Theorem 3.9 are satisfied at points $(-\infty, -1/2] \cup [5/2, +\infty)$. $\square$

Now consider the following example.

Example 3.16. Define $f_3 : \mathbb{R} \to \mathbb{R}$ by

$$f_3(x) = \begin{cases} |x+2| & \text{if } x \leq 0, \\ 4x+2 & \text{if } 0 < x \leq 1, \\ 2|x-2|+4 & \text{if } x > 1. \end{cases}$$

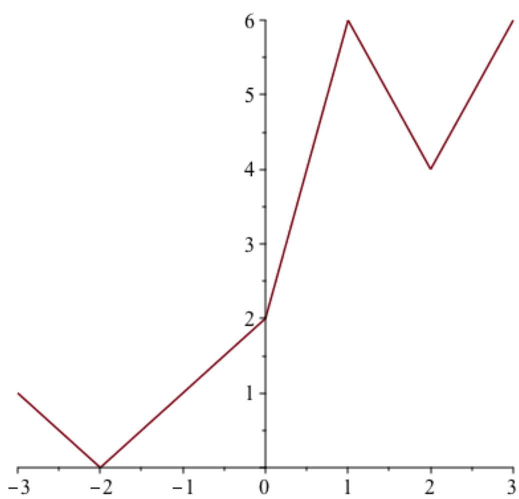

Figure 3.3. *The graph of $f_3(x)$*

Then $f_3'(x;h) = f_3^\circ(x;h)$ for every $x \neq 1$ and for all $h \in \mathbb{R}$. We have:

$$f_3(x) - f_3(\overline{x}) \geq f_3'(\overline{x}; x - \overline{x})$$

for every $\overline{x} < 0, \overline{x} = 1, \overline{x} > 2$ and for all $x \in \mathbb{R}$. Thus, all the assumptions of Theorem 3.13 are satisfied and hence the "max relation" (3.35) remains true for every $\overline{x} < 0, \overline{x} > 2$, see Figure 3.3. We also have:

$$f_3(x) - f_3(\overline{x}) \geq f_3^\circ(\overline{x}; x - \overline{x})$$

for every $\overline{x} < 0, \overline{x} > 2$ and for all $x \in \mathbb{R}$, which shows that the Assumption 3.6, the relation (3.26) and Theorem 3.9 are all satisfied at these points.

The directional derivative $f_3'(\overline{x};\cdot)$ and the Clarke directional derivative $f_3^\circ(\overline{x};\cdot)$ are both convex at the points $\overline{x} = 0$ and $\overline{x} = 2$, but the conditions $f_3(x) - f_3(\overline{x}) \geq f_3'(\overline{x}; x - \overline{x})$ and $f_3(x) - f_3(\overline{x}) \geq f_3^\circ(\overline{x}; x - \overline{x})$ are not satisfied at these points.

For $\overline{x} = 1$ we have:

$$f_3'(1;h) = \begin{cases} 4h & \text{if } h \leq 0, \\ -2h & \text{if } h > 0. \end{cases}$$

Clearly, neither f_3 nor its directional derivative $f_3'(1;\cdot)$ are convex. Nevertheless, it satisfies the condition $f_3(x) - f_3(1) \geq f_3'(1; x-1)$ for all $x \in \mathbb{R}$ and moreover the "max relation" (3.35) of Theorem 3.13 remains true. It is easy to show that (see Example 3.15)

$$\partial^w f_3'(1;0) = \partial^w f_3(1) = \{(v,c) \in \mathbb{R} \times \mathbb{R}_+ : -c + 4 \leq v \leq c - 2\}$$

and

$$f_3'(1;h) = \max\{vh - c|h| : (v,c) \in \partial^w f_3(1), c \leq 4\}.$$

□

Another illustration of Theorem 3.13 is demonstrated in the following example.

Example 3.17. Let $X = \mathbb{R}$ and let

$$f_4(x) = \begin{cases} x^2 - 6 & \text{if } x < -2, \\ x & \text{if } -2 \leq x \leq 2, \\ (x-4)^3 + 10 & \text{if } x > 2. \end{cases}$$

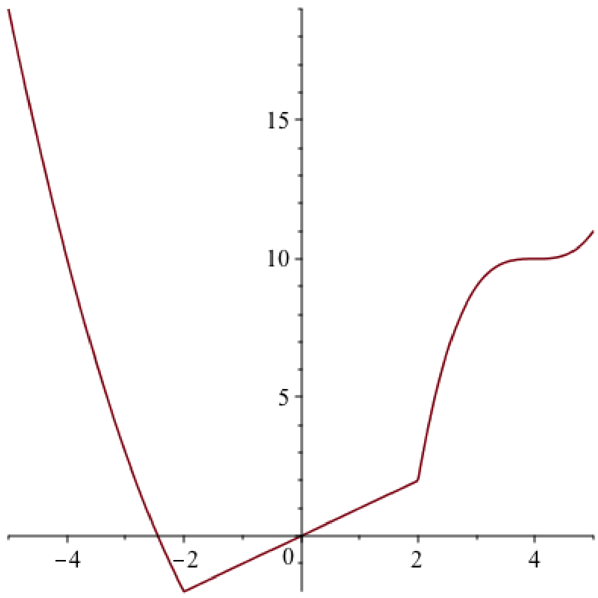

Figure 3.4. *The graph of* $f_4(x)$

It is evident that this function achieves its global minimum at $x = -2$. Note also that the directional derivative of this function at $x = -2$, is as follows:

$$f_4'(-2; h) = \begin{cases} -4h & , h < 0 \\ h & , h \geq 0. \end{cases}$$

Then, clearly

$$\partial^w f_4(-2) = \partial^w f_4'(-2; 0) = \{(x^*, c) \in \mathbb{R} \times \mathbb{R}_+ : -c - 4 \leq x^* \leq c + 1\}$$

and it can easily be shown that

$$f_4'(-2; h) = max\{x^* h - c|h| : (x^*, c) \in \partial^w f(-2), 0 \leq c \leq L\}$$

for every $L > 0$, see Figure 3.4. $\square$

In the next subsection, we will prove the sum rule for the weak subdifferential by using "max relations" established in this and previous subsections. Since these relations require different conditions for different classes of functions, we will prove these theorems for Clarke directionally differentiable and directionally differentiable functions separately.

3.3. Sum Rule

In this subsection, we prove the sum rule for the weak subdifferential. We will show that this property is satisfied for the class of functions satisfying the conditions identified in Assumptions 3.6 and 3.11. The main theorems of this section will be proved by using the indicator and support functions of the weak subdifferential set. Therefore we first give formulations and properties of these functions. Denote the indicator function of the weak subdifferential $\partial^w_{S_L} f(\overline{x})$ defined in (3.29), by $\delta_{\partial^w_{S_L} f(\overline{x})}$:

$$\delta_{\partial^w_{S_L} f(\overline{x})}(x^*, c) = \begin{cases} 0 \text{ if } (x^*, c) \in \partial^w_{S_L} f(\overline{x}), \\ +\infty, \text{ otherwise.} \end{cases}$$

Theorem 3.18. *Let $\sigma_{\partial^w_{S_L} f(\overline{x})} : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R} \cup \{\pm\infty\}$ be the support function of the weak subdifferential $\partial^w_{S_L} f(\overline{x})$, then*

$$\sigma_{\partial^w_{S_L} f(\overline{x})}(h, -\|h\|) = \delta^*_{\partial^w_{S_L} f(\overline{x})}(h, -\|h\|) \text{ for all } h \in S - \overline{x}, \tag{3.39}$$

where $\delta^*_{\partial^w_{S_L} f(\overline{x})}$ denotes the conjugate function of the indicator function $\delta_{\partial^w_{S_L} f(\overline{x})}$.

Proof. By definition,

$$\sigma_{\partial^w_{S_L} f(\overline{x})}(h, -\|h\|) = \begin{cases} \sup\{\langle (x^*, c), (h, -\|h\|)\rangle : (x^*, c) \in \partial^w_{S_L} f(\overline{x})\} \text{ if } h \in S - \overline{x}, \\ -\infty, \text{ otherwise .} \end{cases} \tag{3.40}$$

Note that the value of the support function $\sigma_{\partial^w_{S_L} f(\overline{x})}(h, -\|h\|)$ is equal to $-\infty$ if $h \notin S - \overline{x}$, because in this case the weak subdifferential is not defined and hence its value (as a set-valued map) is taken equal to the empty set $\emptyset$. Now by using the definition of the conjugate function, we obtain:

$$\delta^*_{\partial^w_{S_L} f(\overline{x})}(h, -\|h\|) = sup\{\langle (x^*, c)(h, -\|h\|)\rangle - \delta_{\partial^w_{S_L} f(\overline{x})}(x^*, c) : (x^*, c) \in \mathbb{R}^n \times \mathbb{R}\}$$

$$= \begin{cases} \sup\{\langle (x^*, c), (h, -\|h\|)\rangle - \delta_{\partial^w_{S_L} f(\overline{x})}(x^*, c) : (x^*, c) \in \partial^w_{S_L} f(\overline{x})\} \text{ if } h \in S - \overline{x}, \\ -\infty \text{ otherwise,} \end{cases}$$

which completes the proof. $\square$

The next theorem establishes the sum rule for the weak subdifferential set for the Clarke directionally differentiable functions.

Theorem 3.19. ***(Sum Rule for the Clarke Directionally Differentiable Functions)*** *Let $S \subseteq \mathbb{R}^n$ be star-shaped at $\overline{x} \in S$ and let Assumption 3.6 be satisfied for functions $f_i : S \to \mathbb{R}, i = 1, 2$. Then*

$$\partial^w_{S_L}(f_1 + f_2)(\overline{x}) = \partial^w_{S_L} f_1(\overline{x}) + \partial^w_{S_L} f_2(\overline{x}). \tag{3.41}$$

Proof. Under the hypotheses of the theorem, functions f_1 and f_2 both are Clarke directionally differentiable at $\overline{x} \in S$. By the definition, the Clarke directional derivative satisfies

$$(f_1 + f_2)^\circ(\overline{x}; \cdot) \leq f_1^\circ(\overline{x}; \cdot) + f_2^\circ(\overline{x}; \cdot).$$

Using the "max relation (3.30), we obtain:

$$\begin{aligned} & \max\{\langle (x^*, c)(h, -\|h\|)\rangle : (x^*, c) \in \partial^w_{S_L}(f_1 + f_2)(\overline{x})\} \\ \leq \; & \max\{\langle (x_1^*, c_1)(h, -\|h\|)\rangle : (x_1^*, c_1) \in \partial^w_{S_L} f_1(\overline{x})\} \\ + \; & \max\{\langle (x_2^*, c_2)(h, -\|h\|)\rangle : (x_2^*, c_2) \in \partial^w_{S_L} f_2(\overline{x})\} \end{aligned}$$

or by using (3.40) and the compactness of the weak subdifferential sets, we obtain:

$$\sigma_{\partial^w_{S_L}(f_1+f_2)(\overline{x})}(h, -\|h\|) \leq \sigma_{\partial^w_{S_L} f_1(\overline{x})}(h, -\|h\|) + \sigma_{\partial^w_{S_L} f_2(\overline{x})}(h, -\|h\|).$$

Now considering relation (3.39), we deduce

$$\delta^*_{\partial^w_{S_L}(f_1+f_2)(\overline{x})}(h, -\|h\|) \leq \delta^*_{\partial^w_{S_L} f_1(\overline{x})}(h, -\|h\|) + \delta^*_{\partial^w_{S_L} f_2(\overline{x})}(h, -\|h\|). \tag{3.42}$$

Since

$$\delta^*_{C+D}(x) = \delta^*_C(x) + \delta^*_D(x) \text{ for any two sets } C, D \subseteq \mathbb{R}^n,$$

we obtain

$$\delta^*_{\partial^w_{S_L} f_1(\overline{x})}(h, -\|h\|) + \delta^*_{\partial^w_{S_L} f_2(\overline{x})}(h, -\|h\|) = \delta^*_{\partial^w_{S_L} f_1(\overline{x}) + \partial^w_{S_L} f_2(\overline{x})}(h, -\|h\|).$$

and hence by (3.42) we get

$$\delta^*_{\partial^w_{S_L}(f_1+f_2)(\overline{x})}(h, -\|h\|) \leq \delta^*_{\partial^w_{S_L} f_1(\overline{x}) + \partial^w_{S_L} f_2(\overline{x})}(h, -\|h\|).$$

Now taking the conjugates leads to:

$$\delta^{**}_{\partial^w_{S_L}(f_1+f_2)(\overline{x})}(x^*,c) \geq \delta^{**}_{\partial^w_{S_L}f_1(\overline{x})+\partial^w_{S_L}f_2(\overline{x})}(x^*,c).$$

Since the indicator function of the weak subdifferential is convex and lower semicontinuous for a given function f (because, the weak subdifferential $\partial^w_{S_L}f(\overline{x})$ is convex and compact), by the Fenchel biconjugate theorem (see, e.g. [49, Theorem 11.1, page 474]) we have:

$$\delta_{\partial^w_{S_L}f(\overline{x})} = \delta^{**}_{\partial^w_{S_L}f(\overline{x})}.$$

Hence, considering the last two relations, we get

$$\delta_{\partial^w_{S_L}(f_1+f_2)(\overline{x})}(x^*,c) \geq \delta_{\partial^w_{S_L}f_1(\overline{x})+\partial^w_{S_L}f_2(\overline{x})}(x^*,c)$$

which leads to

$$\partial^w_{S_L}(f_1+f_2)(\overline{x}) \subset \partial^w_{S_L}f_1(\overline{x}) + \partial^w_{S_L}f_2(\overline{x}).$$

On the other hand, if $(x_1^*,c_1) \in \partial^w_{S_L}f_1(\overline{x}), (x_2^*,c_2) \in \partial^w_{S_L}f_2(\overline{x})$, we get

$$f_1(x) - f_1(\overline{x}) \geq \langle x_1^*, x-\overline{x}\rangle - c_1\|x-\overline{x}\| \text{ for all } x \in S$$

and

$$f_2(x) - f_2(\overline{x}) \geq \langle x_2^*, x-\overline{x}\rangle - c_2\|x-\overline{x}\| \text{ for all } x \in S.$$

By adding the last two inequalities we obtain

$$\partial^w_{S_L}f_1(\overline{x}) + \partial^w_{S_L}f_2(\overline{x}) \subset \partial^w_{S_L}(f_1+f_2)(\overline{x}),$$

which completes the proof. □

Now we formulate the sum rule for the weak subdifferential of tangentially convex functions. We will show that this property is satisfied for the class of functions satisfying the conditions identified in Assumption 3.11.

Theorem 3.20. ***(Sum Rule for the Tangentially Convex Functions)*** *Let $S \subseteq \mathbb{R}^n$ be star-shaped at $\overline{x} \in S$. Assume that all conditions of Theorem 3.13 are satisfied for functions $f_i : S \to \mathbb{R}, i = 1, 2$. Then*

$$\partial^w_{S_L}(f_1+f_2)(\overline{x}) = \partial^w_{S_L}f_1(\overline{x}) + \partial^w_{S_L}f_2(\overline{x}). \tag{3.43}$$

Proof. Under hypotheses of the theorem, functions f_1 and f_2 both are directionally differentiable at $\overline{x} \in S$. By the definition, the directional derivative satisfies

$$(f_1 + f_2)'(\overline{x}; \cdot) = f_1'(\overline{x}; \cdot) + f_2^{\circ}(\overline{x}; \cdot).$$

Using the "max relation" (3.36) we obtain:

$$\begin{aligned} & \max\{\langle (x^*, c)(h, -\|h\|)\rangle : (x^*, c) \in \partial^w_{S_L}(f_1 + f_2)(\overline{x})\} \\ = \; & \max\{\langle (x_1^*, c_1)(h, -\|h\|)\rangle : (x_1^*, c_1) \in \partial^w_{S_L} f_1(\overline{x})\} \\ + \; & \max\{\langle (x_2^*, c_2)(h, -\|h\|)\rangle : (x_2^*, c_2) \in \partial^w_{S_L} f_2(\overline{x})\}. \end{aligned}$$

The remaining part of the proof is similar to that of Theorem 3.19, where all subsequent inequalities will be satisfied in the form of equalities. □

In the next section, we apply the sum rule for the weak subdifferential obtained in this section, to generalize the well-known result of convex analysis on the intersection of two convex sets to a nonconvex case.

3.4. Augmented Normal Cones to an Intersection of Two Sets

A well-known result of convex analysis states that the normal cone to the intersection of two convex sets, equals the sum of normal cones to these sets (see e.g. [8, Theorem 4.3.4, page 130]). In this section, by applying the sum rule for the weak subdifferential of the indicator functions of two (nonconvex) sets, we obtain the parallel result for the augmented normal cones to the intersection of nonconvex sets.

The following lemma explains the relation between the weak subdifferential of the indicator function and the augmented normal cone to a given set.

Lemma 3.21. *Let δ_C be the indicator function of a set $C \subset \mathbb{R}^n$ and $\overline{x} \in C$, then $\partial^w \delta_C(\bar{x}) = N^a_C(\bar{x})$.*

Proof. Assume that $(x^*, \alpha) \in \partial^w \delta_C(\overline{x})$ then it implies:

$$\delta_C(z) \geq \delta_C(\bar{x}) + \langle x^*, z - \bar{x}\rangle - \alpha\|z - \bar{x}\| \qquad \text{for all } z \in \mathbb{R}^n.$$

or equivalently,

$$\langle x^*, z - \bar{x}\rangle - \alpha\|z - \bar{x}\| \leq 0 \qquad \text{for all } z \in C.$$

Thus $(x^*, c) \in N^a_C(\bar{x})$. The reverse direction similarly follows from definitions of the weak subdifferential and the augmented normal cone. □

Now, we show that if C is an arbitrary star-shaped set with $\operatorname{core}(C) \neq \emptyset$, then indicator function $\delta'_C(\overline{x}; x-\overline{x})$ of this set satisfies all the conditions of Theorem 3.13.

Lemma 3.22. *Let $C \subset \mathbb{R}^n$ with core$(C) \neq \emptyset$ and let C be star-shaped at $\overline{x} \in$ core(C). Then the indicator function δ_C of this set and its directional derivative $\delta'_C(\overline{x}; \cdot)$ satisfy all the conditions of Theorem 3.13.*

Proof. Let $\overline{x} \in \operatorname{core}(C)$. We first show that $\delta'_C(\overline{x}; \cdot)$ is a convex function on $\mathbb{R}^n$. The definition of the directional derivative implies that:

$$\delta'_C(\overline{x}; x-\overline{x}) = \lim_{\lambda \to 0^+} \frac{\delta_C(\overline{x} + \lambda(x-\overline{x})) - \delta_C(\overline{x})}{\lambda} \quad \text{for all } x \in \mathbb{R}^n. \tag{3.44}$$

Clearly, for every $x \in \mathbb{R}^n$, since $\overline{x} \in \operatorname{core}(C)$, we have $\overline{x}+\lambda(x-\overline{x}) \in C$ for sufficiently small positive λ. This implies that $\delta'_C(\overline{x}; x-\overline{x}) = 0$ for every $x \in \mathbb{R}^n$. Hence $\delta'_C(\overline{x}; x-\overline{x})$ is a convex function of x and thus, δ_C is tangentially convex. Since $\delta'_C(\overline{x}; \cdot)$ is a convex function on the whole $\mathbb{R}^n$, it is continuous. Finally, since $\delta'_C(\overline{x}; x-\overline{x}) = 0$ for every $x \in \mathbb{R}^n$, the relations (3.32) and (3.33) are both satisfied. Thus, the proof is completed. □

Now we can prove the main result of this section.

Theorem 3.23. *Let C_1 and C_2 be two nonempty subsets of $\mathbb{R}^n$ with core$(C_i) \neq \emptyset, i = 1, 2$. Assume that C_1 and C_2 are both star-shaped at $\overline{x} \in$ core$(C_1 \cap C_2)$. Then*

$$N^a_{C_1}(\overline{x}) + N^a_{C_2}(\overline{x}) = N^a_{C_1 \cap C_2}(\overline{x}). \tag{3.45}$$

Proof. By Lemma 3.21, the relation (3.45) is equivalent to the following equality:

$$\partial^w \delta_{C_1 \cap C_2}(\overline{x}) = \partial^w \delta_{C_1}(\overline{x}) + \partial^w \delta_{C_2}(\overline{x}). \tag{3.46}$$

Lemma 3.22 states that the indicator functions δ_{C_1} and δ_{C_2} both satisfy the conditions of Theorem 3.20. By this theorem, we deduce

$$\partial^w \delta_{C_1}(\overline{x}) + \partial^w \delta_{C_2}(\overline{x}) = \partial^w (\delta_{C_1}(\overline{x}) + \delta_{C_2}(\overline{x})).$$

By definition of indicator functions, for any two sets C_1 and C_2, we have:

$$\delta_{C_1}(x) + \delta_{C_2}(x) = \delta_{C_1 \cap C_2}(x) \text{ for all } x \in \mathbb{R}^n. \tag{3.47}$$

Assume that $(x^*, c) \in \partial^w \delta_{C_1 \cap C_2}(\overline{x})$. The definition of the weak subdifferential yields to:

$$\delta_{C_1 \cap C_2}(x) - \delta_{C_1 \cap C_2}(\overline{x}) \geq \langle x^*, x - \overline{x} \rangle - c\|x - \overline{x}\| \text{ for all } x \in \mathbb{R}^n.$$

By using (3.47), we get

$$\delta_{C_1}(x) + \delta_{C_2}(x) - \delta_{C_1}(\overline{x}) - \delta_{C_2}(\overline{x}) \geq \langle x^*, x - \overline{x} \rangle - c\|x - \overline{x}\| \text{ for all } x \in \mathbb{R}^n,$$

which implies that $(x^*, c) \in \partial^w(\delta_{C_1}(\overline{x}) + \delta_{C_2}(\overline{x}))$.

The proof of the converse inclusion is similar and therefore is omitted. $\square$

Note that the assumption that C_1 and C_2 are both star-shaped at $\overline{x} \in \text{core}(C_1 \cap C_2)$ is important for this theorem. We may give a counterexample that if this condition is not satisfied then the theorem may not hold.

Example 3.24. Let us consider the sets;

$$C = \{(x, y) : xy \geq 0, -1 \leq x \leq 1, -1 \leq y \leq 1\}$$

and

$$D = \{(x, y) : xy \leq 0, -1 \leq x \leq 1, -1 \leq y \leq 1\}$$

It is obvious that

$$C \cap D = \{(0, y) : |y| \leq 1\} \cup \{(x, 0) : |x| \leq 1\}$$

Then, if we calculate the augmented normal cone of the set C at $\mathbf{x} = \mathbf{0}$:

$$\begin{aligned} N_C^a(\mathbf{0}) =& \{(x^*, \alpha) | \langle x^*, x \rangle - \alpha\|x\| \leq 0, \quad \forall x \in C\} \\ =& \{(x^*, \alpha) | x_1^* \cdot x_1 + x_2^* x_2 - \alpha\sqrt{x_1^2 + x_2^2} \leq 0, \forall x \in C\} \end{aligned}$$

Considering the following points from the set C:

$$\mathbf{x} = (1,0) \in C \text{ yields to } x_1^* - \alpha \leq 0,$$

$$\mathbf{x} = (0,1) \in C \text{ yields to } x_2^* - \alpha \leq 0,$$

$$\mathbf{x} = (0,1) \in C \text{ yields to } x_1^* + x_2^* - \alpha\sqrt{2} \leq 0,$$

$$\mathbf{x} = (-1,0) \in C \text{ yields to } -x_1^* - \alpha \leq 0,$$

$$\mathbf{x} = (0,-1) \in C \text{ yields } to - x_2^* - \alpha \leq 0,$$

$$\mathbf{x} = (-1,-1) \in C \text{ yields to } -x_1^* - x_2^* - \alpha\sqrt{2} \leq 0.$$

By using the above inequalities, we obtain that:

$x_1^* - \alpha \leq 0$ and $-x_1^* - \alpha \leq 0$ implies that $|x_1^*| \leq \alpha$,

$x_2^* - \alpha \leq 0$ and $-x_1^* - \alpha \leq 0$ implies that $|x_2^*| \leq \alpha$,

$x_1^* + x_2^* - \alpha\sqrt{2} \leq 0$ and $-x_1^* - x_2^* - \alpha\sqrt{2} \leq 0$ together implies that $|x_1^* + x_2^*| \leq \alpha\sqrt{2}$.

Thus we obtain

$$N_C^a(\mathbf{0}) = \left\{(\mathbf{x}^*, \alpha) \middle| \quad |\mathbf{x_1^*}| \leq \alpha, |\mathbf{x_2^*}| \leq \alpha \text{ and } |\mathbf{x_1^*} + \mathbf{x_2^*}| \leq \alpha\sqrt{\mathbf{2}}\right\}.$$

Similary, let us calculate the augmented normal cone of the set D at $\mathbf{x} = \mathbf{0}$:

$$N_D^a(\mathbf{0}) = \{(\mathbf{x}^*, \alpha) | \langle \mathbf{x}^*, \mathbf{x} \rangle - \alpha \|\mathbf{x}\| \leq \mathbf{0}, \quad \forall \mathbf{x} \in \mathbf{D}\}$$
$$= \{(x^*, \alpha) | x_1^* \cdot x_1 + x_2^* \cdot x_2 - \alpha\sqrt{x_1^2 + x_2^2} \leq 0, \forall x \in D\}.$$

$$\mathbf{x} = (1,0) \in D \text{ yields to } x_1^* - \alpha \leq 0,$$

$$\mathbf{x} = (0,1) \in D \text{ yields to } x_2^* - \alpha \leq 0,$$

$$\mathbf{x} = (1,1) \in D \text{ yields to } x_1^* - x_2^* - \alpha\sqrt{2} \leq 0,$$

$$\mathbf{x} = (-1,0) \in D \text{ yields to } s - x_1^* - \alpha \leq 0,$$

$$\mathbf{x} = (0,-1) \in D \text{ yields to } -x_2^* - \alpha \leq 0,$$

$$\mathbf{x} = (-1,-1) \in D \text{ yields to } -x_1^* + x_2^* - \alpha\sqrt{2} \leq 0.$$

By using the above inequalities that we obtain:

$$x_1^* - \alpha \leq 0 \text{ and } - x_1^* - \alpha \leq 0 \text{ we have } |x_1^*| \leq \alpha,$$
$$x_2^* - \alpha \leq 0 \text{ and } - x_1^* - \alpha \leq 0 \text{ we have } |x_2^*| \leq \alpha,$$
$$x_1^* - x_2^* - \alpha\sqrt{2} \leq 0 \text{ and } - x_1^* + x_2^* - \alpha\sqrt{2} \leq 0 \text{ we have } |x_1^* - x_2^*| \leq \alpha\sqrt{2}.$$

Thus, we get

$$N_D^a(\mathbf{0}) = \{(\mathbf{x}^*, \alpha)| \quad |\mathbf{x_1^*}| \leq \alpha, |\mathbf{x_2^*}| \leq \alpha \text{ and } |\mathbf{x_1^*} - \mathbf{x_2^*}| \leq \alpha\sqrt{\mathbf{2}}\}.$$

Augmented normal cone of $C \cap D$ can be calculated as follows:

$$\begin{aligned} N_{C\cap D}^a(\mathbf{0}) =& \{(x^*, \alpha)|\langle x^*, x\rangle - \alpha\|x\| \leq 0, \quad \forall \mathbf{x} \in C \cap D\} \\ =& \{(x^*, \alpha)|x_1^* \cdot x_1 + x_2^* \cdot x_2 - \alpha\sqrt{x_1^2 + x_2^2} \leq 0, \quad \forall x \in C \cap D\}. \end{aligned}$$

$$\mathbf{x} = (1, 0) \in C \cap D \text{ yields to } x_1^* - \alpha \leq 0,$$
$$\mathbf{x} = (0, 1) \in C \cap D \text{ yields to } x_2^* - \alpha \leq 0,$$
$$\mathbf{x} = (-1, 0) \in C \cap D \text{ yields to } - x_1^* - \alpha \leq 0,$$
$$\mathbf{x} = (0, -1) \in C \cap D \text{ yields to } - x_2^* - \alpha \leq 0.$$

By using the above inequalities that we get:

$$x_1^* - \alpha \leq 0 \text{ and } - x_1^* - \alpha \leq 0 \text{ yields to } |x_1^*| \leq \alpha,$$
$$x_2^* - \alpha \leq 0 \text{ and } - x_1^* - \alpha \leq 0 \text{ yields to } |x_2^*| \leq \alpha.$$

Thus, $N_{C\cap D}^a(\mathbf{0}) = \{(\mathbf{x}^*, \alpha)| \quad |\mathbf{x_1^*}| \leq \alpha \text{ and } |\mathbf{x_2^*}| \leq \alpha\}$.

It is clear that $N_{C\cap D}^a(\mathbf{0}) \subset \mathbf{N_C^a}(\mathbf{0}) + \mathbf{N_D^a}(\mathbf{0})$. But the reverse direction is not true. Let $(\alpha, (\sqrt{2}-1)\alpha) \in N_C^a(\mathbf{0})$ and $(\alpha, (1-\sqrt{2})\alpha) \in N_D^a(\mathbf{0})$ but $(\alpha, (\sqrt{2}-1)\alpha) + (\alpha, (1-\sqrt{2})\alpha) = (2\alpha, 0) \notin N_{C\cap D}^a(\mathbf{0})$.

4. DUALITY IN NONCONVEX OPTIMIZATION

In this chapter, we aim to present an explicit formulation of the dual problem and to establish zero duality gap conditions for inequality constrained optimization problems, in terms of positively homogeneous and lower semicontinuous objective and constraint functions which are defined on a conic set. Formulating conditions for the dual problems with inequality constraints in terms of the objective and the constraint functions is not an obvious task. We investigate the zero duality gap condition involving the objective function, the constraint functions ,and the set without any convexity assumption. Hence, we are interested in the conditions on functions f, g, and set S such that the strong duality relation holds.

We consider a nonlinear programming problem in the following form:

$$(P) \quad \begin{cases} \inf\limits_{x \in X} f(x) \end{cases}$$

where

$$X = \{x \in S : g_j(x) \leq 0, j = 1, \ldots, m\} \neq \emptyset.$$

Remark 5. We use the notation $g(x) \leqq y$ to represent $g_j(x) \leq y_j$, for all $j = 1, \ldots, m$ for the sake of simplicity.

Problem (P) is called the primal problem. The infimum for the problem (P) will be denoted $Inf(P)$ and every element $x \in X$ such that $f(x) = Inf(P)$ will be called an optimal solution of (P).

Let G ,i.e, the graph of a perturbed set of feasible solutions, be defined as a set-valued mapping from $\mathbb{R}^m$ into $\mathbb{R}^n$ in the following form:

$$G(y) = \{x \in S : g_j(x) \leq y_j, j = 1, \ldots, m\}.$$

For the problem (P), the associated dualizing parameterization function Φ is defined as follows:

$$\Phi(x, y) = \begin{cases} f(x) & \text{if } x \in S \text{ and } g(x) \leqq y, \\ +\infty & \text{otherwise.} \end{cases}$$

The dualizing parameterization function $\Phi : X \times Y \to \overline{\mathbb{R}}$ defined above satisfies that $\Phi(x, 0) = f(x)$. Then, it is obvious that $\inf\limits_{x \in X} \Phi(x, 0) = Inf(P)$.

We define the perturbation function associated with problem (P) as follows:

$$h(y) = \inf_{x \in S} \Phi(x, y).$$

Using the classical technique to form the dual of a minimization problem [17,45] we form the dual problem as follows:

$$\sup_{(y^*, d) \in Y^* \times \mathbb{R}_+} \{-\Phi^w(0, y^*, d)\} \tag{4.48}$$

is called the dual problem of (P) with respect to Φ. The supremum for problem (P^*) is denoted by $Sup(P^*)$, and any element $(y^*, d) \in Y^* \times \mathbb{R}_+$ such that $\Phi^w(0, y^*, d) = Sup(P^*)$ is called an optimal solution to (P^*).

$$\begin{aligned} \Phi^w(0, y^*, \beta) &= \sup_{(x,y) \in \mathbb{R}^n \times \mathbb{R}^m} \{-\beta \|y\| + \langle y, y^* \rangle - \Phi(x, y)\}, \\ &= \sup_{x \in S} \sup_{y \in \mathbb{R}^m, g(x) \leqq y} \{-\beta \|y\| + \langle y, y^* \rangle - f(x)\}, \end{aligned}$$

where the Euclidean norm is used. Letting $y = g(x) + z, z \in \mathbb{R}^m_+$, we obtain:

$$= \sup_{x \in S} \sup_{z \in \mathbb{R}^m_+} \{-\beta \|g(x) + z\| + \langle g(x) + z, y^* \rangle - f(x)\},$$

To construct the dual problem, we need to calculate $\sup_{(y^*, \beta) \in \mathbb{R}^m \times \mathbb{R}_+} \{-\Phi^w(0, y^*, \beta)\}$. By using the representation

$$\|g(x) + z\| = sup\{\langle g(x) + z, e \rangle : e \in \mathbb{R}^m, \|e\| \leq 1\},$$

we obtain

$$\begin{aligned} &\sup_{(y^*, \beta) \in \mathbb{R}^m \times \mathbb{R}_+} \{-\Phi^w(0, y^*, \beta)\} \\ &= \sup_{(y^*, \beta) \in \mathbb{R}^m \times \mathbb{R}_+} \inf_{x \in S} \{f(x) + \inf_{z \in \mathbb{R}^m_+} [-\langle y^*, g(x) + z \rangle + \sup_{\|e\| \leq 1} \beta \langle g(x) + z, e \rangle]\}. \\ &= \sup_{(y^*, \beta) \in \mathbb{R}^m \times \mathbb{R}_+} \inf_{x \in S} \{f(x) - \langle y^*, g(x) \rangle + \inf_{z \in \mathbb{R}^m_+} \sup_{\|e\| \leq 1} [\beta \langle g(x), e \rangle] + \langle \beta e - y^*, z \rangle\} \\ &= \sup_{(y^*, \beta) \in \mathbb{R}^m \times \mathbb{R}_+} \inf_{x \in S} \{f(x) - \langle y^*, g(x) \rangle + \sup_{\|e\| \leq 1} \inf_{z \in \mathbb{R}^m_+} [\beta \langle g(x), e \rangle] + \langle \beta e - y^*, z \rangle\}. \end{aligned}$$

Letting

$$V = [(y^*, \beta) \in \mathbb{R}^m \times \mathbb{R}_+ : \beta e - y^* \in \mathbb{R}^m_+, \text{ for some } e \in \mathbb{R}^m, \|e\| \leq 1]$$

and

$$\tau(x, y^*, \beta) = \sup\{\beta\langle g(x), e\rangle : e \in \mathbb{R}^m \text{ such that} \|e\| \leq 1, \beta e - y^* \in \mathbb{R}^m_+\}$$

we can now formulate the dual problem $(P)^*$ as follows:

$$(P^*) \qquad \sup_{(y^*,\beta)\in V} \inf_{x\in S}\{f(x) - \langle y^*, g(x)\rangle + \tau(x, y^*, \beta)\}.$$

. In this formulation, the function $L : (\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}_+) \to \overline{\mathbb{R}}$ defined as

$$L(x, y^*, \beta) = f(x) - \langle y^*, g(x)\rangle + \tau(x, y^*, \beta),$$

is called the sharp augmented Lagrangian function associated with problem (P). The function

$$H(y^*, \beta) = \inf_{x\in S} L(x, y^*, \beta)$$

is called the dual function. then, the dual problem $(P)^*$ can also be written in the following form:

$$(P^*) \qquad \sup_{(y^*,\beta)\in V} H(y^*, \beta)$$

The uniform level boundedness definition is given in [49, Definition 1.16] and the zero duality gap condition was established by using this definition.

Definition 4.1. (uniform level boundedness) For each $\overline{y} \in \mathbb{R}^m$ and $\alpha \in \mathbb{R}$, a function $\Phi(x, y)$ from $\mathbb{R}^n \times \mathbb{R}^m$ into $\overline{\mathbb{R}}$ is level-bounded in x locally uniformly in y if there exists a neighborhood $V \subset N(\overline{y})$ together with a bounded set $B \subset \mathbb{R}^n$ satisfying $\{x : \Phi(x, y) \leq \alpha\} \subset B$ for all $y \in V$; or equivalently, there exists a neighborhood $V \subset N(\overline{y})$ such that the set $\{(x, y) : y \in V, \Phi(x, y) \leq \alpha\}$ is bounded in $\mathbb{R}^n \times \mathbb{R}^m$.

In the following theorem, the zero duality gap conditions are formulated over the dualizing parametrization functions for the given problem (P) [49, Theorem 11.59].

Theorem 4.2. *(duality without convexity) We are interested in a problem of minimizing f on $\mathbb{R}^n$, and consider the augmented Lagrangian $\bar{l}(x, u, r)$ associated with a dualizing parameterization $\Phi : \mathbb{R}^n \times \mathbb{R}^m \to \overline{\mathbb{R}}$, ($f = \Phi(\cdot, 0)$,) and some augmenting function $\sigma : \mathbb{R}^m \to \overline{\mathbb{R}}$. Suppose that the dualizing parameterization function $\Phi(x, y)$ is level-bounded in x locally uniformly in y, and let the perturbation function $h(y) := \inf_x \Phi(x, y)$. Let $\inf_x \bar{l}(x, u, r) > -\infty$ for at least one $(u, r) \in \mathbb{R}^m \times (0, \infty)$.*

Then

$$f(x) = \sup_{u,r} \bar{l}(x,u,r), \qquad \overline{\psi}(u,r) = \inf_{u,r} \bar{l}(x,u,r), \tag{4.49}$$

where actually $f(x) = \sup_{u} \bar{l}(x,u,r)$ *for every* $r > 0$, *and in fact*

$$\inf_x f(x) = \inf_x [\sup_{u,r} \bar{l}(x,u,r)] = \sup_{u,r}[\inf_x \bar{l}(x,u,r)] = \sup_{u,r} \overline{\psi}(u,r). \tag{4.50}$$

Furthermore, optimal solutions to the primal and augmented dual problems are characterized as saddle points of the augmented Lagrangian:

$$\begin{aligned} &\overline{x} \in \arg\min_x f(x) \quad \text{and } (\overline{u},\overline{r}) \in \arg\max_{u,r} \overline{\psi}(u,r), \\ \Leftrightarrow \quad &\inf_x \bar{l}(x,\overline{u},\overline{r}) = \bar{l}(\overline{x},\overline{u},\overline{r}) = \sup_{u,r} \bar{l}(\overline{x},u,r), \end{aligned}$$

the elements of $\arg\max_{u,r} \overline{\psi}(u,r)$ *are the pairs* $(\overline{u},\overline{r})$ *satisfying the following property*

$$h(y) \geq h(0) + \langle \overline{u}, y \rangle - \overline{r}\sigma(y) \quad \forall y. \tag{4.51}$$

The following theorem gives the relation between the zero duality gap condition and the weakly subdifferentiable perturbation function at origin.

Theorem 4.3. *([4, Theorem 5]) Suppose that the function* $h(0)$ *is finite and* $\partial^w h(0) \neq \emptyset$. *Then*

i. $inf(P) = sup(P^*)$.

ii. *Any weak subgradient of* h *at* $0 \in Y$ *is an optimal solution of* (P^*).

The following Theorem and Corollary investigate the relation between the lower Lipschitz functions and the weak subdifferentiability.

Theorem 4.4. *[4, Theorem 1] Assume that the function* $h : \mathbb{Y} \to (-\infty, +\infty]$ *is finite at a point* y_0, *then the following properties are equivalent:*

i. h *is lower Lipschitz at* y_0.

ii. h *is weakly subdifferentiable at* y_0.

iii. h *is lower locally Lipschitz at* y_0 *and there exist numbers* $p \geq 0$ *and* q *such that*

$$h(y) \geq -p\|y\| + q \text{ for all } y \in \mathbb{Y}.$$

The following definition of lower semicontinuity of a set-valued mapping F is introduced in [2].

Definition 4.5. ([2, Definition 2, page 108]) We say that F is lower semicontinuous at $\overline{x} \in X$ if for any $\overline{y} \in F(\overline{x})$ and any neighborhood $\mathscr{N}(\overline{y})$ of $\overline{y}$, there exists a neighborhood $\mathscr{N}(\overline{x})$ of $\overline{x}$ such that

$$F(x) \cap \mathscr{N}(\overline{y}) \neq \emptyset, \quad \forall x \in \mathscr{N}(\overline{x}).$$

We say that F is lower semicontinuous if F is lower semicontinuous at every $\overline{x} \in X$.

The following definition of upper semicontinuity of a set-valued mapping F is introduced in [2].

Definition 4.6. ([2, Definition 1, page 108]) We say that F is upper semicontinuous at $\overline{x} \in X$ if for any neighborhood $\mathscr{N}(F(\overline{x}))$ of $F(\overline{x})$, there exist a neighborhood $\mathscr{N}(\overline{x})$ of $\overline{x}$ such that

$$F(x) \subset \mathscr{N}(F(\overline{x})), \quad \forall x \in \mathscr{N}(\overline{x}).$$

We say that F is upper semicontinuous if F is upper semicontinuous at every $\overline{x} \in X$.

We give the following examples to demonstrate the given set-valued mapping definitions.

Example 4.7. Consider the following set-valued mapping H. The mapping H from $\mathbb{R}$ into $\mathbb{R}^2$ is defined as following:

$$H(x) = \begin{cases} [0, +\infty) & \text{if } x = 0, \\ 0 & \text{if } x \neq 0. \end{cases}$$

The graph of set-valued mapping H is follows:

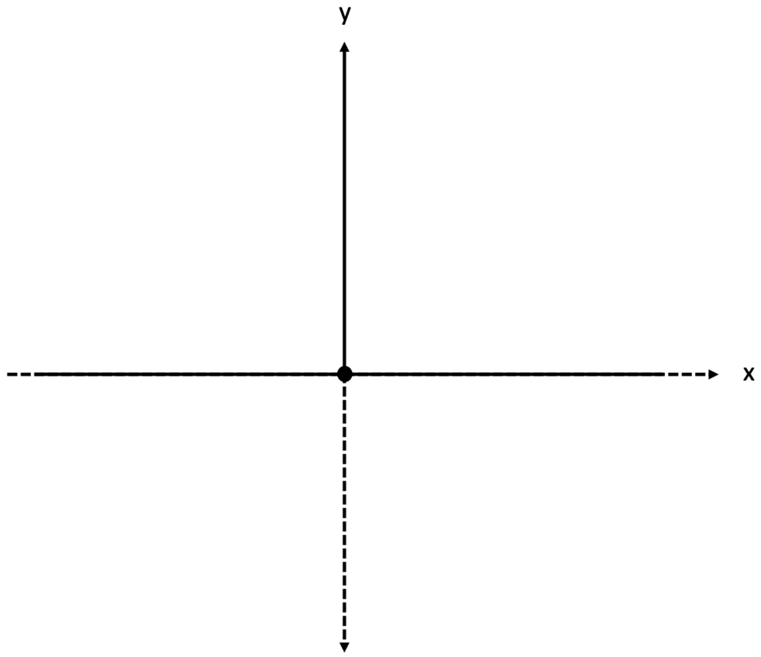

Figure 4.1. *The graph of an upper semicontinuous set-valued mapping H*

The set-valued mapping H whose graph is given in Figure 4.1 is upper semicontinuous but not lower semicontinuous. Consider the point $\overline{x} = 0$, then $H(\overline{x}) = H(0) = [0, +\infty)$. If we take $\overline{y} = 1 \in H(0) = [0, +\infty)$ and let $\varepsilon = \dfrac{1}{2}$, it yields to $\mathscr{N}(\overline{y}, \varepsilon) = \mathscr{N}(1, \varepsilon) = \left(\dfrac{1}{2}, \dfrac{3}{2}\right)$. It is clear that for any neighborhood of $x = 0$ and for all $\tilde{x} \in \mathscr{N}(0) \setminus \{0\}$ we have $H(\tilde{x}) = 0$. It follows that $H(\tilde{x}) \cap \left(\dfrac{1}{2}, \dfrac{3}{2}\right) = \emptyset$. Thus $\forall x \in \mathscr{N}(\overline{x})$, H does not satisfy the condition

$$H(x) \cap \mathscr{N}(\overline{y}) \neq \emptyset,$$

which proves that H is not lower semicontinuous.

Example 4.8. Consider the following set-valued mapping F. The mapping F from $\mathbb{R}$ into $\mathbb{R}^2$ is defined by

$$F(x) = \begin{cases} \left[-\dfrac{1}{2}, \dfrac{1}{2}\right], & \text{if } x = 0, \\ 0, & \text{if } x \neq 0. \end{cases}$$

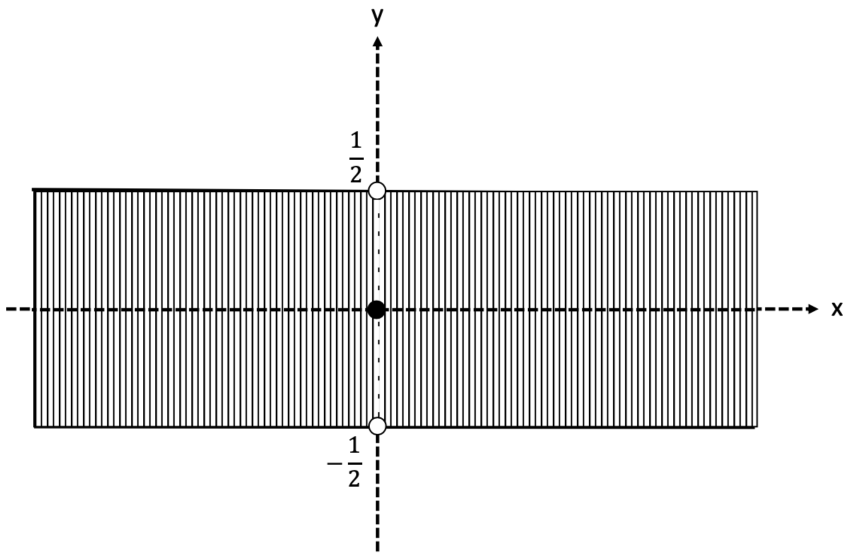

Figure 4.2. *The graph of a lower semicontinuous set-valued mapping F*

The set-valued mapping F whose graph is given in Figure 4.2 is not upper semicontinuous at $\overline{x} = 0 \in X$. For any neighborhood of $F(0)$ we have to find a neighborhood of $\overline{x} = 0$ such that

$$F(x) \subset \mathscr{N}(F(0)) \text{ for all } x \in \mathscr{N}(0).$$

It is clear that letting $\varepsilon = \dfrac{1}{2}$, the neighborhood of $F(0)$, $\mathscr{N}(F(0), \varepsilon)$, yields to $\left(-\dfrac{1}{2}, \dfrac{1}{2}\right)$. We know that $\tilde{x} = \dfrac{1}{4}$ belongs to the ε neighborhood of $\overline{x} = 0$ and $F(\tilde{x}) = \left\{\dfrac{-1}{2}, \dfrac{1}{2}\right\}$. However, $F(\tilde{x}) \not\subset \mathscr{N}(F(0), \varepsilon) = \left(-\dfrac{1}{2}, \dfrac{1}{2}\right)$. Indeed, when $\varepsilon < \dfrac{1}{2}$, there does not exist any neighborhood of $\overline{x} = 0$ satisfying

$$F(x) \subset \mathscr{N}(F(0)) \text{ for all } x \in \mathscr{N}(0),$$

which shows that F is not upper semicontinuous set-valued mapping.

4.1. Zero Duality Gap Condition for Problems with Positively Homogeneous Objective and Constraint Functions

Verifying the zero duality gap conditions formulated in terms of perturbation and dualizing parameterization functions, nor deriving the optimality conditions formulated in terms of objective and constraint functions are not easy tasks. In this subsection, we consider positively homogeneous objective and constraint functions defined on a cone S. Under this given conditions, we obtain that the perturbation function is positively homogeneous as well. We prove that if the perturbation function is positively homogeneous and continuous function then it is Lipschitz. Then by using the Theorem 2.33, it yields that the perturbation function is weakly subdifferentiable at origin. Hence by following Theorem 4.3, the zero duality gap condition satisfies under the given conditions.

The following theorem is proved in [33, Lemma 2.7].

Lemma 4.9. *Let $f : X \to \mathbb{R}$ be a positively homogeneous function bounded below on some neighborhood of zero. Then f is the weakly subdifferentiable at 0_X.*

In the following, lemma we show that the perturbation function corresponding to the problem (P) is positively homogeneous. In establishing the zero duality gap condition this theorem will be helpful.

Lemma 4.10. *Let S be a cone in $\mathbb{R}^n$. Assume that $f : S \to \overline{\mathbb{R}}$ is a positively homogeneous function and let the mapping $g = (g_1, g_2, \dots, g_m) : S \to \mathbb{R}^m$ be positively homogeneous functions. Then the perturbation function $h(y)$ is positively homogeneous.*

Proof. We have to show that $h(\lambda y) = \lambda h(y)$. By the definition of the perturbation function, we have:

$$h(\lambda y) = f(x) \text{ if } x \in S \text{ and } g(x) \leq \lambda y$$

Since g is positively homogeneous function it follows that,

$$h(\lambda y) = f(x) \text{ if } x \in S \text{ and } \frac{1}{\lambda} g(x) = g(\frac{1}{\lambda} x) \leq y.$$

Now, if we denote $\frac{1}{\lambda} x = z$, then we obtain,

$$h(\lambda y) = f(\lambda z) \text{ if } \lambda z \in S \text{ and } g(z) \leq y.$$

By the assumption that f is positively homogeneous function and S is a cone, then it yields that:

$$h(\lambda y) = \lambda f(z) \text{ if } z \in S \text{ and } g(z) \leq y.$$

Hence,

$$h(\lambda y) = \lambda h(y)$$

which proves that the perturbation function $h(y)$ is positively homogeneous.

$\square$

The following theorem shows that the set valued mapping G is continuous under the assumptions that S is a closed set and the constraint function g is a lower semicontinuous function.

Theorem 4.11. *Assume that S is a closed set. If g is a lower semicontinuous function then the set valued mapping G is continuous.*

Proof. Let $y_n \to \overline{y}$, $x_n \in G(y_n)$ and $x_n \to \overline{x}$. Then we need to show that $\overline{x} \in G(\overline{y})$. Since $x_n \to \overline{x}$, it implies that $x_n \in S$ and $g(x_n) \leq y_n$. Passing to the limit as $x_n \to \overline{x}$ and $y_n \to \overline{y}$ yields to

$$g(\overline{x}) \leqq \liminf_{x_n \to \overline{x}} g(x_n) \leqq \overline{y}.$$

Since S is closed we know that $\overline{x} \in S$. Thus, $\overline{x} \in S$ and $g(\overline{x}) \leqq \overline{y}$ results that $\overline{x} \in G(\overline{y})$. $\square$

We have the following lemma which gives conditions for obtaining the lower semicontinuity of the perturbation function.

Lemma 4.12. *Assume that S is a closed set. Suppose also that*

i. f is lower semicontinuous on $x \in S$.

ii. $G(0)$ is compact, g is lower semicontinuous on $x \in S$.

Then the perturbation function $h(y)$ is lower semicontinuous at $y = 0$.

Proof. Let $\varepsilon > 0$ be arbitrary positive number. We have to prove that there exists a neighborhood $\mathscr{N}(0)$ of $y = 0$ such that

$$h(0) - \varepsilon \leq h(y) \tag{4.52}$$

Since f is lower semicontinuous at $x \in S$, there exists a neighborhood $\mathscr{N}(x)$ such that

$$f(x^{'}) - \varepsilon \leq f(x), \quad \forall x^{'} \in \mathscr{N}(x) \tag{4.53}$$

or equivalently,

$$f(x^{'}) \leq f(x) + \varepsilon, \quad \forall x^{'} \in \mathscr{N}(x). \tag{4.54}$$

The compactness of $G(0)$ implies that it can be covered by n neighborhoods $\mathscr{N}(x_i)$. Thus

$$G(0) \subset \mathscr{N} := \bigcup_{i=1}^{n} \mathscr{N}(x_i).$$

Since g is lower semicontinuous on S. By Theorem 4.11 we conclude that the set valued mapping G is continuous. Thus G is upper semicontinuous set-valued mapping. Since the set-valued mapping G is upper semicontinuous then we know that there exists a neighborhood $\mathscr{N}_0(0)$ of $y = 0$ such that

$$G(y) \subset \mathscr{N}, \quad \forall y \in \mathscr{N}_0(0)$$

When y belongs to $\mathscr{N}_0(0)$ and x belongs to $G(0)$, then x belongs to $\mathscr{N}$ and thus to some $\mathscr{N}(x_i)$. For any $x^{'}$ belongs to $G(y)$ it is obvious that

$$h(0) \leq f(x^{'}), \quad \forall x^{'} \in G(y) \tag{4.55}$$

(4.54) and (4.55) together implies that

$$h(0) \leq f(x^{'}) \leq f(x) + \varepsilon \tag{4.56}$$

Taking the infimum when x ranges over $G(y)$ we obtain the inequality (4.52). $\square$

The following theorem states that a positively homogeneous and lower semicontinuous function is the weakly subdifferentiable at $x = 0$.

Theorem 4.13. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be positively homogeneous and lower semicontinuous function on the cone S. Then f is weakly subdifferentiable at $\overline{x} = 0$, that is there exists $(x^*, c) \neq (0, 0)$ where $(x^*, c) \in (\mathbb{R}^n \times \mathbb{R}) \setminus \{(0_{\mathbb{R}^n}, 0)\}$ such that*

$$\langle x^*, x \rangle - c\|x\| \leq f(x) - f(0), \quad \forall x \in S.$$

Proof. Since f is positively homogeneous it implies that $f(0) = 0$. f is bounded from below on $S_1 = \{x \in S : \|x\| = 1\}$ since f is lower semicontinuous. Consider an arbitrary element $x^* \in \mathbb{R}^n \setminus \{0\}$. Then the function $y = \langle x^*, x \rangle$ is continuous and hence is bounded from below on S_1. Then there exists a sufficiently large number $c > 0$ such that

$$\langle x^*, u \rangle - c\|u\| \leq f(u) - f(0), \quad \forall u \in S_1. \tag{4.57}$$

Consider an arbitrary element $x \in S$. Since S is a cone, then there exists $t > 0$ and

$u \in S_1$ such that $x = tu$. For this u, multiply both sides of (4.57) by $t > 0$. Then it yields to

$$\langle x^*, x\rangle - c\|x\| \leq f(x) - f(0), \quad \forall x \in S.$$

The proof is completed. $\square$

In the following theorem we show that if objective function f and the set-valued mapping G are continuous then the perturbation function p is upper semicontinuous.

Lemma 4.14. *Assume that S is a closed set. Suppose also that*

i. f is upper semicontinuous on S.

ii. g is lower semicontinuous on S.

Then the perturbation function p is upper semicontinuous at $\overline{y} = 0$.

Proof. Let $\varepsilon > 0$ be an arbitrary positive real number. By the definition of infimum there exists $\overline{x} \in G(0)$ such that

$$f(\overline{x}) \leq h(0) + \frac{\varepsilon}{2}. \tag{4.58}$$

Since f is upper semicontinuous, there exists a neighborhood $\mathscr{N}(\overline{x})$ such that

$$f(x) \leq f(\overline{x}) + \frac{\varepsilon}{2} \text{ for all } x \in \mathscr{N}(\overline{x}).$$

or equivalently,

$$f(x) - \frac{\varepsilon}{2} \leq f(\overline{x}) \text{ for all } x \in \mathscr{N}(\overline{x}). \tag{4.59}$$

Since g is lower semicontinuous on S. By Theorem 4.11 the set valued mapping G is continuous. Thus G is lower semicontinuous set-valued mapping. Since G is lower semicontinuous at $y = 0$, there exists a neighborhood $\mathscr{M}(0)$ of $y = 0$ such that

$$G(y) \cap \mathscr{N}(\overline{x}) \neq \emptyset \text{ for all } y \in \mathscr{M}(0).$$

For any $y \in \mathscr{M}(0)$, we know there exists $x \in G(y) \cap \mathscr{N}(\overline{x})$. Hence, $x \in G(y)$ and $x \in \mathscr{N}(\overline{x})$. By using (4.58) and (4.59) it follows that that

$$f(x) - \frac{\varepsilon}{2} \leq f(\overline{x}) \leq h(0) + \frac{\varepsilon}{2}.$$

Hence we obtain that

$$f(x) \leq h(0) + \varepsilon.$$

Taking the infimum when x ranges over $G(y)$ yields to

$$\inf_{x \in G(y)} f(x) \leq h(0) + \varepsilon.$$

Thus, we obtain that $h(y) \leq h(0) + \varepsilon$.
The proof is completed.

□

We obtain from Lemma 4.12 and Lemma 4.14 that the perturbation function $h(y)$ is continuous.

The following proposition is proved in [2]. The idea of the Lemma 4.12 is based on this proposition.

Proposition 4.15. *[see 2, Chapter 3, Proposition 21, page 119].*
Suppose that

i. W is upper semicontinuous on $X \times Y$.

ii. $G(y_0)$ is compact, and G is upper semicontinuous at y_0.

Then the marginal function V is upper semicontinuous at y_0.

The following approach was used from the proof of previous proposition to prove Lemma 4.12.

Remark 6. We know that

$$-V(y) = -\sup_{x \in G(y)} W(x,y) = \inf_{x \in G(y)} (-W(x,y)) = \inf_{x \in G(y)} f(x) = h(y).$$

Since $V(y)$ is upper semicontinuous from the previous proposition it follows that $h(y) = -V(y)$ is lower semicontinuous. The proof is completed.

It follows from Lemma 4.12 that the perturbation function $h(y)$ is lower semicontinuous.

Remark 7. The following proposition is proved in [2]. The idea of the proof of the following proposition is used in Lemma 4.14.

Proposition 4.16. *([2, Chapter 3, Proposition 19, page 118])*
Let X and Y be two sets, G is a set-valued map from Y to X, and W a real valued function defined on $X \times Y$. Consider the following the family of maximization problems

$$V(y) := \sup_{x \in G(y)} W(x, y)$$

that depends on the parameter y. The function V is called the marginal function. Suppose that

i. W is lower semicontinuous on $X \times Y$.

ii. G is lower semicontinuous at y_0.

Then the marginal function V is lower semicontinuous at y_0.

Remark 8. Our approach to the Lemma 4.14 as following. By the definition the function V, we have

$$V(y) = \sup_{x \in G(y)} W(x, y).$$

Since V is lower semicontinuous then it follows that

$$-V(y) = -\sup_{x \in G(y)} W(x, y) = \inf_{x \in G(y)} (-W(x, y)) = \inf_{x \in G(y)} f(x) = h(y).$$

then it implies that $V(y)$ is upper semicontinuous. Since f is continuous and G is continuous then

$$h(y) = \inf_{x \in G(y)} f(x)$$

is upper semicontinuous.

The following theorem shows that a positive homogeneous and continuous function is Lipschitz.

Lemma 4.17. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a positively homogeneous and continuous function. Then f is a Lipschitz function.*

Proof. We know that if f is continuous on $S_1 = \{u \in S : \|u\| = 1\}$ then it attains its minimum and maximum on S_1. Thus there exists real numbers m and M with

$$f(u) \geq m, \quad \text{for all } u \in S_1 \tag{4.60}$$

and

$$f(u) \leq M, \quad \text{for all } u \in S_1. \tag{4.61}$$

Let $x \in S$ be arbitrary, then there exists $t > 0, u \in S_1$ such that $x = tu$. Then it follows that

$$f(x) - f(0) = f(x) = f(tu) = tf(u) \geq tm = tm\|u\| = m\|tu\| = m\|x\| \tag{4.62}$$

Now, if $m > 0$, then (4.62) implies that

$$f(x) - f(0) \geq m\|x\| \geq -L\|x\| \text{ where } L > 0 \text{ arbitrary.} \tag{4.63}$$

If $m < 0$ in (4.62) then

$$f(x) - f(0) \geq m\|x\| = -L\|x\| \;\text{ where } L > 0 \text{ and } -L = m < 0. \tag{4.64}$$

Thus (4.63) and (4.64) imply that f is lower Lipschitz. Now since f is bounded above (4.61) implies that

$$f(x) - f(0) = f(x) = f(tu) = tf(u) \leq tM = tM\|u\| = M\|tu\| = m\|x\| \tag{4.65}$$

If $M > 0$, we know that

$$f(x) - f(0) \leq M\|x\| = L\|x\| \tag{4.66}$$

If $M < 0$, then

$$f(x) - f(0) \leq M\|x\| \leq L\|x\| \text{ where } L > 0 \text{ arbitrary.} \tag{4.67}$$

(4.66) and (4.67) implies that f is upper Lipschitz. Thus f is Lipschitz and there exists an $L > 0$ such that

$$|f(x) - f(0)| \leq L\|x - 0\|, \quad \forall x \in S. \tag{4.68}$$

□

The following theorem shows that a continuous and positive homogeneous function is weakly subdifferentiable.

Theorem 4.18. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a positively homogeneous and continuous function. Then f is weakly subdifferentiable.*

Proof. If f is homogeneous and continuous then by Lemma 4.17, f is Lipschitz. It follows from Theorem 2.33 that f is weakly subdifferentiable at $\bar{x} \in int(dom(f))$, that is $\partial^w f(\bar{x}) \neq \emptyset$ and there exists $(x^*, c) \in \partial^w f(\overline{x})$ with $x^* \neq 0$ and $c > 0$. The proof is completed. □

We can now state the main result of this subsection. We have shown that the perturbation function is positively homogeneous and Lipschitz. Then the strong duality easily follows from Lemma 4.17 and Theorem 4.3.

Theorem 4.19. *Let $S \subseteq \mathbb{R}^n$ be a cone and let $f : S \to \mathbb{R}$ and $g_i : S \to \mathbb{R}$,i=1,...,m be real valued functions. Assume that f and g_i, i=1,...,m are positively homogeneous and lower semicontinuous at* 0*. Then the strong duality relation holds for* (P)*.*

Proof. The Lemma 4.17 implies that the function h is Lipschitz. Then Theorem 2.33 implies that h is weakly subdifferentiable. Thus by Theorem 2.33 there exists a pair $(x^*, c) \in \partial^w f(\overline{x})$ with $x^* \neq 0$ and $c > 0$. Finally the strong duality relation holds for problem (P) easily follows from Theorem 4.3. □

We provide the following two examples to demonstrate Theorem 4.19.

Example 4.20. Consider the following nonconvex optimization problem

$$(P) \quad \min\Big\{f(x)\Big|g(x) = -x \leq 4\Big\}$$

where $f(x) = \begin{cases} 3x & \text{if } x \leq 0, \\ \dfrac{x}{2} & \text{if } x > 0. \end{cases}$

Notice that the objective and the constraint function are positively homogeneous. First we show that the optimal values of primal and dual problems are equal. It is easy to see that $Inf(P) = -12$. To formulate the dual problem $(P)^*$ we need to construct the corresponding set V. We have

$$\begin{aligned} V \quad &= \{(y^*, d) \in \mathbb{R} \times \mathbb{R}_+ | de - y^* \geq 0, \text{ for some } e \in \mathbb{R}, |e| \leq 1\}, \\ &= \{(y^*, d) \in \mathbb{R} \times \mathbb{R}_+ | y^* \leq d\}. \end{aligned}$$

Consequently, the dual problem $(P)^*$ is

$$(P)^* \qquad \sup_{(y^*,d)\in V} \inf_{x\in\mathbb{R}} \{f(x) - y^*(-x - 4) + \tau(x, y^*, d)\},$$

where

$$\tau(x, y^*, d) \quad = \sup\{de(-x-4) | |e| \leq 1, de - y^* \geq 0\}$$
$$= \begin{cases} -d(x+4) & \text{if } x \leq -4, \\ -y^*(x+4) & \text{if } x \leq -4, y^*+d \geq 0, \\ d(x+4) & \text{if } x \leq -4, y^*+d < 0, \end{cases}$$

for all $(y^*, d) \in V$ and $x \in \mathbb{R}$. The optimal value of $(P)^*$ is calculated as follows:

$$\sup_{(y^*,d)\in V} \min_{1\leq i\leq 3} \{\varphi_i(y^*, d)\}$$

where $\varphi_i(y^*, d) = \inf_{x\in X_i}$ and $X_1 = [-\infty, -4]$, $X_2 = [-4, 0]$ and $X_3 = [0, +\infty]$. Thus, $\varphi_1(y^*, d) = -12$,

$$\varphi_2(y^*, d) = \begin{cases} -12 & \text{if } y^*+d \geq 0, \\ -12 + 4(y^*+d) & \text{if } y^*+d < 0, \end{cases}$$

and

$$\varphi_3(y^*, d) = \begin{cases} 0 & \text{if } y^*+d \geq 0, \\ -\infty & \text{if } d = 0, y^*+d < 0. \end{cases}$$

Consequently,

$\sup_{(y^*,d)\in V} \min_{1\leq i\leq 3}\{\varphi_i(y^*, d) = 0\}$ and $Sup(P)^* = Inf(P) = -12$.

5. CONCLUSION

In this thesis, zero duality gap conditions are investigated in nonconvex optimization. We consider a nonlinear programming problem in the following form

$$(P) \quad \begin{cases} \inf_{x \in \mathbb{R}^n} f(x) \end{cases}$$

where

$$\{x \in S : g_j(x) \leq 0, j = 1, \ldots, m\} \neq \emptyset.$$

In this class of nonconvex optimization problem, we consider a positively homogeneous and lower semicontinuous objective and constraint functions defined on a conic set. We obtain the perturbation function is the weakly subdifferentiable at origin hence the zero duality gap property satisfies. These conditions are investigated based on the idea of the weak subdifferential.

In this thesis, we consider pointwise supremum of a family of $g_{(x_i^*, x_i, c_i, \alpha_i)}(x) = \langle x_i^*, x \rangle - c_i \|x - x_i\| + \alpha_i$ type of functions and we call these type of functions superlinear. Superlinear functions helps us investigate some properties of abstract convexity known to hold for the classical convex analysis. In this study, superlinear functions are used in order to provide the subdifferentiability theorems.

Additionally, we investigate the sum rule for the weak subdifferential. It is shown that the weak subdifferentials of some classes of the Clarke directionally differentiable functions and the tangentially convex functions, satisfy this property in the form of equality. Moreover we show that the "sup relation" for the weak subdifferential, can be formulated as a "max relation" for the mentioned classes of functions. In this work, we also study a relation between the weak subdifferential of the indicator function and the augmented normal cone to a nonconvex set, which is used to establish some additional properties of nonconvex sets.

REFERENCES

[1] Aubin, J.P. (1982). Contingent derivatives of set-valued maps and existence of solutions to nonlinear inclusions and differential inclusions, in Mathematical Analysis and Applications, Part A, edited by L. Nachbin, Advances in Mathematics: Supplementary Studies, 7A, Academic Press, New York, pp. 160–232.

[2] Aubin, J.P. and Ekeland, I.(1984). *Applied Nonlinear Analysis*, New York: Wiley Interscience in Dicrete Mathematics Series, Springer.

[3] Azimov, A.Y. and Gasimov, R.N. (1999). *On weak conjugacy, weak subdifferentials and duality with zero gap in nonconvex optimization*, International Journal of Applied Mathematics, 1, pp. 171–192.

[4] Azimov, A.Y. and Gasimov, R.N. (2002). *Stability and duality of nonconvex problems via augmented lagrangian*, Cybernatics and Systems Analysis, 3(3), pp. 412–421.

[5] Bagirov, A., Ozturk, G. and Kasimbeyli, R. (2019). *A sharp augmented Lagrangian-based method in constrained nonconvex optimization*, Optimization Methods and Software, 34(3), pp. 462–488.

[6] Bila, S. and Kasimbeyli, R. (2020). *On some Properties of the Weak Subdifferential*, Applied Analysis and Optimization, 4(2), pp. 247–255.

[7] Borwein,J.M. and Lewis, A.S. (2006). *Convex Analysis and Nonlinear Optimization*, CMS Books in Mathematics, Springer Science+Business Media, Inc., New York.

[8] Borwein, J.M. and Zhu, Q.J. (2010). *Techniques of Variational Analysis*, CMS Books in Mathematics, Springer Science+Business Media, Inc., New York.

[9] Browder, F.E. (1965). *Multivalued monotone nonlinear mappings and duality mappings in Banach spaces*, Trans. Amer. Math. Soc., 118, pp. 338–351.

[10] Bulbul, K.G. and Kasimbeyli, R.(2021). *Augmented Lagrangian Based Hybrid Subgradient Method for Solving Aircraft Maintenance Routing Problem*, Computers & Operations Research, published online 2021-03, https://doi.org/10.1016/j.cor.2021.105294

[11] Clarke, F.H. (1973). *Necessary conditions for nonsmooth problems in optimal control and the calculus of variations*, Thesis, University of Washington, Seattle.

[12] Clarke, F.H. (1975). *Generalized gradients and applications*, Trans. Amer. Math. Soc., 205, pp. 247–262.

[13] Clarke, F.H. (1983). *Optimization and nonsmooth analysis*, CMS of Monographs and Advanced Texts, John Wiley & Sons Inc., New York.

[14] Dinc Yalcin, G. and Kasimbeyli, R. (2020). *On weak conjugacy, augmented Lagrangians and duality in nonconvex optimization*, Mathematical Methods of Operations Research, 92(1), pp. 199–228.

[15] Dinc Yalcin, G. and Kasimbeyli, R. (2021). *Weak subgradient method for solving nonsmooth nonconvex optimization problems*, Optimization, 70(7), pp. 1513-1553.

[16] Dini, U. (1878). *Fondamenti per la Teoria della Funzioni di Variabili Reali*, Pisa, Italy.

[17] Ekeland, I. and Temam, R. (1976). *Convex Analysis and Variational Problems*, Elsevier.

[18] Gale,D., Kuhn,W. and Tucker,W. (1951). *Linear programming and the theory of games*, Activity Analysis of Production and Allocation, pp. 317–329.

[19] Gasimov, R.N. (1992). *Duality in nonconvex optimization*, Ph.D. Dissertation, Department of Operations Research and Mathematical Modeling, Baku State University, Baku.

[20] Gasimov, R.N. (2002). *Augmented Lagrangian duality and nondifferentiable optimization methods in nonconvex programming*, J. Global Optimization, 24, pp. 187–203.

[21] Ioffe, A.D. (1981). Nonsmooth analysis: differential calculus of non-differentiable mappings, Transactions of the American Methematical Society, 266, pp. 1–56.

[22] Minty, G.J. (1962). *Monotone (nonlinear) operators in Hilbert space*, Duke Math. J., 29, pp. 341–346.

[23] Moreau, J.J. (1966). *Fonctionelles convexes, lecture notes*, Seminaire "Equations aux Derivees Partielles", College de France.

[24] Kasimbeyli, N. (2015). *Existence and characterization theorems in nonconvex vector optimization*, J. Global Optim., 62, pp. 155–165.

[25] Kasimbeyli, N. (2019). *Separation theorem via superlinear functions and characterization of maximal elements in multiobjective optimization*, Applied Analysis and Optimization, 3(3), pp. 373–382.

[26] Kasimbeyli, N. and Kasimbeyli, R. (2017). *A representation theorem for Bishop-Phelps cones*. Pacific J. Optim., 13(1), pp. 55–74.

[27] Kasimbeyli, N., Kasimbeyli, R. and Mammadov, M. (2016). *A Generalization of a Theorem of Arrow, Barankin and Blackwell to a Nonconvex Case*, Optimization, 65(5), pp. 937–945.

[28] KASIMBEYLI, R. (2009). *Radial Epiderivatives and Set-Valued Optimization*, Optimization, 58, pp. 521–534.

[29] KASIMBEYLI, R. (2010). *A Nonlinear Cone Separation Theorem and Scalarization in Nonconvex Vector Optimization*, SIAM J. on Optimization, 20, pp. 1591–1619.

[30] KASIMBEYLI, R. (2013). *A conic scalarization method in multi-objective optimization*, Journal of Global Optimization, 56(2), pp. 279–297.

[31] KASIMBEYLI, R. AND KARIMI, M. (2019). *Separation theorems for nonconvex sets and application in optimization*, Operations Research Letters, 47, pp. 569–573.

[32] KASIMBEYLI, R. AND KARIMI, M. (2021). *Duality in Nonconvex Vector Optimization*, Journal of Global Optimization, 80(1), pp. 139–160.

[33] KASIMBEYLI, R. AND MAMMADOV, M. (2009). *On weak subdiferentials, directional derivatives, and radial epiderivatives for nonconvex functions*, SIAM J. Optimization, 20, pp. 841–855.

[34] KASIMBEYLI, R. AND MAMMADOV, M. (2011). *Optimality conditions in nonconvex optimization via weak subdifferentials*, Nonlinear Analysis: Theory, Methods and Applications, 74, pp. 2534–2547.

[35] LEMARECHAL, C. (1986). *An introduction to the theory of nonsmooth optimization*, Optimization, 17(6), pp. 827–858.

[36] MARTINEZ LEGAZ, J.E. (2014). *Optimality conditions for pseudoconvex minimization over convex sets defined by tangentially convex constraints*, Optimization Letters, 9(5), pp. 1017–1023.

[37] MICHEL, P. AND PENOT, J.P. (1984). Calcul sous-differentiel pour les fonctions lipschitziennes et non lipschitziennes, Comptes Rendus de l'Academie des Sciences de Paris, 298, pp. 269–272.

[38] MORDUKHOVICH, B.S. (1976). Maximum principle in the problem of time optimal response with nonsmooth constraints, Journal of Applied Mathematics and Mechanics, 40, pp. 960–969.

[39] MORDUKHOVICH, B.S. (1984). Nonsmooth analysis with nonconvex generalized differentials and conjugate mappings, Dokl. Acad. Nauk BSSR (Russian), 28, pp. 976–979.

[40] PENOT, J.P. (1974). Sous-differentielles de fonctions numeriques non-convexes, Comptes Rendus de l'Academie des Sciences de Paris, 278, pp. 1553–1555.

[41] PSHENICHNYI, B.N. (1971). *Necessary conditions for an extremum*, Nauka, Moscow, (Russian).

[42] Pshenichnyi, B.N. (1971). *Necessary conditions for an extremum*, Marcel Dekker, Inc., New York.

[43] Robinson, S.M. (1987). *Local structure of feasible sets in nonlinear programming, part III: stability and sensitivity*, Mathematical Programming Studies, 30, pp. 45–66.

[44] Rockafellar, R.T. (1963). *Convex functions and dual extremum problems*, Thesis, Harvard.

[45] Rockafellar, R.T. (1970). *Convex analysis*, Princeton University Press, Princeton.

[46] Rockafellar, R.T. (1974). *Conjugate duality and optimization*, Regional Conference Series in Applied Math. 16, SIAM Publications, Philadelphia.

[47] Rockafellar, R.T. (1979). *Clarke's tangent cones and the boundaries of closed sets in R^n*, Nonlinear Analysis: Theory, Methods and Applications, 3, pp. 145–154.

[48] Rockafellar, R.T. (1980). *Generalized directional derivatives and subgradients of nonconvex functions*, Can. J. Math., 32(2), pp. 257–280.

[49] Rockafellar, R.T. and Wets, R. J-B. (2009). *Variational analysis*, Springer-Verlag Berlin Heidelberg.

[50] Rubinov, A.M. (2000). Abstract Convexity and Global Optimization, Kluwer, Dordrecht, Boston.

[51] Rubinov, A.M. and Gasimov, R.N. (2003). *Strictly increasing positively homogeneous functions with application to exact penalization*, Optimization, 52(1), pp. 1–28.

[52] Ustun, O. and Kasimbeyli, R. (2012). *Combined forecasts in portfolio optimization: A generalized approach*, Computers & Operations Research, 39, pp. 805–819.

Samet Bila

CURRICULUM VITAE

Educational Background

- 2016-2022, PhD, Eskişehir Technical University, Graduate School of Sciences, Mathematics Department
- 2012-2015, M.Sc., Clemson University, Mathematical and Statistical Sciences, Mathematics Department
- 2010-2012, M.Sc., Indiana University Purdue University at Indianapolis, School of Science, Mathematics Department
- 2004-2008, BS, Mimar Sinan Fine Arts University, Science and Letters Faculty, Mathematics Department

Work Experience

- 2015-current, Research Assistant, Eskişehir Technical University, Faculty of Science, Mathematics Department
- 2012-2015, Teaching Assistant, Clemson University, Mathematical and Statistical Sciences, Mathematics Department

Publications

- Working Paper: Kasimbeyli, R. and Bila, S., Zero Duality Gap Conditions for classes of Optimization Problems.
- Submitted paper: Bila, S., Kasimbeyli, R., and Farajzadeh, A. (2021) On Sum Rule for the Weak Subdifferential and Some Properties of Augmented Dual Cones, Journal of Global Optimization, Under Editorial Review.
- Journal Paper: Bila, S. and Kasimbeyli, R., On Some Properties of the Weak Subdifferential (2020) Applied Analysis and Optimization.

Rewards

- 2009-2012, YLYS, A Government Scholarship to Study Abroad for Higher Education.