Geometry in Structured Optimisation Problems

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

Tian SANG
Master of Science, The University of Melbourne
Bachelor of Science, The University of Melbourne

School of Science
College of Science, Engineering and Health
RMIT University

October, 2018
Declaration

I certify that except where due acknowledgement has been made, the work is that of the author alone; the work has not been submitted previously, in whole or in part, to qualify for any other academic award; the content of the thesis is the result of work which has been carried out since the official commencement date of the approved research program; any editorial work, paid or unpaid, carried out by a third party is acknowledged; and, ethics procedures and guidelines have been followed.

Tian Sang

October 31st, 2018
Acknowledgements

I’d like to acknowledge the support I have received for my research through the provision of an RMIT DECRA Scholarship tied to the ARC DECRA grant from Dr.Vera Roshchina. The DECRA grant number is DE150100240 with the title “Geometry and Conditioning in Structured Conic Problems”.

Thanks to my supervisors Dr.Vera Roshchina and Prof. Andrew Eberhard. Vera and Andrew have supported my research by providing funding for me to go to conferences during my PhD candidature, including my collaborations in Spain and Sweden during Heidelberg Laureate Forum.

Thanks to Australian Mathematical Sciences Institute (AMSI) and Australian Mathematical Society (AustMS) for their support on attending various research activities.

Thanks to Dr.Yousong Luo, Prof. Marc Demange, Dr. Haydar Demirhan, Prof. Lewi Stone, and Prof. Asha Rao, who have been supportive to me at RMIT University.

Special thanks to some of my dearest friends and family (including my father who passed away just before my PhD), whose names and stories I had written down and really want to share here, but not able to due to the length of this section. Thanks for everything these dearest ones have done to support me in my life. They are the reason this PhD thesis exists today.
Dedication

I would like to in particular dedicate this PhD thesis to my beloved mum Daijie Tian, and a few of my dearest friends, Greg Sully, John Arcaro, and Bell Foozwell. Sometimes I may not be able to show or good at expressing myself, but I love you all, without defining whatever the word “love” really means, or however it takes its form. You all have my love and my heart, always have, always will.
## Contents

Declaration ii  
Acknowledgements iii  
Dedication iv  
Abstract vii  
Publications viii  

1 Introduction 1  
1.1 Motivation 4  

2 Background 8  
2.1 Convex Geometry 8  
2.1.1 General Convex Sets and Polytopes 8  
2.1.2 Facial Structure 10  
2.1.3 Cones 11  
2.2 Basics on Subdifferentials 12  
2.2.1 Lipschitz Continuity 14  
2.2.2 Subdifferentials for Convex Functions 14  
2.2.3 Regular and Limiting Subdifferentials 24  

3 Demyanov-Ryabova Conjecture 28  
3.1 Background 29
CONTENTS

3.2 Preliminaries ................................................................. 31
3.3 Affinely Independent Case ............................................. 34
3.4 The Simplified Demyanov Convertor ................................. 37
3.5 Algebraic Reformulation and the Main Result ...................... 43
3.6 General convex sets ....................................................... 48
3.7 The counterexample and future work ................................. 50
3.8 Conclusion .................................................................. 51

4 Outer Limits of Subdifferentials ........................................... 52
  4.1 Introduction ................................................................. 52
  4.2 Preliminaries ............................................................... 53
  4.3 Limiting subdifferential for pointwise minima ..................... 59
  4.4 Exact representations for piecewise affine functions .......... 70
  4.5 Conclusion .................................................................. 75

5 Facial Structure for Convex Sets ......................................... 76
  5.1 Introduction ................................................................. 76
  5.2 Main Result ................................................................. 79
  5.3 Fractal convex sets ......................................................... 82
  5.4 Future work ................................................................. 85
  5.5 Conclusion .................................................................. 85

6 Conclusion ..................................................................... 87
Abstract

In this thesis, we start by providing some background knowledge on importance of convex analysis. Then, we will be looking at the Demyanov-Ryabova conjecture. This conjecture claims that if we convert between finite families of upper and lower exhausters with the given convertor function, the process will reach a cycle of length at most two. We will show that the conjecture is true in the affinely independent special case, and also provide an equivalent algebraic reformulation of the conjecture.

After that, we will generalise the outer subdifferential construction for max type functions to pointwise minima of regular Lipschitz functions. We will also answer an open question about the relation between the outer subdifferential of the support of a regular function and the end set of its subdifferential.

Lastly, we will address the question of what kind of dimensional patterns are possible for the faces of general closed convex sets. We show that for any finite increasing sequence of positive integers, there exist convex compact sets which only have faces with dimensions from this prescribed sequence. We will also discuss another approach to dimensionality by considering the dimension of the union of all faces of the same dimension. We will demonstrate that the problems arising from this approach are highly nontrivial by providing some examples of convex sets where the sets of extreme points have fractal dimensions.
Publications


Chapter 1

Introduction

In this chapter, we will give an outline of the structure of this thesis, including the main focus on each project worked on, as well as some background knowledge about these projects. We will start by introducing a few simple examples to demonstrate the ideas underlying our work.

In Chapter 2, we will introduce some background knowledge for the later chapters. We start with fundamentals in convex geometry with definitions and examples of general convex sets, polytopes, faces and cones, as well as a few standard results in convex geometry. Then, we will introduce the basics on subdifferentials, including the concepts of directional derivative, support function, Lipschitz continuity, and outer limits. These concepts will appear in the work in Chapter 4. We will also give some results related to these concepts.

In Chapter 3, our work focuses on the Demyanov-Ryabova conjecture. This conjecture is about converting finite exhausters, which are multiset objects that generalise the subdifferential of a convex function. This work is motivated by the calculus of generalised subdifferentials. The study on subdifferentials is crucial in nonsmooth analysis. Exhausters are useful for formulating optimality conditions and finding directions of steepest ascent and descent for wide range of nonsmooth functions. Therefore, it is a very powerful tool for finding the minima and maxima of functions. There are two types of exhausters, upper exhausters and lower exhausters, and there is a one to one correspondence between them.
We would like to relate these two sets to each other, and the Demyanov-Ryabova conjecture
provides the insight into upper and lower exhauster representations.

The Demyanov-Ryabova conjecture claims that when we convert between finite families
of upper and lower exhausters with given convertor function, the process will reach a cycle
of length at most two. We are going to show in Chapter 3 that the conjecture is true in the
affinely independent setup. Then, we will also construct an equivalent combinatorial version
of the conjecture.

The conjecture was first published in 2011 (see [18]), and is well known in the constructive
nonsmooth analysis community. Another published work done in this direction is [13], in
which a special case of the problem was resolved. It was shown that when the collection of
sets includes all subsets of the minimal cardinality, the conjecture is true. This condition
is different from the affine independence assumption that we are going to show in Chapter
3. However, recently a counterexample was found in [51] which disproved the conjecture.
The counterexample was constructed in \( \mathbb{R}^2 \), and it has the minimal cycle of length 4. The
existence of this counter example and the existence of two different sets of conditions under
which a two cycle does exist means some future work is needed to find the minimal set of
conditions for the existence of a two cycle.

In Chapter 4, we study outer limits of subdifferentials. This chapter is motivated by the
importance of error bounds. The concept of error bound is important in optimisation theory,
in particular, it is strongly related to sensitivity analysis that is a key notion in optimisation
complexity and numerical analysis. Sensitivity analysis studies the uncertainty of the data of
a mathematical model, and how this uncertainty potentially impacts the results. To be able
to compute error bounds, we will need outer limits of Fréchet subdifferentials. In particular,
we know that for continuous functions, the error bound is bounded below by the Euclidean
distance from 0 to the outer limit of Fréchet subdifferentials. The focus of this work is to
find good ways to evaluate error bounds via outer limits.

One of our main references is by Cánovas, Henrion, López, and Parra in [11]. The work by
Cánovas, Henrion, López, and Parra on outer limits of subdifferentials and calmness moduli
concerns with max type functions, which are the maximum of a finite set of continuously
differentiable functions of \( n \) real variables. The authors have obtained some results for
deriving an upper bound for the calmness modulus of nonlinear systems. Also, if we consider the convex case, a lower bound also can be obtained. Apart from that, the computation on calmness modulus of linear programming problems is restricted to optimal sets that are singletons. However, the authors were able to come up with a point-based formula for the calmness modulus using Karush-Kuhn-Tucker index set method, which is given in terms of the calmness moduli of certain sublevel multifunctions. It has no assumption on optimal set. Outer limit of subdifferentials is the main tool for computing the calmness modulus for a $C^1$-system at some feasible point. Firstly, the authors approached this problem by considering the special case of polyhedral functions, after that, the more general continuously differentiable functions were being analysed.

We are going to generalise the outer subdifferential for max type functions to pointwise minima of regular Lipschitz functions. We will also answer an open question from Li, Meng, and Yang’s paper about the relation between the outer subdifferential of the support of a regular function.

In Chapter 5, we present some results about the facial structure of convex sets. A significant body of research has emerged recently which digs deeper into the structure of convex sets, particularly focussing on exploring the facial structure (for example, see [3, 32, 46, 52, 56]). The classic convex analysis books give only a cursory overview of the facial structure of convex sets (see [27, 47, 48]). This kind of analysis is crucial in understanding the interplay between geometry and numerical performance, because it highlights the difference between our three-dimensional intuition and the complexity of higher-dimensional geometry, which becomes ever more prominent in large-scale problems. One of the properties we care about is the presence of faces of different dimensions. The facial structure of traditional polyhedral sets is well studied. For example, we know that given a polyhedral set, the faces with every dimensionality are presented in the set, that is, there are no “gaps” when we look at dimensions of faces in the set. However, this is not true for general convex sets. In particular, we will show some results on dimensional patterns for the faces of general closed convex sets. We will also approach the dimensionality by considering the dimension of the union of all faces of the same dimension. One of the interesting open questions to consider for the future is whether any dimensional pattern can be represented by a spectrahedron.
CHAPTER 1. INTRODUCTION

1.1 Motivation

For a smooth function, we often analyse it by studying its gradient. By using its gradient as a tool, we can describe its optimality conditions, obtain its maximum and minimum subject to certain constraints, which motivated the development of differential calculus.

If the function is convex, a useful tool as a substitute for the nonexisting derivative is a subgradient. The set of all subgradients of the function at the point is called the subdifferential.

For our work in this thesis, Demyanov-Ryabova Conjecture is motivated from the study of optimality condition for nonsmooth functions, and exhausters are employed to describe optimality conditions. Our work on Outer Limits of Subdifferentials is strongly related to calculus of subdifferentials and construction of error bounds. The notion of error bound plays an important role in analysis, as it measures whether a given function is steep enough outside of its level set and gives a lower bound for the relevant slope, which is crucial for the convergence of numerical methods.

Example 1.1.1 Consider a function that can be represented as a maximum over a family of linear functions. As explained later in Chapter 2 the subdifferential of this function at zero is the convex hull of all gradients of these linear component functions. If this family of linear functions is finite, then this subdifferential is a convex polytope.

At an arbitrary point (not necessarily zero) the subdifferential of this max-function is a face of the subdifferential that is maximal in the direction given by this point (i.e. it maximises the inner product with this point over the entire convex hull/polytope).

Let \( f(x, y) = \max\{2x+4y, 3x+2y, 6x+3y, 5x+7y, 8x+5y, \frac{7}{2}x+\frac{36}{5}y\} = \max\{f_1, f_2, f_3, f_4, f_5, f_6\} \).

Then the gradients of each function correspond to the vertices of the polygon.
The theory of convex sets and functions as well as optimality conditions in terms of subdifferential has been developed very well in literature. The classic calculation of derivatives often assuming the functions to be differentiable, the subdifferential calculus allows us to proceed calculations even if the function is nonsmooth. Subdifferential calculus can also characterise the Lipschitz continuity of the function.

In Chapter 3, the Demyanov-Ryabova conjecture is about using the convex hull of support faces to form a chain of collections of sets, and then apply the transformation to it, which is the Demyanov convertor.

**Example 1.1.2** For the example shown in Figure 1.2, the three functions have the following expression:

\[
f_1 = (x - 2)^2, \quad f_2 = \frac{1}{2}(x - 5)^2 + 1, \quad f_3 = (x - 4)^2 + 3.
\]

Now, consider

\[
f = \max\{f_1, f_2, f_3\}.
\]

It is clear from the Figure that the max function is obtained from \(f_1\) and \(f_3\). This means the active index set (defined in Definition 2.2.13), denoted as \(I(x)\) is \(I(x) = \{1, 3\}\) at the point
$x$ where $f_1$ and $f_3$ intersect.

![Figure 1.2: Active index set](image)

The active index set is important in calculus for computing subdifferentials of maxfunctions. For example, in the Steepest-Descent Method [26], we are required to obtain all the active indices at the given point, compute the corresponding gradients, and then the subdifferential is the set of all convex combination of the gradients.

**Example 1.1.3** In the Figure 1.3 below, the dotted lines represent the hyperplanes that contain support faces for the given convex set. Then the outer limit is formed by considering all subsets of the active index set, if the inner product of the point and the directional vector equals to 1, then point is in the outer limit.

Geometrically, this means the outer limit is obtained by taking the union of all support faces in convex sets which are on the other side of the “0”.
The motivation of this example is to illustrate the computational process of error bounds (See Section 4.1), which requires outer limits of subdifferentials. In particular, the error bound of a continuous function is bounded below by the Euclidean distance from 0 to the outer limit of subdifferentials.

In Chapter 4, we apply similar method with the active index set and support faces to construct the error bound. The following example gives the idea on the computation process of the error bound, which will be explained more in details in Chapter 4.

**Example 1.1.4** Let $f : \mathbb{R} \to \mathbb{R}$ be defined as:

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

Then, differentiate the function $f$, we have,

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

*The error bound of $f(x)$ at $x = 0$ is 1 in this case.*
Chapter 2

Background

2.1 Convex Geometry

2.1.1 General Convex Sets and Polytopes

A convex set is a set with the property that for every pair of points in the set, the line segment connecting these two points is also contained in the set.

Definition 2.1.1 (Convex set) A set $C \subseteq \mathbb{R}^n$ is called convex if for any pair of points $x, y \in C$ and any $\alpha \in [0, 1]$, we have,

$$\alpha x + (1 - \alpha)y \in C, \quad \alpha \in [0, 1].$$

Example 2.1.2 In Figure 2.1 are examples of some convex sets and nonconvex sets.

![Figure 2.1: On the top: some convex sets; on the bottom: some nonconvex sets.](image-url)
Let the line segment which connects the pair of points \( x, y \in \mathbb{R}^n \) be denoted by \([x, y]\).

That is,

\[
[x, y] := \{ \alpha x + (1 - \alpha)y \mid \alpha \in [0, 1] \}.
\]

We can generalise the line segment connecting any pair of points in the Euclidean space by *convex combination*.

**Definition 2.1.3 (Convex combination)** We say \( x \in \mathbb{R}^n \) is a convex combination of \( \{x_1, x_2, ..., x_k\} \subseteq \mathbb{R}^n \) if there exist \( \alpha_1, ..., \alpha_k \) such that:

\[
x = \sum_{i=1}^{k} \alpha_i x_i, \quad \alpha_i \geq 0, \forall i \in \{1, ..., k\}, \text{ and } k \sum_{i=1}^{k} \alpha_i = 1.
\]

**Definition 2.1.4 (Convex hull)** Let \( S \) be a finite set. We define the convex hull of \( S \) to be the set of all convex combinations of points in \( S \). i.e.,

\[
\text{conv}(S) = \left\{ x = \sum_{i=1}^{k} \alpha_i x_i \mid x_i \in S, \alpha_i \geq 0 \forall i = 1, ..., k, \sum_{i=1}^{k} \alpha_i = 1 \right\}.
\]

**Example 2.1.5** For the example shown in Figure 2.2, \( C \) is the convex hull of four sets.

![Figure 2.2: Convex hull](image)

In other words, the convex hull of \( S \) is the smallest convex set that \( S \) is contained in. To see this, we need the following theorem.
Theorem 2.1.6 Let \( S \subseteq \mathbb{R}^n \), then \( \text{conv}(S) \) is the intersection of all convex sets in \( \mathbb{R}^n \) which contain \( S \).

Proof See [10] for the proof.

Note that, an intersection of a family of convex sets is convex.

Lemma 2.1.7 Let \( I \) be an index set. Let \( C_i, i \in I \) be a collection of convex sets in \( \mathbb{R}^n \). Then, the intersection of \( C_i \)'s,

\[
C = \bigcap_{i \in I} C_i
\]

is also convex.

Proof See any standard text on convex analysis such as [48] or [27].

Definition 2.1.8 (Minkowski sum) Let \( A, B \subseteq \mathbb{R}^n \) be two arbitrary sets. The Minkowski sum of sets \( A \) and \( B \) is formed by adding every vector from \( A \) to every vector in \( B \). That is,

\[
A + B := \{ a + b \mid a \in A, \ b \in B \}.
\]

Example 2.1.9 Below is an example of the Minkowski sum of a disk and a square.

2.1.2 Facial Structure

For polytopes, we have the notion of vertices, edges, and facets. For convex sets, we can generalise these notions to faces. Intuitively, the study of facial structure involves analysing properties of faces and how these faces are joined together.

Definition 2.1.10 (Face) Let \( C \subseteq \mathbb{R}^n \) be a convex set. Then a subset \( F \subseteq C \) is called a face of \( C \) if:
• $F$ is convex

• Let $[x, y] \subseteq C$ be a line segment. If $(x, y) \cap F \neq \emptyset$, then $[x, y] \subseteq F$.

**Example 2.1.11** In Figure 2.3 are some examples of faces.

---

**2.1.3 Cones**

**Definition 2.1.12 (Cone)** A set $K \subseteq \mathbb{R}^n$ is called a cone if for any point $x \in K$ and $\lambda > 0$, we have $\lambda x \in K$.

**Definition 2.1.13 (Polar cone)** Let $K$ be a cone in $\mathbb{R}^n$. Then the polar cone of $K$ is defined to be:

$$K^o = \{ y \in \mathbb{R}^n \mid \langle y, x \rangle \leq 0, x \in K \}.$$

**Example 2.1.14** On the left in Figure 2.4 we have a set $S$ and the smallest cone $K$ that contains $S$, then $K^o$ is the polar cone for $S$.

On the right in Figure 2.4 is the diagram that illustrates the polar cone $K^o$ to the original cone $K$ in the 3 dimensional space.
Corollary 2.1.15 Let $\mathcal{K}$ be a closed convex cone. Then,

$$\forall s \in \mathcal{K}$$

Proof See [9] and [47] for the proof.

2.2 Basics on Subdifferentials

We will start this section by defining the max-function, which we will be needing this in the Chapter 4.

Definition 2.2.1 (Max-function) Let $I$ be a finite index set and $i \in I$. Let $g_i : \mathbb{R}^n \to \mathbb{R}$ be continuously differentiable functions. Then, the max-function $g(x)$ is defined to be:

$$g(x) = \max_{i \in I} g_i(x), \ x \in \mathbb{R}^n.$$  

In Chapter 4, we will study outer limits of subdifferentials, which is motivated by the problem of constructive evaluation of error bounds. To be able to formally define error bound, we need first to define the notion of a sublevel set of a function.
**Definition 2.2.2 (Sublevel set)** Let \( f : X \rightarrow \mathbb{R} \) be a continuous function defined on the open set \( X \subseteq \mathbb{R}^n \). Then, the **sublevel set** of \( f \) corresponding to \( \bar{x} \in X \) is:

\[
S(\bar{x}) := \{ x \in X \mid f(x) \leq f(\bar{x}) \},
\]

where \( \bar{x} \in X \).

Suppose \( f \) is the function as defined above, and \( S(\bar{x}) \) is the sublevel set. Then, we can define the **local (linear) error bound**.

**Definition 2.2.3 (Local (linear) error bound)** The function \( f \) has a local (linear) error bound at \( \bar{x} \) if there exists a constant \( L > 0 \) such that

\[
L \text{ dist}(x, S(\bar{x})) \leq \max\{0, f(x) - f(\bar{x})\}
\]

for all points \( x \) in a sufficiently small neighbourhood of \( \bar{x} \).

The error bound modulus measures whether a given function is steep enough outside of its level set, and it gives a lower bound for the relevant slope. In our work in Chapter 4, we focus on the construction of error bound moduli of structured continuous functions. Now we can formally define **error bound modulus**.

**Definition 2.2.4 (Error bound modulus)** Let \( f : X \rightarrow \mathbb{R} \) be a continuous function which has a local (linear) error bound at \( x \). Let \( L \) to be the constant that satisfies the condition in Definition 2.2.3. Then, the error bound modulus of \( f \) at \( \bar{x} \) is defined to be the supremum of the constants \( L \) over all neighbourhoods of \( \bar{x} \). It can be expressed explicitly as

\[
Erf(\bar{x}) := \liminf_{x \to \bar{x}, f(x) > f(\bar{x})} \frac{f(x) - f(\bar{x})}{\text{dist}(x, S(\bar{x}))}.
\]
2.2.1 Lipschitz Continuity

**Definition 2.2.5 (Lipschitz continuity)** Let $f : \mathbb{R}^n \to \mathbb{R}$ be a real-valued function. Then we say $f$ is Lipschitz continuous if for all $x, y \in \text{dom}(f)$, there exists $c \in \mathbb{R}$ such that,

$$|f(x) - f(y)| \leq c||x - y||,$$

where $c$ is called Lipschitz constant.

In general, we know that a differentiable function with bounded derivative must be Lipschitz, and any finite valued convex function is also Lipschitz on closed and bounded sets. (Theorem 3.1.2 [27])

2.2.2 Subdifferentials for Convex Functions

The concept of “subdifferentiaitonal” is a generalisation of the classical differential. Before we give the formal definition of various subdifferentials, we will go back to some definitions from ordinary differential calculus and convex analysis first.

Intuitively, a convex function means for every line segment that connects two points of the function graph of the function, the line segment lies above or on the graph.

**Definition 2.2.6 (Convex function)** Let $f : \mathbb{R}^n \to \mathbb{R}$ be a continuous real-valued function. Then, $f$ is convex if for all $x, y \in \mathbb{R}^n$ and $t \in [0, 1]$:

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y).$$

We will assume the function $f$ is convex for the rest of this section unless specified otherwise.

**Definition 2.2.7 (Directional derivative)** Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function. Let $d \in \mathbb{R}^n$ be a direction. Then the directional derivative of $f$ at $x$ is defined to be:

$$f'(x, d) := \lim_{t \to 0^+} \frac{f(x + td) - f(x)}{t} = \inf_{t>0} \frac{f(x + td) - f(x)}{t}.$$
Note that a reference for the second equality can be found in [47] Theorem 2.4.1.

**Definition 2.2.8 (Sublinear function)** A function \( \sigma : \mathbb{R}^n \to \mathbb{R} \) is called sublinear if it satisfies the following two conditions:

- **Positive Homogeneity:** \( \sigma(tx) = t\sigma(x) \) for all \( x \in \mathbb{R}^n \) and \( t > 0 \).
- **Subadditivity:** \( \sigma(x + y) \leq \sigma(x) + \sigma(y) \) for all \( x, y \in \mathbb{R}^n \).

**Proposition 2.2.9** Let \( x \in \mathbb{R}^n \) be fixed and \( f \) a convex function. Then the directional derivative of \( f \) at \( x \) is sublinear.

**Proof** We can check two sublinearity conditions for the directional derivative of \( f \) at \( x \).

- **Positive Homogeneity:** Let \( t_1 > 0 \) and \( d \in \mathbb{R}^n \), then we have

\[
f'(x, t_1 d) = \lim_{t \to t_1^+} \frac{f(x + td) - f(x)}{t} \\
= t_1 \lim_{t_2 \to t_1^+} \frac{f(x + t_2 d) - f(x)}{t_2} \\
= t_1 f'(x, d)
\]

where \( t_2 = t_1 t \)

- **Subadditivity:** Let \( d_1, d_2 \in \mathbb{R}^n \), and let \( \alpha_1 + \alpha_2 = 1 \). Apply the convexity of \( f \), then we have,

\[
f'(x, (\alpha_1 d_1 + \alpha_2 d_2)) = \lim_{t \to 0^+} \frac{f(x + t(\alpha_1 d_1 + \alpha_2 d_2)) - f(x)}{t} \\
= \lim_{t \to 0^+} \frac{f(\alpha_1(x + td_1) + \alpha_2(x + td_2)) - \alpha_1 f(x) - \alpha_2 f(x)}{t} \\
\leq \alpha_1 \lim_{t \to 0^+} \frac{f(x + td_1) - f(x)}{t} + \alpha_2 \lim_{t \to 0^+} \frac{f(x + td_2) - f(x)}{t} \\
= \alpha_1 f'(x, d_1) + \alpha_2 f'(x, d_2) \\
= f'(x, \alpha_1 d_1) + f'(x, \alpha_2 d_2)
\]
Therefore, the directional derivative of a convex function is sublinear.

Intuitively, the directional derivative represents the rate of change of the function at a given point along a vector in the space. We know from Proposition 2.2.9 that the directional derivative of a convex function is always sublinear.

The concept of subdifferential generalises the gradient to functions that are not differentiable everywhere, giving rise to nonsmooth analysis. Intuitively, the subdifferential captures the normals to all supporting hyperplanes of the epigraph, which are called subderivatives. In other words, the set of all subderivatives is called the subdifferential, therefore, the subdifferential is a set-valued function or mapping.

For convex functions, there are a few different ways to define a subdifferential. One of them is to compute the directional derivative first, and then relate it to the set which the directional derivative supports, so we will introduce the support function and Minkowski duality.

**Definition 2.2.10 (Support function)** Let $C$ be a convex set, and let $v \in \mathbb{R}^n$. Then we define the support function of $C$ to be:

$$S_C(v) := \sup\{\langle x, v \rangle \mid x \in C\}.$$ 

Then, Minkowski duality ensures the one-to-one correspondence between a compact convex set and its support function. In other words, the map between the compact convex set and its support function is a bijection.

We note that $S_C(v)$ is finite valued if and only if $C$ is compact.

Now we have the following definition for subdifferential.

**Definition 2.2.11 (Subdifferential)** The subdifferential of $f$ at $x$ is defined to be the set of vectors $w \in \mathbb{R}^n$ that satisfies:

$$\partial f(x) = \{w \in \mathbb{R}^n \mid f(u) \geq f(x) + \langle w, u - x \rangle, \forall u \in \mathbb{R}^n\}.$$
Alternatively, we have a description of the subdifferential via directional derivative, as follows.

**Proposition 2.2.12 (Subdifferential via directional derivative)** Let \( f \) be a finite-valued locally Lipschitz convex function. Then, the subdifferential \( \partial f(x) \) of \( f \) at \( x \) is the nonempty compact convex set in \( \mathbb{R}^n \) which has \( f'(x, \cdot) \) as its support function. That is,

\[
\partial f(x) := \{ a \in \mathbb{R}^n \mid \langle a, d \rangle \leq f'(x, d) \forall d \in \mathbb{R}^n \}.
\]  

(2.1)

**Proof** Let \( w \in \mathbb{R}^n \) be a member of (2.1). Then we have,

\[
\langle w, d \rangle \leq f'(x, d), \forall d \in \mathbb{R}^n.
\]

Now apply the second half of the result on the directional derivative (Proposition 2.2.7) of \( f \) at \( x \), we now have,

\[
\langle w, d \rangle \leq \inf_{t>0} \frac{f(x + td) - f(x)}{t}, \forall d \in \mathbb{R}^n.
\]  

(2.2)

Let \( u := x + td \). Since the directional derivative is defined for all \( d \in \mathbb{R}^n \), and \( t \in \mathbb{R}^+ \), we know that \( u \) ranges over whole of \( \mathbb{R}^n \). Then (2.2) implies,

\[
\langle w, u - x \rangle = \langle w, td \rangle \leq f(x + td) - f(x) = f(u) - f(x), \forall u \in \mathbb{R}^n
\]

implying

\[
w \in \partial f(x) \text{ (by Definition 2.2.11)}
\]

Considering Definition 2.2.11, we have the following:

\[
f(u) \geq f(x) + \langle w, u - x \rangle, \forall u \in \mathbb{R}^n
\]

\[
\Rightarrow f(u) - f(x) \geq \langle w, u - x \rangle, \text{ for } u = x + td, d \in \mathbb{R}^n, t > 0,
\]

\[
\Rightarrow f(x + td) - f(x) \geq \langle w, x + td \rangle - \langle w, x \rangle
\]

\[
\Rightarrow f(x + td) - f(x) \geq \langle w, x \rangle + t \langle w, d \rangle - \langle w, x \rangle
\]

\[
\Rightarrow \frac{f(x + td) - f(x)}{t} \geq \langle w, d \rangle, \forall d \in \mathbb{R}^n, t > 0
\]

\[
\Rightarrow w \in \{ w \mid f'(x, d) \geq \langle w, d \rangle, \forall d \in \mathbb{R}^n \}
\]

Therefore, the two characterisations of the subdifferentials are equivalent. \( \square \)
When $f$ is locally Lipschitz $f'(x, \cdot)$ is finite valued so $\partial f(x)$ is compact.

Later we would like to generalise subdifferential to non-convex functions.

Let $I$ be a finite index set, and $\{f_i\}_{i \in I}$ is a collection of convex functions with $f : \mathbb{R}^n \to \mathbb{R}$.

Suppose that

$$f(x) := \max \{ f_i(x) \mid i \in I \} < \infty, \ \forall \ x \in \mathbb{R}^n. \quad (2.2.1)$$

**Definition 2.2.13 (Active index-set)** Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function defined in 2.2.1 and let $x \in \mathbb{R}^n$. Then we can define the active index-set as below:

$$I(x) := \{ i \in I \mid f_i(x) = f(x) \}.$$

**Example 2.2.14 (Active index-set)** Let $f_1, f_2, f_3$ be three functions with graphs as below. Then, the max function consists of $f_1$ and $f_3$, which means the active index-set at the point $x^*$ (where the graphs of $f_1$ and $f_3$ intersect) is:

$$I(x^*) = \{1, 3\}$$

![Figure 2.5: Index set for max function](image-url)
CHAPTER 2. BACKGROUND

Definition 2.2.15 (Epigraph) Let $x \in \mathbb{R}$ and $f : \mathbb{R}^n \rightarrow \mathbb{R}$ with $f(x) < \infty$. Then, the epigraph of $f$ is the set

$$\text{epi}(f) := \{(x,y) \mid f(x) \leq y\}.$$ 

We observe that every convex function defined on the whole space is closed, that is, for every $a \in \mathbb{R}$, the set $\{x \mid f(x) \leq a\}$ is a closed set.

Lemma 2.2.16 Let $\{f_i\}_{i \in I}$ be a collection of closed convex functions with $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Let $f$ be the max function as defined previously. Then, we have

$$\partial f(x) \supseteq \text{conv}\{\cup \partial f_i(x) \mid i \in I(x)\}.$$ 

Note: This result is also true for Fréchet subdifferential (See Definition 2.2.20 later).

Proof Let $i \in I(x)$ and let $w \in \partial f_i(x)$. Then by Definition 2.2.11 of subdifferentials and max function, we have $f_i(x) = f(x)$ and for $w \in \partial f_i(x)$,

$$f(u) \geq f_i(u) \geq f_i(x) + \langle w, u-x \rangle, \quad \forall u \in \mathbb{R}^n.$$ 

Therefore, we have

$$\partial f(x) \supseteq \partial f_i(x).$$ 

Since the $\partial f_i$ are closed and convex, we know that $\partial f(x)$ must contain the closed convex hull of all $\partial f_i(x)$. Therefore,

$$\partial f(x) \supseteq \text{conv}\{\cup \partial f_i(x) \mid i \in I(x)\}.$$ 

Now we want to consider the other inclusion in Lemma 2.2.16, as we want to know when $\partial f(x)$ exactly equal to the closed convex hull of the subdifferentials $\partial f_i(x)$ at the active indices.
Definition 2.2.17 (Upper semi-continuity) A real-valued function $f$ is said to be upper semi-continuous at $x'$ if for every $\epsilon > 0$, there exists a neighbourhood $U_{x'}$ such that for all $x \in U_{x'}$, we have

$$f(x) \leq f(x') + \epsilon.$$ 

Intuitively, upper semi-continuity condition restricts the value of the function near $x'$ to be either close enough to $f(x')$ or less than $f(x')$.

The upper semi-continuity can also be expressed as:

$$\limsup_{x \to x'} f(x) \leq f(x').$$

A function is called upper semi-continuous if it is upper semi-continuous at every point on its domain.

Example 2.2.18

- The floor function $f(x) = \lfloor x \rfloor$ is upper semi-continuous.

- The indicator function (in measure theory) of any closed set is upper semi-continuous. Let $A \subseteq \mathbb{R}^n$ be a closed subset of $\mathbb{R}^n$, then the indicator function on $A$ is defined to be $\mathbb{1}_A : \mathbb{R}^n \to \{0, 1\}$:

$$\mathbb{1}_A := \begin{cases} 1, & \text{if } x \in A, \\ 0, & \text{if } x \notin A. \end{cases}$$

- A step function may be upper semi-continuous, but not left or right continuous. For example,

$$f(x) = \begin{cases} 1, & \text{if } x < 1, \\ 2, & \text{if } x = 1, \\ \frac{3}{2}, & \text{if } x > 1. \end{cases}$$

- Another interesting example of a function that is upper semi-continuous, while the left and right limit does not exist is $\sin(1/x)$, that is,

$$g(x) = \begin{cases} \sin \left( \frac{1}{x} \right), & \text{if } x < 1, \ x \neq 0, \\ 1, & \text{if } x = 1. \end{cases}$$
Lemma 2.2.16 tells us that: \( \partial f(x) \) contains the closed convex hull of subdifferentials \( \partial f_i(x) \) at the active indices. If we add some extra conditions, then we can obtain the other inclusion as in the following theorem.

**Theorem 2.2.19** Let \( I \) to be a compact index set, and \( \{f_i\}_{i \in I} \) a collection of convex functions. Let the function \((x, i) \mapsto f_i(x)\) to be jointly continuous, and let \((x, i, h) \mapsto f'_i(x, h)\) be jointly upper semi-continuous. Then, for \( f(x) = \sup_{i \in I} f_i(x) \), we have

\[
\partial f(x) = \text{conv}\{ \cup \partial f_i(x) \mid i \in I(x) \}.
\]

**Proof** Firstly, we want to show that \( f'(x, h) = \sup_{i \in I} f'_i(x, h) \).

To do this, we also need the following result:

**Claim**: Assume the conditions in Theorem 2.2.19, then \( f(x) \) is continuous.

**Proof of the claim**: As each \( f_i \) is lower semi-continuous, we have

\[
\{ x \mid f_i(x) > \alpha \}
\]

open for all \( \alpha \in \mathbb{R} \). Then,

\[
\{ x \mid f(x) > \alpha \} = \{ x \mid \sup_{i \in I} f_i(x) > \alpha \} = \bigcup_{i \in I} \{ x \mid f_i(x) > \alpha \},
\]

which is a union of open sets for all \( \alpha \in \mathbb{R} \). Therefore,

\[
\{ x \mid f(x) > \alpha \}
\]

is open.

So when \( f_i(x) \) is lower semi-continuous, \( f(x) \) is also lower semi-continuous.

Now we need to show that \( x \mapsto f(x) \) is upper semi-continuous.

Since \( I \) is compact, without loss of generality, taking the subsequence \( i_k \in I \), we can assume \( i_k \to i \in I \). Let \( x_k \to x \) attain \( \limsup_{y \to x} f(y) = \lim_{k \to \infty} f(x_k) \) with \( x_k \to x \), and let
We also need the following result:

**Claim:** \( x \rightarrow I(x) \) has a closed graph.

**Proof of the claim:** We want to show if \( i_k \in I(x_k) \) and \( (x_k, i_k) \rightarrow (x, i) \), then \( i \in I(x) \). For all \( k \), we have \( f_{i_k}(x_k) = f(x_k) \rightarrow f(x) \) as \( k \rightarrow \infty \) by continuity provided in the previous lemma. Since \( (x, i) \mapsto f_i(x) \) is jointly continuous,

\[
f(x) = \lim_k f_{i_k}(x_k) = f_i(x),
\]
which implies \( i \in I(x) \). □

Now we are ready to prove the Theorem 2.2.19. We will use the Mean Value Theorem from reference [12] Theorem 2.3.7 as below:

\[
f_i(y) - f_i(x) \in \langle \partial f_i(x + \mu(y-x)), y - x \rangle
\]

(2.3)

We also need Proposition 2.2.12, which implies

\[
\sup\{\langle x, y - x \rangle \mid x \in \partial f_i(x + \mu(y-x))\} = f'(x + \mu(y-x), y - x)
\]

(2.4)

Given all the assumptions from the statement of the theorem, the Mean Value Theorem as in (2.3), as well as the regularity condition in (2.4), we want to show the following equality:

\[
f'(x, h) = \lim_{t \downarrow 0} \frac{1}{t} \left( f(x + th') - f(x) \right) = \max_{i \in I(x)} f'_i(x, h).
\]

We will show both inequalities, and then we are done.

• Let \( i \in I(x) \), then we have

\[
\frac{1}{t} \left( f(x + th') - f(x) \right) \geq \frac{1}{t} \left( f_i(x + th') - f_i(x) \right)
\]

\[
\Rightarrow f'(x, h) \geq f'_i(x, h)
\]

\[
\Rightarrow f'(x, h) \geq \max_{i \in I(x)} f'_i(x, h)
\]
For the other inclusion, we need to take \((τ_k, h_k)\) to achieve the limit infimum, that is,

\[
f'(x, h) = \lim_{k} \frac{1}{τ_k} \left( f(x + τ_k h_k) - f(x) \right).
\]

Let \(i_k \in I(x + τ_k h_k)\), then,

\[
f'(x, h) \leq \lim_{k} \frac{1}{τ_k} \left( f_{i_k}(x + τ_k h_k) - f_{i_k}(x) \right).
\]

By Mean Value Theorem, we have

\[
f_{i_k}(x + τ_k h_k) - f_{i_k}(x) \in τ_k \langle ∂f_{i_k}(x + μ_k τ_k h_k), h_k \rangle, \quad μ_k \in (0, 1)
\]

\[
⇒ \frac{1}{τ_k} \left( f_{i_k}(x + τ_k h_k) - f(x) \right) \leq \sup \{ ⟨x^*_k, h_k⟩ \mid x^*_k ∈ ∂f_{i_k}(x + μ_k τ_k h_k) \}
\]

\[
= f'_{i_k}(x + μ_k τ_k h_k, h_k)
\]

Now, let \(k \to ∞\). Since \(I\) is compact, we can take a subsequence in the set so that \(i_k \to i \in I(x)\). Therefore,

\[
f'(x, h) \leq f'_i(x, h), \quad i \in I(x)
\]

which gives the inequality

\[
f'(x, h) \leq \sup_{i \in I(x)} f'(x, h).
\]

Now, we can use the result from Theorem 2.2.19. Then we have the following equivalent expression for subdifferential.

Let \(g : \mathbb{R}^n \to \mathbb{R}\) be a max-function, then we can express the subdifferential \(∂g(x)\) as:

\[
∂g(x) = \text{conv}\{∇g_i(x) \mid i \in I(x)\},
\]

where \(I(x) := \{i = 1, ..., m \mid g_i(x) = g(x)\}\).

In Chapter 4, we will require the following definition of family \(D(\bar{x})\). The subsets \(D \subseteq I(\bar{x})\) are used to obtain the hyperplane which has union of the support faces, and zero is strictly on the same side of the subdifferential.

We can also restrict the directional derivative to be a linear function, which gives us the \(Gâteaux\ differential\).
2.2.3 Regular and Limiting Subdifferentials

In this section, we will talk about subdifferentials for non-convex functions.

**Definition 2.2.20 (Fréchet subdifferential)** Let $f$ be a function such that $f : \mathbb{R}^n \to \mathbb{R}$, and $f$ is finite at the point $x \in \mathbb{R}^n$. Then, the Fréchet Subdifferential of $f$ at $x$ is defined as the following:

$$\partial f(x) = \left\{ w \in \mathbb{R}^n \mid \liminf_{u \to x} \frac{f(u) - f(x) - \langle w, u - x \rangle}{||u - x||} \geq 0 \right\}.$$

If $\partial f(x) \neq \emptyset$, then we say $f$ is Fréchet subdifferentiable at $x$.

**Proposition 2.2.21** Let $f$ be a function that is Fréchet differentiable at $x$ and the derivative of $f$ at $x$ is $\nabla f(x)$. Then, we have,

$$\partial f(x) = \{ \nabla f(x) \}.$$

**Proof** See [27] Corollary 4.4.4 and Example 3.4.

We can also define Fréchet superdifferential using the similar idea.

**Definition 2.2.22 (Fréchet superdifferential)** Let $f$ be a function such that $f : \mathbb{R}^n \to \mathbb{R}$, and $f$ is finite at the point $x \in \mathbb{R}^n$. Then, the Fréchet superdifferential is defined as the following set:

$$\partial^+ f(x) = \left\{ w \in \mathbb{R}^n \mid \limsup_{u \to x} \frac{f(u) - f(x) - \langle w, u - x \rangle}{||u - x||} \leq 0 \right\}.$$

If $f$ is not differentiable at $x$, it may happen that both $\partial f(x)$ and $\partial^+ f(x)$ are empty. For example, $|x| - |y|$ at $0$ has $\partial^+ f(x) = \partial f(x) = \emptyset$.

In [33], we have the following proposition.
Proposition 2.2.23 \( f \) is Fréchet differentiable at \( x \) if and only if both \( \partial f(x) \) and \( \partial^+ f(x) \) exist. That is, by Proposition 2.2.21, we have,

\[
\partial f(x) = \partial^+ f(x) = \{ \nabla f(x) \}.
\]

Example 2.2.24 Consider the function \( f : \mathbb{R} \to \mathbb{R} \) defined to be:

\[
f(x) = x \sin \left( \frac{1}{x} \right), \quad x \neq 0.
\]

Then, both \( \partial f(x) \) and \( \partial^+ f(x) \) are empty at \( x = 0 \).

Now, we would like to define outer limits, which is the main focus of Chapter 4.

Firstly, we denote \( \mathbb{N}^\# \) to be all subsequences of \( \mathbb{N} \).

Definition 2.2.25 (Outer limits) Let \( \{A_n\}_{n \in \mathbb{N}} \) to be a sequence of subsets of \( \mathbb{R}^n \). Then, the outer limit is defined to be

\[
\operatorname{Lim sup}_{n \to \infty} A_n := \{ x \mid \exists N \in \mathbb{N}^*, \exists x_n \in A_n(n \in N), \text{ with } x_n \to x \}.
\]

The concept of limiting subdifferentials can be viewed as the limits of usual subdifferentials. Let \( f : \mathbb{R}^n \to \mathbb{R} \) be lower semicontinuous in a neighborhood of \( \bar{x} \in \mathbb{R}^n \). Let \( \{x_k\} \) be a sequence in \( \mathbb{R}^n \) and \( x_k \to \bar{x} \) with \( f(x_k) \to f(\bar{x}) \).

Then we can define the limiting subdifferentials by using the two conditions for sequences above.

Definition 2.2.26 (Limiting subdifferentials) Let \( f : \mathbb{R}^n \to \mathbb{R} \) be lower semicontinuous in a neighbourhood of \( \bar{x} \in \mathbb{R}^n \). Then, the limiting subdifferential is defined to be the set:

\[
\bar{\partial} f(\bar{x}) := \operatorname{Lim sup}_{x \to \bar{x}} \partial f(x),
\]

where \( \partial f(x) \) is the Fréchet subdifferential.
To be able to define Clarke regularity, firstly, we need to define tangent cone and regular normal cone.

**Definition 2.2.27 (Tangent cone)** Let $C$ be a set in $\mathbb{R}^n$ and $x \in C$. Then the tangent cone can be defined as

$$T_C(x) := \{ y \in \mathbb{R}^n \mid \exists x_k \subset C, t_k \subset \mathbb{R} : t_k \to 0, x_k \to x, \frac{x_k - x}{t_k} \to y \}.$$ 

**Definition 2.2.28 (Regular normal cone)** Let $T_C(x)$ be the tangent cone of $C$. Then, the regular normal cone to $C$ at $x$ is the polar cone of $T_C(x)$, denoted by $\hat{N}_C(x)$.

Now, consider a vector $v \in \mathbb{R}^n$, if there exists a sequence $x_k \to x$ and $v_k \to v$ with $x_k \in C$ and $v_k \in \hat{N}_C(x_k)$ for all $k$. Then, we say that $v$ belongs to the normal cone $N_C(x)$ to $C$ at $x$.

**Definition 2.2.29 (Clarke regularity)** A closed and nonempty set $X \subseteq \mathbb{R}^n$ is said to be Clarke regular if $\hat{N}_C(x) = N_C(x)$.

**Example 2.2.30** Consider the point $(1,0)$ in the boundary of the set

$$C = \{(x,y) \mid (x-1)^3 - y^2 \leq 0\}.$$ 

Then, $(t + 1, t^{3/2}) \in C$ and therefore

$$\lim_{t \to 0} \frac{(t + 1, t^{3/2}) - (1,0)}{t} = (1,0).$$

Therefore, we have

$$T_C(1,0) = \{(1,0)\}.$$ 

Hence,

$$N_C(1,0) = \hat{N}_C(0,1) = \{(u,v) \mid u \leq 0\}.$$ 

Therefore $(1,0)$ is a regular point of $C$ is the sense of Clarke regularity.
Now we are going to define *regular function*, which will be used in Chapter 4 in our answer to an open problem.

**Definition 2.2.31 (Regular function)** Let \( f : \mathbb{R}^n \to \mathbb{R} \) to be a function, we say \( f \) is regular at \( \bar{x} \in \mathbb{R}^n \) if the epigraph of \( f \) is regular in the sense of Clarke at \((\bar{x}, f(\bar{x}))\) as a subset of \( \mathbb{R}^n \times \mathbb{R} \).

Regularity conditions are important in optimisation. Some problems must satisfy some regularity conditions to be able to allow the minimum solution to satisfy necessary conditions such as primal feasibility, dual feasibility, or complementary slackness. In particular, regularity of a function enables a much tighter description of the local geometry of the graph of the function to be deduced from its limiting subdifferential. More details can be found in some classic review papers [20] [24] [34] [44].
Chapter 3

Demyanov-Ryabova Conjecture

In general, we can study any smooth function with the help of a main tool, the gradient. By using the gradient, we can obtain a first-order approximation of the function, describe its optimality conditions, find the steepest ascent and descent directions, and so on.

However, with nonsmooth functions, we have to use different tools, one of them is directional derivatives. For convex and max-type functions, their directional derivatives are convex. Therefore, by Minkowski duality, the optimality conditions can be stated in geometric terms, and the steepest descent directions can also be derived in this case.

Exhausters can be employed to describe optimality conditions, in particular, they are very useful for finding directions of steepest ascent and descent for a very wide range of nonsmooth functions. If a point is not stationary, then directions of the steepest ascent and descent can also be calculated by means of exhausters. In other words, exhauster is a powerful tool for finding minima and maxima of functions.

The notions of upper and lower exhauster of a positively homogeneous function $h : \mathbb{R}^n \to \mathbb{R}$ are introduced. The upper exhauster and lower exhauster are more explicitly introduced in Section 3.1. An upper exhauster can be converted into a lower one and vice versa using a convertor operator introduced in [15]. Upper exhauster is a more convenient tool for checking the conditions for a minimum (and vice versa, lower exhauster is better suited for maximum); conversion is also necessary for the application of some calculus rules.

In this chapter, we are going to look at Demyanov-Ryabova conjecture. The conjecture claims that, suppose we convert between finite families of upper and lower exhausters with
the given convertor function (Demyanov convertor), then such process will reach a cycle of length at most two. We will show that the conjecture is true in affinely independent case first. Then, we will construct an equivalent combinatorial reformulation of the conjecture.

3.1 Background

Given a finite collection of polytopes, we can obtain its dual by taking convex hulls of the support faces for every nonzero direction. Then, if we continue this process, it will inevitably reach a cycle due to the finiteness of the problem. The Demyanov-Ryabova conjecture states that such a cycle will have length at most two. Essentially, we want to establish the uniqueness of a dual characterization of a function by establishing a steady 2-cycle in the relevant dynamical system defined by the conversion operator.

Exhausters are multiset objects, that generalise the subdifferential of a convex function. An upper or lower exhauster (See Definition 3.1.1 and Definition 3.1.2) is the collection of sub- or superdifferentials, that correspond to the relevant upper or lower representations of the directional derivative. Introduced by Demyanov in [14], exhausters attracted a noticeable following in the optimization community [1, 2, 16, 21, 36, 38, 43, 58]. Such constructions are popular in applied optimisation as they allow for exact calculus rules and easy conversion from “upper” to “lower” characterisations of the directional derivative. The Demyanov-Ryabova conjecture gives an elegant interpretation in terms of lower convex and upper concave representations of positively homogeneous functions. One of the main challenges in the calculus of exhausters is their lack of uniqueness, and whilst some works are dedicated to finding minimal objects [50], it is shown that a minimal exhauster may not exist (there exist different representations, that cannot be reduced further; see [23]). The resolution of the Demyanov-Ryabova conjecture is key to identifying “stable” upper and lower exhauster representations.

For each collection of compact convex sets, the Minkowski duality gives a corresponding unique collection of support functions. The infimum over this collection is some positively homogeneous function. The dual family of sets obtained after the conversion, when interpreted as superdifferentials of superlinear functions, gives a symmetric lower representation
of the same homogeneous function as the supremum over this family of superlinear functions. This approach emerges from the calculus of exhausters, which serve as constructive generalisations of the gradient and are used for optimality conditions, construction of numerical methods in nonsmooth problems and notably for the computation of other generalised subdifferentials.

Exhausters and other constructive generalisations of the convex subdifferential such as quasi- and codifferentials allow for straightforward generalisation of Minkowski duality in the sense described above that is not available for other classic constructions [31,33]. Neither the essentially primal graphical derivatives [48] nor dual coderivative objects [41] allow for well-defined dual characterisations. Constructive nonsmooth subdifferentials are well suited for practical applications, especially in finite dimensional continuous problems, and have been utilised successfully both in applied problems such as data classification (see an overview [8]) and in theoretical problems coming from other fields, such as spline approximation [55].

**Definition 3.1.1** Given a positively homogeneous function \( h : \mathbb{R}^m \rightarrow \mathbb{R} \), its upper exhauster \( E^* \) is a family of closed convex sets such that \( h \) has an exact representation

\[
h(x) = \inf_{C \in E^*} \max_{v \in C} \langle v, x \rangle,
\]

so that \( h \) is the infimum over a family of sublinear functions.

**Definition 3.1.2** An upper exhauster \( E^* \) is the collection of subdifferentials of these functions, with the following representation,

\[
h(x) = \sup_{C \in E} \min_{w \in C} \langle w, x \rangle.
\]

The lower exhauster \( E_*h \) is defined symmetrically as a supremum over a family of superlinear functions.

Exhausters constructed for first order homogeneous approximations of nonsmooth functions (such as Dini and Hadamard directional derivatives) provide sharp optimality conditions, moreover, exhausters enjoy exact calculus rules which makes them an attractive tool for applications.
When the positively homogeneous function $h$ is piecewise linear, it can be represented as a minimum over a finite set of piecewise linear convex functions described by the related finite family of polyhedral subdifferentials. The exhauster conversion operator allows to obtain symmetric local representation as the maximum over a family of polyhedral concave functions, and vice versa, where the families of sets remain finite and polyhedral. The Demyanov-Ryabova conjecture states that if this conversion operator is applied to a family of polyhedral sets sufficiently many times, the process will stabilise with a 2-cycle. Here we focus on a geometric formulation of this conjecture that does not rely on nonsmooth analysis background. Consider the Example 1.1.1, the polygon in $\mathbb{R}^2$ is the subdifferential at zero, and each face can be obtained after applying the exhauster conversion operator given the corresponding directional derivative.

The contribution of this work is twofold: firstly, we prove that the conjecture is true in the special affinely independent case, when all vertices of the polytopes in the collection form a simplex (Theorem 3.3.4). So, we will restrict the conjecture to the case with $n + 1$ affinely independent vertices in an $n$ dimensional space, and prove it is true. Secondly, we will reformulate this geometric problem and obtain a combinatorial formulation by considering the orderings on the vertex set and forming a simplified map (Theorem 3.5.4). Then, we will show the combinatorial formulation and the geometric problem are equivalent. Apart from these, we have also done some numerical experiments on some random polytopes in 2D.

### 3.2 Preliminaries

**Definition 3.2.1** Let $d \in S^{m-1}$ and let $P \subset \mathbb{R}^m$ be a polytope. We define the maximal face or supporting face of $P$ for the direction $d$ as $P_d := \arg \max_{x \in P} \langle x, d \rangle$.

Geometrically, the maximal face $P_d$ in the direction $d$ is the subset of $P$ of all points that project ‘the furthest’ along this direction.
We note here that for any polytope $P$, its face $P_d$ is the convex hull of a subset of vertices of $P$.

Consider a finite collection $\Omega = \{P_1, P_2, \ldots, P_k\}$, where $P_i \subset \mathbb{R}^m$ is a polytope for each $i \in \{1, \ldots, k\}$, and $k \in \mathbb{N}$. For a direction $d \in S^{m-1}$, let $\Omega(d) := \text{conv}\{P_d : P \in \Omega\}$.

Here we fix a direction $d$, and take the convex hull of the maximal faces in this direction for all polytopes in the family. The new collection of sets generated in this fashion from all directions $d \in S^{m-1}$ is the output of the Demyanov convertor,

$$F(\Omega) := \{\Omega(d) : d \in S^{m-1}\}. \quad (3.2.1)$$

It is not difficult to observe that for every $d \in S^{m-1}$ the set $\Omega(d)$ is a polytope, since each maximal face $P_d$ is a convex hull of finitely many vertices of the polytope $P$, and there are finitely many such polytopes. Moreover, since the total number of vertices of all polytopes in the collection is finite, there are finitely many subsets of these vertices that give finitely many possibilities to form different convex hulls. Hence, $F(\Omega)$ is also a collection of finitely many polytopes. We can now define a sequence $(\Omega_i)$ recursively as

$$\Omega_{i+1} = F(\Omega_i), \quad i \in \{0, 1, \ldots\}, \quad \Omega_0 := \Omega. \quad (3.2.2)$$

**Conjecture 3.2.2 (Demyanov-Ryabova)** Given a finite collection of polytopes $\Omega_0$, the sequence $\{\Omega_i\}_{i \in \mathbb{N}}$ defined via the recursive application of Demyanov convertor $F$ to $\Omega_0$ eventually reaches a cycle of length 2, i.e. there exists $N \in \mathbb{N}$ such that $\Omega_{n+2} = \Omega_n$ for all $n > N$. 

Figure 3.1: Maximal face of a polytope $P$ for a given direction
We can also state the conjecture in the following two ways:

**Restatements of the Conjecture 3.2.2:**

1. There exist an $N \in \mathbb{Z}_{>0}$ such that if $n > N$, then any polytope $P$ satisfies $P \in \Omega_n \iff \Omega_{n+2}$.

2. Let $P$ be a polytope. Then there exist $N \in \mathbb{Z}_{>0}$ such that if $n > N$, then $P \in \Omega_n \iff P \in \Omega_{n+2}$.

In the statement (1), we are given a specific index, and claiming on the existence of the polytope $P$. While in the statement (2), we are given an arbitrary polytope $P$ as the condition, then claiming on the existence of the index.

**Definition 3.2.3** Let $\Omega$ be a set of polytopes. We define $V$ to be the set of all vertices of all polytopes in the collection $\Omega$, i.e. $V = \bigcup_{P \in \Omega} \text{ext } P$.

The following lemma is quite straightforward, but we will provide a rigorous proof for completeness.

**Lemma 3.2.4** Let $\Omega_0$ be a finite family of polytopes in $\mathbb{R}^m$. Then, the following statements are equivalent.

1. There exists an $N \in \mathbb{N}$ such that $\Omega_n = \Omega_{n+2}$ for all $n > N$.

2. For any polytope $P \subset \mathbb{R}^m$, there exists an $N_P \in \mathbb{N}$ such that $\forall n > N_P : P \in \Omega_n \Rightarrow P \in \Omega_{n+2}$.

**Proof** It is evident that (i) yields (ii) by letting $N_P = N$ for all $P$. Now, we want to show the reverse implication.

By the construction of our conversion process the only polytopes that feature in any of the collections in $(\Omega_i)_i$ are the convex hulls of subsets of $V$. There are finitely many such subsets. Therefore for all but finitely many polytopes we can safely let $N_P = 0$.

Consider the remaining finite set of polytopes $\mathcal{P}$ that appear at least once in some of the collections in our conversion sequence. If $P \in \mathcal{P}$ appears in the sequence finitely many
times, then \( N_P \) has to be larger than the index \( n \) of the last collection \( \Omega_n \) that contains \( P \).

If \( P \) features in infinitely many \( \Omega_i \)'s, then it has to be present in each \( \Omega_{N_P+2k} \) for \( i \geq N_P \) and \( k \in \mathbb{N} \), where \( N_P \) is the index that satisfies \((ii)\).

It remains to assign the maximal number \( N_P \) over all polytopes to \( N \), and observe that \((i)\) holds for this \( N \). \(\square\)

Note that, while a large part of the discussion in this section can be repeated verbatim for the case when \( \Omega \) is a bounded family of compact convex sets, the reformulation of the conjecture given in Lemma 3.2.4 \((ii)\) is not necessarily true for this case.

### 3.3 Affinely Independent Case

We first demonstrate the result that the convex hull \( \text{conv} \Omega_i \) is a constant.

**Proposition 3.3.1** Let \( \Omega \) be a finite collection of polyhedral sets in \( \mathbb{R}^m \), then, \( \text{conv} \Omega = \text{conv} F(\Omega) \), where \( \text{conv} \Omega = \text{conv}\{P : P \in \Omega\} \) is the convex hull of all polytopes in \( \Omega \), and \( F(\Omega) \) is the output of the Demyanov convertor \((3.2.1)\).

**Proof** It is evident that \( \text{conv} F(\Omega) \subset \text{conv} \Omega \); we only need to show that no points of the convex hull are lost in the conversion. For this it is sufficient to prove that for every vertex \( v \) of \( C = \text{conv} \Omega \) we have \( v \in \text{conv} F(\Omega) \).

Since \( C \) is a polytope, every vertex \( v \) of \( C \) is exposed. Thus, there exists \( d \in S^{m-1} \) such that

\[
C_d = \text{Arg max}_{x \in C} \langle x, d \rangle = \{v\}.
\]

Since \( C = \text{conv}\{P : P \in \Omega\} \), there exists a polytope \( P \in \Omega \), such that \( v \in P \) (since \( v \) is an extreme point, it can not be represented as a convex combination of any other points in \( C \)).

On the other hand, since \( P \subset C \), we have

\[
\langle v, d \rangle \leq \max_{x \in P} \langle x, d \rangle \leq \max_{y \in C} \langle y, d \rangle = \langle v, d \rangle,
\]

hence, \( v \in P_d \subset \Omega(d) \), and therefore \( v \in \text{conv} F(\Omega) \). \(\square\)
Remark 3.3.2 According to the Proposition 3.3.1, the sequence \((\text{conv } \Omega_i)_{i \in \mathbb{R}}\) is constant. Thus, we can define \(C\) to be the convex hull of all polytopes in the collection, i.e.

\[
C = \text{conv}(\Omega_0) = \text{conv}(\Omega_1) = \ldots = \text{conv}(\Omega_i) \quad \forall \ i
\]

see Fig. 3.2.

![Figure 3.2: Convex hull C of 5 sets](image)

The set \(C\) is a polytope by the finiteness argument. Our goal is to prove the Demyanov-Ryabova conjecture for the special case when \(C\) is a convex hull of an affinely independent set, and all vertices of the polytopes in the collection belong to this set.

We denote \(C_d\) as the maximal face of \(C\) in direction \(d \in S^{m-1}\).

Recall that a finite set of points \(V = \{v_0, \ldots, v_k\} \subset \mathbb{R}^m\) is affinely independent if the vectors

\[
p_i = v_i - v_0, \quad i \in \{1, 2, \ldots, k\}
\]

span a \(k\)-dimensional linear subspace of \(\mathbb{R}^m\), or equivalently if the convex hull of \(V\) has dimension \(k\). The following definition of a simplex will be useful for us in the sequel.

**Definition 3.3.3 (Simplex)** Let \(k + 1\) points \(v_0, v_1, \ldots, v_k \in \mathbb{R}^m\) be affinely independent.
The simplex determined by this set of points is their convex hull:

\[ C = \{ \lambda_0 v_0 + \cdots + \lambda_k v_k : \lambda_i \geq 0, \sum_{i=0}^{k} \lambda_i = 1 \}. \]

Thus, a \( k \)-simplex is a \( k \)-dimensional polytope that is the convex hull of its \( k + 1 \) vertices. Observe that every face of a simplex, called sub-simplex, is also a simplex of a lower dimension. We also mention here that every sub-simplex of a \( k \)-simplex is an exposed face of this simplex, i.e. it is a maximal face for some direction \( d \in S^{m-1} \) (see [61]).

Our next goal is to prove the following special case of Conjecture 3.2.2.

**Theorem 3.3.4** Let \( \Omega_0 \) be a finite collection of polytopes in \( \mathbb{R}^m \). Assume that there exists an affinely independent set \( V = \{ v_0, v_1, \ldots, v_k \} \subset \mathbb{R}^m \), such that \( C := \text{conv}(\Omega_0) = \text{conv}\{ v_0, v_1, \ldots, v_k \} \), that is, \( C \) is a \( k \)-simplex on the set of vertices \( V \). Assume that every polytope \( P \in \Omega_0 \) is a sub-simplex of \( C \), i.e. there exists \( V_P \subset V \) such that \( P = \text{conv} V_P \). Then, the conversion process (3.2.2) reaches a cycle of length 2, that is, there exists a sufficiently large \( N \in \mathbb{N} \) such that we have, \( \Omega_{n+2} = \Omega_n \), for all \( n > N \).

The proof of Theorem 3.3.4 is based on a technical claim that we prove in Lemma 3.3.5.

**Lemma 3.3.5** Let \( \Omega \) be a finite collection of polytopes, let the sequence \((\Omega_i)\), be defined by (3.2.2) with \( \Omega_0 = \Omega \) and let \( C = \text{conv}(\Omega) \). If for some \( d \in S^{m-1} \) and \( n \in \mathbb{N} \) we have \( C_d \in \Omega_n \), then \( C_d \in \Omega_{n+2} \).

**Proof** Let \( C_d \in \Omega_n \), where \( d \in S^{m-1} \). Observe that \( (C_d)_d = C_d \), hence,

\[ C_d = (C_d)_d \subseteq \Omega_n(d) = \text{conv}\{ P_d : P \in \Omega_n \}, \]

therefore,

\[ C_d = (C_d)_d \subseteq \Omega_n(d)_d. \quad (3.3.1) \]

On the other hand, since \( \Omega_n(d) \subseteq C \), we have

\[ \Omega_n(d)_d \subseteq C_d. \quad (3.3.2) \]

Putting (3.3.1) and (3.3.2) together, we obtain \( C_d = \Omega_n(d)_d \). Since \( \Omega_n(d) \in \Omega_{n+1} \), we have one inclusion \( \Omega_{n+1}(d) \supseteq \Omega_n(d)_d = C_d \).
To finish the proof it remains to show the reverse inclusion $\Omega_{n+1}(d) \subseteq C_d$ (then $C_d = \Omega_{n+1}(d) \in \Omega_{n+2}$ and we are done). This is equivalent to showing that, if $P \in \Omega_{n+1}$, then $P_d \subseteq C_d$.

For any $P \in \Omega_{n+1}$ there exists $d' \in S^{m-1}$ such that $P = \Omega_n(d')$, therefore,

$$P = \Omega_n(d') = \operatorname{conv}\{P_{d'} : P \in \Omega_n\} \supseteq (C_d)_{d'}.$$

Since $(C_d)_{d'}$ is nonempty, this yields $P \cap C_d \neq \emptyset$, and hence, $P_d = (P \cap C_d) \subseteq C_d$. □

**Proof** (of Theorem 3.3.4) We will show the equivalent claim (see Lemma 3.2.4 (ii)) that for every polytope $P$ there exists $N_P \in \mathbb{N}$ such that

$$\forall n > N_P : \ P \in \Omega_n \Rightarrow P \in \Omega_{n+2}. \tag{3.3.3}$$

First consider the case when $P \neq C = \operatorname{conv}(\Omega)$. There must be a direction $d \in S^{m-1}$, such that $P$ is a maximal face of $C$, i.e. $P = C_d$. Then (3.3.3) follows from Lemma 3.3.5.

It remains to show that (3.3.3) is true for $P = C$, i.e. there exists $N \in \mathbb{N}$, such that for all $n > N$, $C \in \Omega_n \Rightarrow C \in \Omega_{n+2}$.

Assume the contrary, then $C \in \Omega_N$ and $C \notin \Omega_{N+2}$ for some $N$. By the construction of our sequence $\{\Omega_i\}$ we know that $C = \Omega_{N-1}(d)$ for some $d \in S^{m-1}$. However, $\Omega_{N+1}(d) \neq C$ since $C \notin \Omega_{N+2}$ by the assumption.

There exists a vertex $a \in V$ of $C$ such that $a \notin \Omega_{N+1}(d)$. Since $a \in C \subset \Omega_{N-1}(d)$, this implies $a \in P_d$ for some $P \in \Omega_{N-1}$, hence, $P \notin \Omega_{N+1}$. Indeed, assuming the contrary, we would have $a \in P_d \subset \Omega_{N+1}(d)$, which contradicts our choice of the vertex $a$.

Since $P \in \Omega_{N-1}$ and $P \notin \Omega_{N+1}$, we have $P = C$ (otherwise we obtain a contradiction to Lemma 3.2.4). Hence, $C \in \Omega_{N-1}$ and $C \notin \Omega_{N+1}$. Now we can repeat the same argument with $N' = N - 1$ and deduce that $C \in \Omega_{N-2}$, $C \notin \Omega_N$, but the latter is a contradiction to our assumption that $C \in \Omega_N$. □

### 3.4 The Simplified Demyanov Convertor

In this section, we prove that Conjecture 3.2.2 has an equivalent formulation, which will be the combinatorial problem we are going to explain in detail in Section 3.5. The first step in
this reformulation is the observation that it is enough to consider a certain dense subset of \( S^{m-1} \) to obtain the conversion sequence \((\Omega_i)\), from the initial collection of polytopes \( \Omega = \Omega_0 \).

The second step consists of further simplifications by mapping these directions to a finite subset of the symmetric group \( S_k \), where \( k \) is the cardinality of the set \( V \) of all vertices, and realising the convertor as a transformation of collections of subsets of integers in \( \{1, 2, \ldots, k\} \).

Our first step is to introduce a reduced transformation \( F' \), that ignores all directions for which some of the support faces in the collection of polytopes are not singletons. Given a finite collection of polytopes \( \Omega \), and let \( V \) be the set of all vertices of all polytopes in this collection, \( V(\Omega) = V = \bigcup_{P \in \Omega} \text{ext}\ P \).

We throw away all directions that can potentially result in a nonsingleton maximal face at some step of the conversion, and let

\[
\widehat{S}^{m-1} := \{ d \in S^{m-1} : \langle v, d \rangle \neq \langle w, d \rangle, \forall v \neq w \in V(\Omega) \}.
\]

We will use the fact that \( \widehat{S}^{m-1} \) is a dense subset of \( S^{m-1} \) in later proofs. This is easy to see by noting that, if \( v_1, \ldots, v_k \) are the vertices of all polytopes in \( \Omega \), then each of the sets

\[
V_{ij} = \{ x : \langle v_i - v_j, x \rangle = 0 \}, \quad i \neq j, \ i, j \in \{1, 2, \ldots, k\}
\]

is a hyperplane in \( \mathbb{R}^m \), so the cone \( V = \mathbb{R}^m \setminus \bigcup_{i \neq j} V_{ij} \) is dense in \( \mathbb{R}^m \), and hence \( \widehat{S}^{m-1} = S^{m-1} \cap V \) is dense in \( S^{m-1} \).

We define a modified transformation \( F' \) by ignoring the directions in \( S^{m-1} \setminus \widehat{S}^{m-1} \),

\[
F'(\Omega) := \{ \Omega(d) : d \in \widehat{S}^{m-1} \}, \tag{3.4.1}
\]

and build the modified sequence \( \Omega'_1, \Omega'_2, \ldots \) obtained by the recursive application of \( F' \) to \( \Omega_0 = \Omega \),

\[
\Omega'_{i+1} = F(\Omega'_i), \quad i \in \{0, 1, \ldots\}, \quad \Omega_0 := \Omega. \tag{3.4.2}
\]

It is natural to ask whether \( F'(\Omega_i) = F'(\Omega'_i) \) and \( F(\Omega_i) = F(\Omega'_i) \) for all \( i \in \mathbb{N} \). We answer this in the affirmative later on (see Theorem 3.4.6).

We will use two well known results (see [61]), which we prove here for completeness.

**Lemma 3.4.1** Let \( P \subset \mathbb{R}^m \) be a polytope, and let \( d \in S^{m-1} \). There exists a neighbourhood \( N_d \) of \( d \), such that \( \forall d' \in N_d : P_{d'} \subset P_d \).
CHAPTER 3. DEMYANOV-RYABOVA CONJECTURE

Proof Assume the contrary, without lost of generality, there is a sequence \((d_k), d_k \to d\) and a vertex \(v\) of \(P\), such that \(v \in P_{d_k}\), but \(v \notin P_d\). This is impossible because

\[
\max_{x \in P} \langle x, d \rangle = \lim_{k \to \infty} \max_{x \in P} \langle x, d_k \rangle = \lim_{k \to \infty} \langle v, d_k \rangle = \langle v, d \rangle.
\]

Therefore, \(v \in P_d\). \(\square\)

We prove the following via elementary methods but it is also an application of Radamacher’s Theorem as applied to the differentiability of the support function. The support is convex and bounded and hence Lipschitz. By Rademacher’s Theorem, it is almost everywhere differentiable. It is well known that in a direction of differentiability of the support function the gradient gives us the unique exposed point.

Lemma 3.4.2 Let \(P\) be a polytope, and let \(v\) be a vertex of \(P_d\) for some \(d \in S^{m-1}\). Then, in any neighbourhood \(N_d\) of \(d\) there exists \(d' \in N_d \cap S^{m-1}\) such that \(P_{d'} = \{v\}\).

Proof Observe that if \(P_d = \{v\}\), there is nothing to prove. We hence assume that the dimension of \(P_d\) is at least 2. Let \(d'' \in S^{m-1}\) be some direction, that exposes the vertex \(v\) within the affine hull of \(P_d\), i.e. \(d'' \in \text{aff } P_d - v\) and \((P_d)_{d''} = \{v\}\). Consider the parametric family \(d_t := d + t(d'' - d), t \in [0, \infty)\).

Since \(d_t \to d\) as \(t \to 0\), by Lemma 3.4.1, there exists a sufficiently small \(t_0\) such that

\[
\forall t \in [0, t_0) : P_{d_t} \subset P_d. \tag{3.4.3}
\]

On the other hand, observe that for any \(x \in P_d\), we have

\[
\langle x, d_t \rangle = (1 - t)\langle x, d \rangle + t\langle x, d'' \rangle = (1 - t)\langle v, d \rangle + t\langle x, d'' \rangle,
\]

hence,

\[
\text{Arg max}_{x \in P_d} \langle x, d_t \rangle = \text{Arg max}_{x \in P_d} \langle x, d'' \rangle = \{v\},
\]

and together with (3.4.3), we have \(\forall t \in [0, t_0) : P_{d_t} = \{v\}\).

Observe that this relation is also true for the normalised vectors \(d_t/\|d_t\|\), which converge to \(d\) as well:

\[
\lim_{t \to 0} \frac{d_t}{\|d_t\|} = d,
\]
This result also follows from the Rademacher Theorem.

Hence we are able to choose $d'' \in S^{m-1}$ arbitrarily close to $d$ so that $P_{d''} = \{v\}$. □

**Proposition 3.4.3** Let $\Omega$ be a finite collection of polytopes in $\mathbb{R}^m$, and let $d \in S^{m-1}$. For any $P \in \Omega$ and any vertex $v$ of $P_d$, there exists another direction $d' \in S^{m-1}$ and a neighbourhood $N_{d'}$ of $d'$, such that

$$\forall d'' \in N_{d'}, \forall P' \in \Omega : \quad P_{d''} = \{v\} \quad \text{and} \quad P'_{d''} \subset P'_{d'}.$$  \hfill (3.4.4)

**Proof** Let $\Omega$, $P$, $v$ and $d$ as in the statement of the proposition. By Lemma 3.4.1 for every polytope $P' \in \Omega$, there exists a sufficiently small neighbourhood $N_{d'}$ of $d$, such that, 

$$\forall d' \in N_d : \quad P'_{d'} \subset P'_{d'}.'$$

We let $N_d := \bigcap_{P' \in \Omega} N_{d'}$, then, $\forall d' \in N_d, \forall P' \in \Omega : \quad P'_{d'} \subset P'_{d'}.$

Since $v$ is a vertex of $P_d$, by Lemma 3.4.2, we can find another direction $d'' \in N_d$, such that $P_{d''} = \{v\}$.

Applying Lemma 3.4.1 to the direction $d''$, and to our vertex $v$ of $P_{d''}$, we deduce that, there is a neighbourhood $N_{d''}$, such that, for all $d'' \in N_{d''}, P_{d''} \subset P_{d''} = \{v\}.$

Hence, $\forall d'' \in N_{d''} : \quad P_{d''} = \{v\}.$

To finish the proof we observe that, the neighbourhood $N_{d''} := N_{d''} \cap N_d$ satisfies (3.4.4). □

**Proposition 3.4.4** Let $\Omega$ be a finite collection of polytopes in $\mathbb{R}^m$, and let $\tilde{S}$ be a dense subset of $S^{m-1}$. Let $v$ be a vertex of $\Omega(d)$ for some $d \in S^{m-1}$. Then, there exists $\tilde{d} \in \tilde{S}$, such that, $\Omega(\tilde{d}) \subseteq \Omega(d)$ and $v \in \Omega(\tilde{d})$.

**Proof** If $v$ is a vertex of $\Omega(d)$, then there exists a polytope $P \in \Omega$, such that $v$ is a vertex of $P_d$. By Proposition 3.4.3, there is another direction $d' \in S^{m-1}$, such that,

$$\forall d' \in N_{d'}, \forall P' \in \Omega : \quad P_{d'} = \{v\} \text{ and } P'_{d'} \subset P'_{d'}.$$  \hfill (3.4.5)

Since the set $\tilde{S}$ is dense in $S^{m-1}$, there exists $\tilde{d} \in \tilde{S} \cap N_{d'} \neq \emptyset$ that satisfies (3.4.5). We have

$$\Omega(\tilde{d}) = \text{conv}\{P_d, P \in \Omega\} \subseteq \text{conv}\{P_d, P \in \Omega\},$$

and also $v \in \{v\} = P_d \subset \Omega(\tilde{d})$. □
Chapter 3. Demyanov-Ryabova Conjecture

Proposition 3.4.5 For any finite collection of polytopes $\Omega$, and any dense subset $\tilde{S}$ of $S^{m-1}$, we have

$$F(F(\Omega)) = F(F'(\Omega)) \quad \text{and} \quad F'(F'(\Omega)) = F'(F(\Omega)), \quad (3.4.6)$$

where $F'$ is the reduced mapping associated with $\tilde{S}$, that is, $F'(\Omega) = \{\Omega(d) : d \in \tilde{S}\}$.

Proof Observe that by construction $F'(\Omega) \subset F(\Omega)$. Therefore, for any direction $d \in S^{m-1}$, we have

$$(F'(\Omega))(d) = \text{conv}\{P_d : P \in F'(\Omega)\} \subseteq \text{conv}\{P_d : P \in F(\Omega)\} = (F(\Omega))(d).$$

We will show that in fact $(F'(\Omega))(d) = (F(\Omega))(d)$ for any $d \in S^{m-1}$. Notice that, this proves both relations in (3.4.6). It only remains to demonstrate the inclusion

$$(F'(\Omega))(d) \supseteq (F(\Omega))(d). \quad (3.4.7)$$

For a fixed $d$ choose any other direction $d' \in S^{m-1}$, and let $v$ be a vertex of $(\Omega(d'))_d$. By Proposition 3.4.4, there exists a direction $d'' \in \tilde{S}$, such that $\Omega(d'' \subset \Omega(d')$ and $v \in \Omega(d'')$.

This means that $v \in (F'(\Omega))(d)$, hence, by the arbitrariness of $v$, we have (3.4.7). \[\square\]

Our next goal is to prove that different ‘paths’ of transformations starting with $\Omega_0$ yield equivalent outcomes, so it does not matter if at some intermediate steps we use $F$ or $F'$, the $i$-th application of $F$ will lead to $\Omega_i$.

Theorem 3.4.6 The following diagram commutes.

$$\begin{array}{ccccccc}
\Omega_0 & \xrightarrow{F} & \Omega_1 & \xrightarrow{F} & \Omega_2 & \xrightarrow{F} & \Omega_3 & \xrightarrow{F} & \Omega_4 & \cdots \\
\downarrow F' & & \downarrow F & & \downarrow F & & \downarrow F & & \downarrow F & \\
\Omega_1' & \xrightarrow{F'} & \Omega_2' & \xrightarrow{F'} & \Omega_3' & \xrightarrow{F'} & \Omega_4' & \cdots \\
\end{array}$$

In other words, for any $i \in \mathbb{N}$ we have $F'(\Omega_i) = \Omega_{i+1}'$ and $F(\Omega_i') = \Omega_{i+1}$.

Before we proceed with the proof, we consider an example to clarify the meaning of the commutative diagram in Theorem 3.4.6.
Example 3.4.7 Let $\Omega = \Omega_0$ be a collection of two dimensional polytopes, that consists of two opposite edges of a square, for instance,

$$\Omega_0 = \{(1, -1), (1, 1)\}, \{(-1, -1), (-1, 1)\}.$$ 

It is not difficult to verify that our commutative diagram reduces to the following chain of transformations.

Proof of Theorem 3.4.6 The set $\hat{S}^{m-1}$ is dense in $S^{m-1}$, hence, Proposition 3.4.5 yields

$$\Omega_2 = F(\Omega_1) = F(F(\Omega_0)) = F(F'((\Omega_0))) = F(\Omega'_1)$$

and

$$\Omega'_2 = F'(\Omega'_1) = F'(F'(\Omega_0)) = F'(F(\Omega_0)) = F'(\Omega_1),$$

which gives us the induction base. Assuming that for some $i \geq 2$

$$\Omega_i = F(\Omega_{i-1}) \quad \text{and} \quad \Omega'_i = F'(\Omega_{i-1}), \quad (3.4.8)$$

the relations (3.4.6) and (3.4.8) together yield

$$\Omega_{i+1} = F(F(\Omega_{i-1})) = F(F'(\Omega_{i-1})) = F(\Omega'_i)$$
and
\[ \Omega'_{i+1} = F'(F'(\Omega'_{i-1})) = F'(F(\Omega'_{i-1})) = F'(\Omega_i). \]

The desired relations follow by induction on \( i \).

\[ \square \]

### 3.5 Algebraic Reformulation and the Main Result

We have introduced some necessary components of our combinatorial formulation in the Section 3.4. We will state the main result of this section in Theorem 3.5.6 and present the proof.

We label all vertices in \( V \), and encode each direction \( d \in \hat{S}^{m-1} (\Omega) \) according to the order of the projections of the vertices on this direction. In other words, given an ordered list of vertices \( V = \{v_1, \ldots, v_k\} \) to every direction \( d \in \hat{S}^{m-1} \), we assign an element \( \tau(d) \in S_k \), where \( S_k \) is the set of orderings on the sequence \((1, 2, \ldots, k)\), that corresponds to the order of the projections of the vertices in the direction \( d \),

\[ \tau(d) = (i_1, i_2, \ldots, i_k) \in S_k \quad \text{such that} \quad (v_{i_1}, d) > (v_{i_2}, d) > \cdots > (v_{i_k}, d). \quad (3.5.1) \]

Observe that, this is well defined as we have discarded all directions for which we may encounter vertices, that project onto the same point. We can also encode each polytope \( P \in \Omega_i \) as a subset of the vertex indices, \( p \subset \{1, 2, \ldots, k\} \). Now that we have encoded our data in the discrete format, we are ready to explain the combinatorial equivalent of the conversion procedure. We first illustrate the ideas by a simple example.

**Example 3.5.1** Consider a collection \( \Omega \) that contains a line segment and a disjoint singleton in \( \mathbb{R}^2 \). We label the relevant vertices as \( A, B, \) and \( C \) as shown in Fig. 3.4 (one can think of the labelling 1, 2, 3 instead).
Choose a direction $d$, and encode it using the order of the projections of the vertices along this direction (see Fig. 3.5).

![Figure 3.4: Two sets with three vertices in $\mathbb{R}^2$](image)

It is not difficult to observe, that for our example, we obtain 6 different encodings of the reduced set of directions, $ACB$, $ABC$, $BCA$, $BAC$, $CAB$ and $CBA$. (3.5.2)

We also encode the initial collection of polytopes as $\omega_0 = \{AB, C\}$.

Now, suppose we want to construct $\Omega_1(d)$ for some direction $d$ that is encoded as $ACB$. To construct the ‘maximal faces’ for each encoded polytope in $\omega_0$, we find the vertex that appears the earliest in the sequence that encodes our direction. For the first polytope $AB$ this
is $A$, and for the second one we have the only possibility $C$, hence $\Omega_1'(d)$ corresponds to the encoded polytope $AC$. It is evident from Fig. 3.5, that this produces the same polytope as the geometric construction.

If we apply this algorithm to every direction in (3.5.2), we end up with 6 polytopes

$$\{AC, AC, BC, BC, AC, BC\}.$$ Removing the repetitions, we obtain $\omega_1 = \{AC, BC\}$.

Observe that the geometric conversion with the reduced set of directions yields exactly the same result.

We can continue this procedure using the same 6 directions with the set $\omega_1$ to obtain another 6 polytopes, which are

$$\{AC, AB, CB, AB, C, C\},$$

and we have $\omega_2 = \{AC, AB, CB, C\}$.

If we keep applying the same procedure, we get $\omega_3 = \{AC, ABC, CB\}$ and $\omega_4 = \{AC, AB, CB, AB, C\} = \{AC, AB, CB, C\}$. We have reached a cycle of length 2.

Our encoding of the set $V$ for finite family $\Omega$ of polytopes in $\mathbb{R}^m$ results in the set of directions $T = \{\tau_1, \tau_2, \ldots, \tau_r\} \subset S_k$, that correspond to the orderings of the projections of the vertices onto the directions in $\hat{S}^{m-1}$. We also encode our family $\Omega$ of polytopes $P \in \mathbb{R}^m$ as subsets $p$ of the set $\{1, 2, \ldots, k\}$ of the vertex labels. We introduce the following discrete conversion procedure.

Given a collection $w$ of nonempty subsets $p$ of $\{1, 2, \ldots, k\}$, for each ordering $\tau \in S_k$. Let $w(\tau) := \bigcup_{p \in w} \{\max(\tau)\}$, where $\max_\tau(p)$ is the maximal element in $p$ with respect to the ordering $\tau$. In other words, it is the element of $p$ that has the earliest place in the sequence $\tau$.

The discrete conversion operator $f$ is defined as

$$f(w) = \{w(\tau) : \tau \in T\}. \quad (3.5.3)$$

Note that this operator maps elements of the space $X$ of sets of nonempty subsets of $\{1, 2, \ldots, k\}$ onto the elements of the same space.
We next explicitly identify the constructions, that we have just defined for the example considered previously.

Example 3.5.2 We have the following structure based on our abstract algebraic formulation for the data given in the previous Example 3.5.1:

- \( V = \{v_1, v_2, v_3\} \) (corresponding to \( A, B \) and \( C \));
- \( w = w_0 = \{\{1, 2\}, \{3\}\} \);
- \( T = \{(1, 3, 2), (1, 2, 3), (2, 3, 1), (2, 1, 3), (3, 1, 2), (3, 2, 1)\} \);
- \( \tau_1 = \tau(d) = (1, 3, 2) \);
- \( w(\tau_1) = \bigcup_{p \in w} \{\max_{\tau_1}(p)\} = \{\max_{\tau_1}(\{1, 2\}), \max_{\tau_1}(\{3\})\} = \{1, 3\} \);
- \( w_1 = \{\{1, 3\}, \{2, 3\}\} \).

We can then continue the process to obtain \( w_2, w_3, w_4 \).

Definition 3.5.3 We say that the pair \((w, T)\), where \( w \) is a collection of subsets of \( \{1, 2, \ldots, k\} \) and \( T \subset S_k \) for some \( k \in \mathbb{N} \) is geometric if there exists a finite set of vertices \( V = \{v_1, v_2, \ldots, v_k\} \subset \mathbb{R}^m \), such that for the collection \( \Omega = \{\text{conv}_{i \in p}\{v_i\}, p \in w\} \), the set \( T \) corresponds to the encoding of the set of reduced directions \( \hat{S}^{m-1} \) as in (3.5.1).

In other words, the data \((w, T)\) can be obtained from a collection of polytopes.

We have the following combinatorial equivalent to Demyanov-Ryabova conjecture.

Conjecture 3.5.4 If a pair \((w, T)\), where \( w \) is a collection of subsets of \( \{1, 2, \ldots, k\} \) and \( T \subset S_k \) for some \( k \in \mathbb{N} \) is geometric, then the sequence \( w_1, w_2, \ldots \) obtained via the iterative application of the discrete convertor (3.5.3),

\[
    w_{i+1} = f(w_i), \quad i \in \{0, 1, 2, \ldots\}, \quad w_0 = w, \quad (3.5.4)
\]

achieves a cycle of length 2, i.e. for some \( N \in \mathbb{N} \), \( w_{n+2} = w_n \), for all \( n > N \).
It is not clear to us whether the conjecture is false, if we drop the geometric assumption. Our main goal however is to show that, it is equivalent to the Demyanov-Ryabova conjecture.

We first make the following obvious remark:

**Proposition 3.5.5** Let $\Omega$ be a finite collection of polytopes in $\mathbb{R}^n$. Let $(\Omega'_i)_i$ be the modified process starting from $\Omega$, and $(\omega_i)_i$ be the associated discrete process. Then, the sequence $(\Omega'_i)_i$ reaches a cycle of length 2 if and only if the sequence $(\omega_i)_i$ reaches a cycle of length 2.

The main challenge is to show that the modified process $(\Omega'_i)_i$ reaches a cycle of length 2 if and only if $(\Omega_i)_i$ does. The key to this is Theorem 3.4.6, which we have proved in the previous section of this chapter.

**Theorem 3.5.6** The Demyanov-Ryabova conjecture (Conjecture 3.2.2) is true if and only if Conjecture 3.5.4 is true.

We first note that the equivalence of the modified process $(\Omega'_i)_i$ to the discrete process $(\omega_i)_i$ is evident: at each iteration, $\Omega'_i$ can be reconstructed from $\omega_i$, and vice versa. We hence have the following trivial claim.

**Proof of Theorem 3.5.6** If follows immediately from Theorem 3.4.6, that the process $(\Omega'_i)_i$ reaches a cycle of length 2 if and only if $(\Omega_i)_i$ does. Since the process $(\Omega'_i)_i$ is equivalent to the discrete process $(\omega_i)_i$ by Proposition 3.5.5, we are done. □

We end this section by showing the following proposition which gives a bound on diameter in the Conjecture 3.2.2, which is valid in dimension two.

**Proposition 3.5.7** Let $n$ be the number of vertices in $\mathbb{R}^2$, then, we have no more than $n(n-1)$ different directions in $T$.

**Proof** Let $d$ be an arbitrary direction. Then, we can rotate $d$ clockwise to obtain all directions. We can encode $d$ by writing the vertex set in order of furthest along $d$ to closest along $d$. As we rotate the direction $d$ clockwise, each pair of letters swaps exactly twice. This implies that, there are $2 \times \binom{n}{2} = n(n-1)$ swaps in total, however, some swaps can occur simultaneously, so we only have an upper bound. □
Note: If there are three or more vertices collinear, or two or more pairs of collinear vertices are parallel to each other, then the number of orders for vertex set would be less than $n(n - 1)$, as some of the swaps would happen the same time as we rotate the direction around the $\mathbb{R}^2$ plane.

Therefore, the upper bound of the number of the directions is $n(n - 1)$ for general cases.

### 3.6 General convex sets

**Example 3.6.1** Consider the following special example with two circles $A$ and $B$ as the starting set of convex sets. Two circles share the common center.

Then after the transformation is applied first time, we end up with an infinite number of line segments between the outer circle and the inner circle, and if we extend all the line segments, they meet at the center of the circles.

![Diagram of circles and line segments](image)

If we apply the transformation again, we will have an infinite number of convex sets of the same shape in the collection. They rotate around the centre of the circle.
Then we can keep applying the transformation to the previous step.

Conjecture 3.6.2 Let $A$, $B$ be two convex sets in $\mathbb{R}^2$, and $A \subseteq B$. Then the conjecture is true.

Ideas: Consider the example above with two circles, we can get the idea that if we generalise two circles to two general convex sets, then keep applying the transformation, we will end up with sets with very similar structure.
Following the same process as the example above, we should have the cycle between following two collections of infinite number of convex sets.

3.7 The counterexample and future work

Subsequent to the publication of the work in this chapter in [54], the general Demyanov-Ryabova conjecture was shown to be false in [51] 2018. A counterexample consists of four polytopes was constructed in $\mathbb{R}^2$, which has the minimal cycle of length 4.

Although the general Demyanov-Ryabova conjecture is false, it is still of importance to understand the conditions under which the conjecture is true, for example, we have established in this thesis that the conjecture is true under the affine linear independence condition. Some interesting future work on this problem might be: What are the other special conditions to allow the conjecture to be true? How can these properties be expressed in terms of the classes of functions that give rise to the associated upper and lower exhausters? Does there exist a broad class of functions which admit upper and lower exhausters that possess the two cycle? What operations within a class of functions give rise to new function
that also possess two cycles for their associated exhausters? Given exhausters that admit a two cycle, can one develop a duality theory utilising this correspondence?

3.8 Conclusion

We have proved that the Demyanov-Ryabova conjecture is true assuming an affine independence condition, that is, when we restrict the number of vertices of polytopes in the collection to \( n + 1 \) affinely independent points for an \( n \) dimensional space.

We have also obtained a combinatorial reformulation of the conjecture by ordering vertices in the collection. The combinatorial formulation allows us to work on the conjecture using algebraic approaches. After we obtain the set of orderings on the vertex set that correspond to the set of restricted directions, we are able to forget about the geometry of the sets, and proceed with the equivalent algebraic version of the conjecture. This means that, we can try to apply powerful algebraic and combinatorial tools to this problem. We know that from the recent counterexample found in [51] has shown that the general conjecture is false. The combinatorial tool should advance insight for the future work on general conjecture using a purely algebraic approach.
Chapter 4

Outer Limits of Subdifferentials

4.1 Introduction

Our desire to study outer limits of subdifferentials is motivated by the problem of constructive evaluation of error bounds. The error bound modulus measures whether a given function is steep enough outside of its level set and gives a lower bound for the relevant slope. This idea stems from the works of Hoffman [28] and Łojasiewicz [39]. Error bounds are crucial for a range of stability questions, for the existence of exact penalty functions, and for the convergence of numerical methods. The literature on error bounds is vast, and we refer the reader to the following selection of recent works and classic review papers for more details [5–7,20,24,35,37,45,60]. In this work we focus on the constructive evaluation of error bound modulus for structured continuous functions.

Recall the Definition 2.2.2, we define the sublevel set

\[ S(\bar{x}) = \{ x \in X \mid f(x) \leq f(\bar{x}) \}, \]

where \( \bar{x} \in X \). We say that \( f \) has a local (linear) error bound at \( \bar{x} \) if there exists a constant \( L > 0 \) such that

\[ L \operatorname{dist}(x,S(\bar{x})) \leq \max\{0, f(x) - f(\bar{x})\} \]

(4.1.1)

for all points \( x \) in a sufficiently small neighbourhood of \( \bar{x} \). Here \( \operatorname{dist}(x,A) \) is the Euclidean distance. Taking the supremum of the constants \( L \) that satisfy (4.1.1) over all neighbourhoods of \( \bar{x} \) we arrive at an exact quantity called the error bound modulus of \( f \) at \( \bar{x} \), which
can be explicitly expressed as \( \text{Er}_f(\bar{x}) := \liminf_{x \to \bar{x}} \frac{f(x) - f(\bar{x})}{\text{dist}(x, S(\bar{x}))} \). It is possible to obtain sharp estimates of the error bound modulus \( \text{Er}_f(\bar{x}) \) for sufficiently structured functions by means of subdifferential calculus. For continuous functions that we are considering in this thesis the error bound modulus is bounded from below by the distance from zero to the outer limits of Fréchet subdifferentials (see [20]),

\[
\text{Er}_f(\bar{x}) \geq \text{dist}
\begin{pmatrix}
0, \limsup_{x \to \bar{x}} f(x) \\
\partial f(x)
\end{pmatrix},
\]

with equality holding when \( f \) is sufficiently regular (for instance convex), see [20, Theorem 5 and Proposition 10]. In [40] this equality is proved for a lower \( C^1 \) function and an additional upper estimate of the error bound of a regular locally Lipschitz function is given via the distance to the outer limits of the Fréchet subdifferentials of the subdifferential support function, and such limits are in turn expressed using the notion of the end of a closed convex set introduced in [29].

In this chapter we generalise some of the constructive results of [11] to the case of min-max type functions, providing an exact description for the outer limits of subdifferentials in the case of polyhedral functions and sharp bounds for a more general case (see Theorems 4.3.2 and 4.4.2). We also strengthen Theorem 3.2 of [11] in Corollary 4.3.7 by dropping the affine independence assumption (although the latter result can probably be obtained from the findings of [40]). Finally, we answer in the affirmative the open question of [40] for the case of functions with sublinear Hadamard directional derivative (see Corollary 4.3.5 and Remark 4.3.6).

Throughout the chapter we use the Euclidean norm, and we denote the closed unit ball and the unit sphere by \( \mathcal{B} \) and \( \mathcal{S} \) respectively.

4.2 Preliminaries

Recall that a function \( f : X \to \mathbb{R} \) is Hadamard directionally differentiable if for every \( x \in X \) and \( p \in \mathbb{R}^n \) the limit

\[
f'(x; p) = \lim_{t \downarrow 0} \frac{f(x + tp) - f(x)}{t}
\]
exists and is finite. The quantity \( f'(x; p) \) is called the (Hadamard) \textit{directional derivative} of \( f \) at \( x \) in the direction \( p \). It follows from the definition that the directional derivative is a positively homogeneous function of degree one, i.e.

\[
f'(x; \lambda p) = \lambda f'(x; p) \quad \forall x \in X, \ p \in \mathbb{R}^n, \ \lambda > 0. \tag{4.2.1}
\]

Hadamard directionally differentiable functions enjoy certain continuity properties that we summarise in the next proposition. These properties are well-known (e.g. see [17]), but we provide a proof here for convenience.

\textbf{Proposition 4.2.1} Let \( f : X \to \mathbb{R} \) be Hadamard directionally differentiable at \( \bar{x} \in X \). Then the directional derivative \( f'(%x; \cdot) \) is a continuous function; moreover,

\[
f(x + s) = f(x) + f'(x; s) + o(s), \quad \frac{o(s)}{\|s\|} \xrightarrow{s \to 0} 0. \tag{4.2.2}
\]

\textbf{Proof} We first show that the Hadamard directional derivative is continuous. Fix some \( \bar{x} \in X \), choose an arbitrary \( p \in \mathbb{R}^n \) and a sequence \( \{p_k\} \), \( p_k \to p \). By the definition of the directional derivative for every \( k \in \mathbb{N} \) there exist \( t_k \) such that \( 0 < t_k < 1/k \) and

\[
f'(\bar{x}; p_k) = \frac{f(\bar{x} + t_k p_k) - f(\bar{x})}{t_k} + \delta_k, \quad |\delta_k| < 1/k \tag{4.2.3}
\]

(notice we keep \( p_k \) fixed). Since \( t_k \downarrow 0 \) and \( p_k \to p \), we have

\[
\lim_{k \to \infty} \frac{f(\bar{x} + t_k p_k) - f(\bar{x})}{t_k} = f'(\bar{x}; p). \tag{4.2.4}
\]

Now passing to the limit on both sides of (4.2.3) and using (4.2.4), we obtain

\[
\lim_{k \to \infty} f'(\bar{x}; p_k) = \lim_{k \to \infty} \frac{f(\bar{x} + t_k p_k) - f(\bar{x})}{t_k} + \lim_{k \to \infty} \delta_k = f'(\bar{x}; p),
\]

and so the directional derivative is continuous.

It remains to show the relation (4.2.2). Assume the contrary. Then there is \( \bar{x} \in X \), a constant \( c > 0 \) and a sequence \( \{s_k\} \), \( s_k \to 0 \) such that

\[
\frac{|f(\bar{x} + s_k) - f(\bar{x}) - f'(\bar{x}; s_k)|}{\|s_k\|} > c \quad \forall k \in \mathbb{N}.
\]
Without loss of generality we can assume that \( s_k/\|s_k\| =: p_k \rightarrow p \in \mathcal{S} \), then from the continuity of \( f'(x; \cdot) \) we get

\[
0 = \left| f'(\bar{x}; p) - \lim_{k \to \infty} f'(\bar{x}; p_k) \right|
= \left| f'(\bar{x}; p) - \lim_{k \to \infty} \frac{f'(\bar{x}; s_k)}{\|s_k\|} \right| \quad \text{(using (4.2.1))}
= \lim_{k \to \infty} \left| \frac{f(\bar{x} + s_k) - f(\bar{x})}{\|s_k\|} - \lim_{k \to \infty} \frac{f'(\bar{x}; s_k)}{\|s_k\|} \right|
\geq c,
\]
which is impossible by our assumption that \( c > 0 \).

In this work our focus is on the functions with sublinear Hadamard directional derivatives and finite minima of such functions. The former are called subdifferentiable functions in [17], however this terminology is not universally accepted (e.g. in [33] subdifferentiable functions are the ones with nonempty Fréchet subdifferential). To avoid possible confusion with definitions, throughout the paper we sacrifice brevity for clarity and use the full description. Regular Lipschitz functions have sublinear Hadamard directional derivatives, see [48, Theorem 9.16], therefore all results obtained here for functions with sublinear Hadamard directional derivatives also apply to regular Lipschitz functions.

Recall that the Fréchet subdifferential of a function \( f : X \to \mathbb{R} \) at \( \bar{x} \in X \) is the set

\[
\partial f(\bar{x}) = \left\{ v \in \mathbb{R}^n \mid \liminf_{x \to \bar{x}, x \neq \bar{x}} \frac{f(x) - f(\bar{x}) - \langle v, x - \bar{x} \rangle}{\|x - \bar{x}\|} \geq 0 \right\}.
\]

For Hadamard directionally differentiable function \( f : X \to \mathbb{R} \) one has (see [33, Proposition 1.17])

\[
\partial f(\bar{x}) = \{ v \in \mathbb{R}^n \mid f'(\bar{x}; p) \geq \langle v, p \rangle \ \forall p \in \mathbb{R}^n \}.
\]

(4.2.5)

When the directional derivative is sublinear, the Fréchet subdifferential of \( f \) at \( \bar{x} \in X \) coincides with the subdifferential of the directional derivative at 0, so we have

\[
f'(\bar{x}; p) = \max_{v \in \partial f(\bar{x})} \langle v, p \rangle \ \forall p \in \mathbb{R}^n;
\]

(4.2.6)

moreover, for a convex function \( f : \mathbb{R}^n \to \mathbb{R} \) the Fréchet subdifferential coincides with the
classic Moreau-Rockafellar subdifferential,
\[
\partial f(x) = \{v \in \mathbb{R}^n \mid f(y) - f(x) \geq \langle v, y - x \rangle \forall y \in \mathbb{R}^n\} = \{v \in \mathbb{R}^n \mid f'(x;p) \geq \langle v, p \rangle \forall p \in S\}.
\]
(4.2.7)

We will be using the following result explicitly (see [26, Chap. VI, Example 3.1]).

**Proposition 4.2.2** Let \( h : \mathbb{R}^n \to \mathbb{R} \) be a sublinear function,
\[
h(x) = \max_{v \in C} \langle x, v \rangle,
\]
where \( C \subset \mathbb{R}^n \) is a compact convex set. Then
\[
\partial h(x) = \text{Arg} \max_{v \in C} \langle x, v \rangle.
\]
(4.2.8)

We will also utilise the following optimality condition (see Corollary 1.12.3 in [33]).

**Proposition 4.2.3** Let \( f_1 : X \to \mathbb{R} \) and \( f_2 : X \to \mathbb{R} \) and assume that \( f_1 \) is Fréchet differentiable at \( x \). If \( f_1 + f_2 \) attains a local minimum at \( x \), then \(-\nabla f_1(x) \subset \partial f_2(x)\).

Let \( f : X \to \mathbb{R} \) be a pointwise minimum of a finite set of functions with sublinear Hadamard directional derivatives. We have explicitly
\[
f(x) = \min_{i \in I} f_i(x) \quad \forall x \in X,
\]
(4.2.9)

where \( f_i : X \to \mathbb{R} \) are Hadamard directionally differentiable on \( X \) with sublinear directional derivatives, so that (see (4.2.6)) for every \( \bar{x} \in X \)
\[
f_i'(\bar{x};p) = \max_{v \in \partial f_i(\bar{x})} \langle v, p \rangle \quad \forall p \in \mathbb{R}^n,
\]
(4.2.10)

and \( I \) is a finite index set. Observe that the relation (4.2.2) is valid for each individual function \( f_i, i \in I \), so that we have
\[
f_i(x + s) = f_i(x) + f_i'(x;s) + o_i(s), \quad \frac{o_i(s)}{\|s\|} \xrightarrow{s \to 0} 0.
\]
(4.2.11)

Throughout this section we use the following two active index sets.
\[
I(x) = \{i \in I \mid f_i(x) = f(x)\},
\]
\[
I(x,p) = \left\{i_0 \in I(x) \left| \max_{v \in \partial f_{i_0}(x)} \langle v, p \rangle = \min_{i \in I(x)} \max_{v \in \partial f_i(x)} \langle v, p \rangle\right.\right\}.
\]

We will need the following well known relation (see [17]).
Proposition 4.2.4 Let $f : X \to \mathbb{R}$ be a pointwise minimum of a finite number of functions with sublinear Hadamard directional derivatives, as in (4.2.9). Then

$$f'(\bar{x}; p) = \min_{i \in I(\bar{x})} f'_i(\bar{x}; p) \quad \forall p \in S.$$  

**Proof** The proof follows from the definition of directional derivative. Indeed, for all $x$ in a sufficiently small neighbourhood of $\bar{x}$ we have $I(x) \subset I(\bar{x})$. Therefore

$$\lim_{t \downarrow 0} \frac{f(\bar{x} + tp') - f(\bar{x})}{t} = \lim_{t \downarrow 0} \frac{\min_{i \in I(\bar{x})} f_i(\bar{x} + tp') - f(\bar{x})}{t}$$

$$= \lim_{t \downarrow 0} \frac{\min_{i \in I(\bar{x})} [f_i(\bar{x} + tp') - f(\bar{x})]}{t}$$

$$= \min_{i \in I(\bar{x})} \lim_{t \downarrow 0} \frac{f_i(\bar{x} + tp') - f_i(\bar{x})}{t} = \min_{i \in I(\bar{x})} f'_i(\bar{x}, p).$$

The next relation is well known (see [49] for the discussion of more general calculus rules for Fréchet subdifferentials) and follows directly from the definition of the Fréchet subdifferential and the observation that the directional derivative of a pointwise minimum of a finite number of functions equals the pointwise minimum of directional derivatives of the active subset of these functions. We provide proof here for the sake of completeness.

Proposition 4.2.5 Let $f : X \to \mathbb{R}$ be a finite minimum of Hadamard directionally differentiable functions with sublinear derivatives. Then the Fréchet subdifferential of $f$ at $\bar{x} \in X$ is the intersection of the Fréchet subdifferentials of the active functions. In other words, given $f : X \to \mathbb{R}$ such that

$$f(x) = \min_{i \in I} f_i(x) \quad \forall x \in X,$$

where $I$ is a finite index set, and $f_i : X \to \mathbb{R}$ are Hadamard directionally differentiable with sublinear directional derivatives, one has

$$\partial f(\bar{x}) = \bigcap_{i \in I(\bar{x})} \partial f_i(\bar{x}) \quad \forall \bar{x} \in X,$$

where

$$I(\bar{x}) = \{i \in I \mid f(\bar{x}) = f_i(\bar{x})\}.$$
CHAPTER 4. OUTER LIMITS OF SUBDIFFERENTIALS

Proof First of all, from Proposition 4.2.4 we have
\[
f'(\bar{x}; p) = \min_{i \in I(\bar{x})} f'_i(\bar{x}; p) \quad \forall p \in S,
\]
which follows from the definition of directional derivative. Indeed, for all \(x\) in a sufficiently small neighbourhood of \(\bar{x}\) we have \(I(x) \subset I(\bar{x})\). Therefore
\[
\lim_{t \to 0} \frac{f(\bar{x} + tp') - f(\bar{x})}{t} = \lim_{t \to 0} \frac{\min_{i \in I(\bar{x})} f_i(\bar{x} + tp') - f(\bar{x})}{t} = \min_{i \in I(\bar{x})} \lim_{t \to 0} \frac{f_i(\bar{x} + tp') - f_i(\bar{x})}{t} = \min_{i \in I(\bar{x})} f'_i(\bar{x}, p).
\]
Using (4.2.5) and (4.2.12) we have
\[
\partial f(\bar{x}) = \left\{ v \in \mathbb{R}^n \mid \min_{i \in I(\bar{x})} f'_i(\bar{x}; p) \geq \langle v, p \rangle \ \forall p \in \mathbb{R}^n \right\} = \{ v \in \mathbb{R}^n \mid f'_i(\bar{x}; p) \geq \langle v, p \rangle \ \forall p \in \mathbb{R}^n \ \forall i \in I(\bar{x}) \} = \bigcap_{i \in I(\bar{x})} \partial f_i(\bar{x}).
\]
\]

Proposition 4.2.6 Let \(g : X \to \mathbb{R}\) be a pointwise maximum of a finite number of \(C^1(X)\) functions, i.e. \(g(x) = \max_{j \in J} g_j(x)\), \(g_j \in C^1(X)\), and \(|J| < \infty\). Then \(g\) has a nonempty Fréchet subdifferential that can be expressed explicitly as
\[
\partial g(x) = \text{conv} \{ \nabla g_j(x) \},
\]
where \(J(x)\) is the active index set. Moreover, the function \(g\) is Hadamard directionally differentiable with
\[
g'(x; p) = \max_{j \in J(x)} g'_j(x; p) = \max_{v \in \partial g(x)} \langle v, p \rangle = \max_{j \in J(x)} \langle \nabla g_j(x), p \rangle.
\]
Proof This result is well known, and its proof can be easily deduced from the fact that the Hadamard directional derivative is the support function of the convex hull in (4.2.13). □
CHAPTER 4. OUTER LIMITS OF SUBDIFFERENTIALS

Proposition 4.2.7 Let \( f = \min_i f_i \), where \( f_i \) are Hadamard directionally differentiable with sublinear directional derivatives on \( X \). For every \( \bar{x} \in X \) and every \( p \in S \) there exists \( \varepsilon = \varepsilon(\bar{x}, p) > 0 \) such that

\[
I(x) \subseteq I(\bar{x}, p) \quad \forall x = \bar{x} + tp + t^2 u, \quad t \in (0, \varepsilon), u \in B.
\]

Proof Suppose that the claim is not true. Then there exists a sequence \( \{t_k\}, t_k \downarrow 0 \) and a sequence \( \{u_k\}, u_k \in B \) such that for

\[
x_k = \bar{x} + t_k p + t_k^2 u_k
\]

we have \( I(x_k) \setminus I(\bar{x}, p) \neq \emptyset \). Without loss of generality assume that there is an \( i_0 \in I \) such that \( i_0 \in I(x_k) \setminus I(\bar{x}, p) \). Observe that

\[
f_{i_0}(x_k) \leq f(x_k),
\]

hence, by the continuity of \( f \), \( i_0 \in I(\bar{x}) \); moreover, observing that \( p + t_k u_k \rightarrow p \), we have

\[
f'_{i_0}(\bar{x}; p) = \lim_{k \rightarrow \infty} \frac{f_{i_0}(x_k) - f_{i_0}(\bar{x})}{t_k} \leq \lim_{k \rightarrow \infty} \frac{f(x_k) - f(\bar{x})}{t_k} = f'(\bar{x}; p).
\]

We then have from Proposition 4.2.4

\[
\max_{v \in \partial f_{i_0}(\bar{x})} \langle v, p \rangle = f'_{i_0}(\bar{x}; p) \leq f'(\bar{x}; p) = \min_{i \in I(\bar{x})} f'_{i}(\bar{x}; p) = \min_{i \in I(\bar{x})} \max_{v \in \partial f_i(\bar{x})} \langle v, p \rangle,
\]

hence, \( i_0 \in I(\bar{x}, p) \), which contradicts our assumption. \( \Box \)

4.3 Limiting subdifferential for pointwise minima

Our results rely on the following technical lemma, whose proof is inspired by the proofs of fuzzy mean value theorems for Fréchet subdifferential (see [31,42]). To show the existence of a nearby point with a desired subgradient, an auxiliary function is constructed which attains a local minimum at such point.

Lemma 4.3.1 Let \( f : X \rightarrow \mathbb{R} \) be a pointwise minimum of finitely many functions with sublinear Hadamard directional derivatives, as in (4.2.9). Then for every \( \bar{x} \in X \), \( p \in S \) and

\[
y \in \bigcap_{i \in I(\bar{x}, p)} \text{Arg} \max_{v \in \partial f_i(\bar{x})} \langle v, p \rangle \tag{4.3.1}
\]
there exist sequences \( \{x_k\} \) and \( \{y_k\} \) such that

\[
x_k \to_{k \to \infty} \bar{x}, \quad \frac{x_k - \bar{x}}{\|x_k - \bar{x}\|} \to_{k \to \infty} p, \quad y_k \in \partial f(x_k), \quad y_k \to_{k \to \infty} y.
\]

**Proof** Fix \( \bar{x} \in X, p \in S \) and \( y \) such that

\[
y \in \bigcap_{i \in I(\bar{x}, p)} \text{Arg} \max_{v \in \partial f_i(\bar{x})} \langle v, p \rangle.
\]

Observe that by the relation (4.2.10) and by the positive homogeneity of the directional derivative (4.2.1) we have for all \( i \in I(\bar{x}, p) \)

\[
f'_i(\bar{x}; \lambda p) = \lambda f'_i(\bar{x}; p) = \lambda \max_{v \in \partial f_i(\bar{x})} \langle v, p \rangle = \langle y, \lambda p \rangle \quad \forall \lambda > 0.
\]

(4.3.2)

For any \( \lambda > 0 \) define the function \( \varphi_\lambda : X \to \mathbb{R} \) as follows

\[
\varphi_\lambda(x) = f(x) - f(\bar{x}) - \langle y, x - \bar{x} \rangle + \frac{1}{\lambda} \|x - (\bar{x} + \lambda p)\|^2.
\]

We will show that for sufficiently small \( \lambda \) a minimum of the function \( \varphi_\lambda \) on the ball \( \bar{x} + \lambda p + \lambda \varepsilon B \) is attained at an interior point (here \( \varepsilon \in (0, \min(\varepsilon(\bar{x}, p), 1)) \), where \( \varepsilon(\bar{x}, p) \) comes from Proposition 4.2.7). Note here that since \( X \) is an open set, there exists \( r \in (0, 1) \) such that \( \bar{x} + rB \subset X \). If \( \lambda \) is smaller than \( r/2 \), then \( \bar{x} + \lambda p + \lambda \varepsilon B \subset \bar{x} + rB \subset X \). We will assume that our \( \lambda \) is always chosen small enough to satisfy this condition, and also that \( \lambda < \varepsilon(\bar{x}, p) \) (see Proposition 4.2.7).

Observe that the function \( \varphi_\lambda \) is continuous, and hence it attains its minimum on the ball \( \bar{x} + \lambda p + \lambda \varepsilon B \). Assume that contrary to what we want to prove, there exist \( \lambda_k \downarrow 0, u_k \in S \) such that

\[
\bar{x} + \lambda_k p + \lambda_k \varepsilon u_k \in \text{Arg} \min_{x \in \bar{x} + \lambda_k p + \lambda_k \varepsilon B} \varphi_\lambda_k(x).
\]

We therefore have

\[
\varphi_\lambda_k(\bar{x} + \lambda_k p + \lambda_k \varepsilon u_k) \leq \varphi_\lambda_k(\bar{x} + \lambda_k p),
\]

or explicitly

\[
\min_{i \in I} f_i(\bar{x} + \lambda_k p + \lambda_k \varepsilon u_k) - f(\bar{x}) - \langle y, \lambda_k p + \lambda_k \varepsilon u_k \rangle + \frac{1}{\lambda_k} \|\varepsilon \lambda_k u_k\|^2
\]

\[
\leq \min_{i \in I} f_i(\bar{x} + \lambda_k p) - f(\bar{x}) - \langle y, \lambda_k p \rangle.
\]

(4.3.3)
By our choice of $\varepsilon$ and $\lambda$, Proposition 4.2.7 yields that

$$I(\bar{x} + \lambda k p + \lambda k \varepsilon u_k) \subseteq I(\bar{x}, p) \quad \forall k \in \mathbb{N}.$$ 

Without loss of generality, due to the finiteness of $I(\bar{x}, p)$, we can assume that the index set is constant, i.e.

$$I(\bar{x} + \lambda k p + \lambda k \varepsilon u_k) = \tilde{I}.$$ 

which together with (4.3.3) yields

$$f_i(\bar{x} + \lambda k p + \lambda k \varepsilon u_k) - f_i(\bar{x}) - \langle y, \lambda k p \rangle - f_i(\bar{x} + \lambda k p + \lambda k \varepsilon u_k; \lambda k p + \lambda k \varepsilon u_k) \leq f_i(\bar{x} + \lambda k p) - f_i(\bar{x}) - \langle y, \lambda k p \rangle \quad \forall i \in \tilde{I}. \quad (4.3.4)$$

Notice that by our choice of $y$ we have from the relations (4.2.10) and (4.3.2) for all $i \in \tilde{I} \subset I(\bar{x})$

$$\langle y, \lambda k p \rangle \leq \max_{v \in \partial f_i(\bar{x})} \langle v, \lambda k p + \lambda k \varepsilon u_k \rangle \leq f_i'(\bar{x}; \lambda k p + \lambda k \varepsilon u_k) \quad (4.3.5)$$

and

$$\langle y, \lambda k p \rangle = \max_{v \in \partial f_i(\bar{x})} \langle v, \lambda k p \rangle = f_i'(\bar{x}; \lambda k p) \quad (4.3.6)$$

Noticing that $\|u_k\| = 1$, substituting (4.3.5) and (4.3.6) into (4.3.4), dividing the whole expression by $\lambda_k$, we obtain for every $i \in \tilde{I}$

$$\varepsilon^2 \leq \frac{f_i(\bar{x} + \lambda k p) - f_i(\bar{x}) - f_i'(\bar{x}; \lambda k p)}{\lambda_k} - \frac{f_i(\bar{x} + \lambda k p + \lambda k \varepsilon u_k) - f_i(\bar{x}) - f_i'(\bar{x}; \lambda k p + \lambda k \varepsilon u_k)}{\lambda_k} = o_i(\lambda k p) - o_i(\lambda k p + \lambda k \varepsilon u_k) = o_i(\lambda k p) - \frac{\lambda k}{\lambda_k} - \frac{o_i(\lambda k p + \lambda k \varepsilon u_k)}{\|\lambda k p + \lambda k \varepsilon u_k\|} \|p + \varepsilon u_k\|$$

where $o_i(\cdot)$'s are as in (4.2.11). It is not difficult to see that the right hand side goes to zero as $k \to \infty$, and hence $\varepsilon^2 = 0$, which contradicts our choice of a fixed positive $\varepsilon$.

We have shown that our assumption is wrong, and given a fixed $\varepsilon > 0$ for sufficiently small $\lambda(\varepsilon)$ the function $\varphi_\lambda$ has a local minimum in the interior of the ball $\bar{x} + \lambda p + \varepsilon \lambda B$ for all $\lambda \in (0, \lambda(\varepsilon))$; in other words, it attains an unconstrained local minimum at this point.

Let $\{\varepsilon_k\}$ be such that $\varepsilon_k \downarrow 0$, and choose

$$\lambda_k = \min\{\varepsilon_k, \lambda(\varepsilon_k)\} \quad \forall k \in \mathbb{N}.$$
For each $k \in \mathbb{N}$ there exists a point $u_k$ in the interior of $B$ such that $\bar{x} + \lambda_k p + \varepsilon_k \lambda_k u_k$ is a minimum of the function $\varphi_{\lambda}$ on the ball $\bar{x} + \lambda_k p + \varepsilon_k \lambda_k B$. From the optimality condition in Proposition 4.2.3 we have

$$y_k := y - 2\varepsilon_k u_k \in \partial f(\bar{x} + \lambda_k p + \lambda_k \varepsilon_k u_k).$$

Observe that $y_k \to y$ and for $x_k := \bar{x} + \lambda_k p + \lambda_k \varepsilon_k u_k$ we have $x_k \to \bar{x}$ and

$$\frac{x_k - \bar{x}}{\|x_k - \bar{x}\|} = \frac{p + \varepsilon_k u_k}{\|p + \varepsilon_k u_k\|} \to p, \quad k \to \infty,$$

so we are done. □

We use Lemma 4.3.1 to obtain an inclusion relation for the outer limits of subdifferentials.

**Theorem 4.3.2** Let $f : X \to \mathbb{R}$ be as in (4.2.9). Then

$$\bigcup_{f'(\bar{x}; p) > 0} \bigcap_{i \in I(x, p)} \operatorname{Arg max}_{v \in \partial f_i(\bar{x})} \langle v, p \rangle \subseteq \limsup_{x \to \bar{x}} \partial f(x), \quad (4.3.7)$$

where $I(x, p)$ is the index set as defined earlier.

**Proof** Fix $\bar{x} \in X$, choose any direction $p \in S$ such that $f'(\bar{x}; p) > 0$. We will show that for

$$y \in \bigcap_{i \in I(\bar{x}, p)} \operatorname{Arg max}_{v \in \partial f_i(\bar{x})} \langle v, p \rangle$$

we have

$$y \in \limsup_{x \to \bar{x}} \partial f(x).$$

By Lemma 4.3.1 there exist sequences $\{x_k\}$ and $\{y_k\}$ such that $y_k \in \partial f(x_k)$, $y_k \to y$, $x_k \to \bar{x}$ and

$$p_k := \frac{x_k - \bar{x}}{\|x_k - \bar{x}\|} \to p.$$

It remains to show that for sufficiently large $k$ we have $f(x_k) > f(\bar{x})$. Assume this is not so. Then without loss of generality, $f(x_k) \leq f(\bar{x})$ for all $k \in \mathbb{N}$, and

$$f'(\bar{x}; p) = \lim_{k \to \infty} \frac{f(\bar{x} + \|x_k - \bar{x}\| p_k) - f(\bar{x})}{\|x_k - \bar{x}\|} = \lim_{k \to \infty} \frac{f(x_k) - f(\bar{x})}{\|x_k - \bar{x}\|} \leq 0,$$

which contradicts our choice of $p$. □
 CHAPTER 4. OUTER LIMITS OF SUBDIFFERENTIALS

We would like to point out that the result of the computation of the expression on the left hand side of (4.3.7) depends on the position of zero with respect to the subdifferential. Also evidently the left hand side expression is not necessarily closed. The following example illustrates both facts.

Example 4.3.3

\[ \phi(x) = \begin{cases} 
0, & \text{if } x \leq 0 \\
2^{-n}, & 2^{-n} \leq x \leq 2^{-n} \text{ n odd} \\
3x - 2^{-n}, & 2^{-n} \leq x \leq 2^{-n} \text{ n even} \\
x, & \text{otherwise}
\end{cases} \]

Note:

- \( 3(2^{-n}) - 2^{-n} = 2^{-n}(3 - 1) = 2^{-n} \cdot 2 = 2^{-(n-1)} \)
- \( 3(2^{-n-1}) - 2^{-n} = 2^{-n-1}(3 - 2) = 2^{-n-1} \)

Differentiate the function \( \phi \), we have,

\[ \nabla \phi(x) = \begin{cases} 
0, & \text{if } x \leq 0 \\
0, & \text{if } 2^{-n-1} \leq x \leq 2^{-n} \text{ n odd} \\
3, & \text{if } 2^{-n-1} \leq x \leq 2^{-n} \text{ n even}
\end{cases} \]

\( \partial \phi(0) = [0, 1] \) and \( \partial^> \phi(0) = [0, 3] \)

\( \text{end} \phi(0) = \{3\} \) and \( \text{ebm}(0) = 1 \)

\[ 0 = d(0, \partial^> \phi(0)) \leq \text{ebm}(0) = \text{Er} \phi(0) = 1 < d(0, \partial^> \sigma_{\phi(0)}(0)) = 3 \]

where \( \sigma \) is the support function.

Example 4.3.4 Consider the two functions

\[ f_1(x, y) = \sqrt{x^2 + y^2} + \frac{1}{2} x, \quad f_2(x, y) = \sqrt{x^2 + y^2} - \frac{1}{2} x. \]
Figure 4.1: On the left: the functions $f_1$ and $f_2$; on the right: $f(x,y) = \min\{f_1(x,y), f_2(x,y)\}$.

Notice that $f = \min\{f_1, f_2\}$ is nonnegative everywhere, and $f(x,y) = 0$ iff $(x,y) = 0$ (see the Mathematica plots in Fig. 4.1).

It is not difficult to observe that the subdifferentials of $f_1$ and $f_2$ at zero are unit disks centred at $(\frac{1}{2}, 0)$ and $(-\frac{1}{2}, 0)$ respectively (see Example 3.2 in [49] for detailed explanation). The left-hand side in (4.3.7) for the function $f$ at 0 is the union of two (open) semi-circles, see Fig. 4.2. Observe that the closure of this set coincides with the outer limit on the right hand side of (4.3.7).

We next modify this example by translating the subdifferentials and obtaining a different set on the left hand side of (4.3.7).

Consider the modified functions

$$\tilde{f}_1(x,y) = \sqrt{x^2 + y^2} + \frac{3}{2} x, \quad \tilde{f}_2(x,y) = \sqrt{x^2 + y^2} + \frac{1}{2} x.$$

The minimum function $f(x,y) = \min\{\tilde{f}_1(x,y), \tilde{f}_2(x,y)\}$ is no longer nonnegative (see Fig. 4.3. Similar to the previous example, the subdifferentials of $f_1$ and $f_2$ at zero are unit disks centred at $(\frac{3}{2}, 0)$ and $(\frac{1}{2}, 0)$ respectively. The left-hand side in (4.3.7) is the union of one semi-circle and two smaller segments of the other circle, see Fig. 4.4.

We have the following useful special case of Theorem 4.3.2.
Corollary 4.3.5 Let \( f : X \to \mathbb{R} \) be Hadamard directionally differentiable at every point \( x \in X \), assume that the directional derivative \( f'(\bar{x}; \cdot) \) is a sublinear function for every fixed \( \bar{x} \in X \), then

\[
\bigcup_{\substack{v \in \partial f(\bar{x}) \\ f'(\bar{x}; p) > 0}} \operatorname{Arg \ max}_{p \in \partial f} \langle v, p \rangle \subset \limsup_{x \to \bar{x}} \partial f(x).
\]  

(4.3.8)

Remark 4.3.6 Observe that in the notation of [40] the closure of the union on the left hand side of (4.3.8) coincides with the outer limit of the Fréchet subdifferentials of the support of \( \partial f(\bar{x}) \). Hence we answer affirmatively the open question of [40] on whether such outer limit is a subset of the right hand side of (4.3.8).

In Corollary 4.3.7 we strengthen Theorem 3.2 of [11], dropping the affine independence assumption. We first recall the notation from [11] and the related geometric constructions.

Let \( g : X \to \mathbb{R} \) be a pointwise maximum of smooth functions, i.e.

\[
g(x) = \max_{j \in J} g_j(x), \quad g_j \in C^1(X) \quad \forall j \in J,
\]

where \( J \) is a finite index set. Define as usual the active index set

\[
J(x) = \{ j \in J \mid g(x) = g_j(x) \},
\]
Figure 4.3: On the left: the functions \( \tilde{f}_1 \) and \( \tilde{f}_2 \); in the middle: \( \tilde{f}(x, y) = \min\{\tilde{f}_1(x, y), \tilde{f}_2(x, y)\} \); on the right: the plot of \( \tilde{f} \) is shown together with \( \{z = 0\} \).

and following [11] define the collection \( \mathcal{D}(\bar{x}) \) of index subsets \( D \subset J(\bar{x}) \) such that the following system is consistent with respect to \( d \)

\[
\begin{cases}
\langle \nabla g_j(\bar{x}), d \rangle = 1, & j \in D, \\
\langle \nabla g_j(\bar{x}), d \rangle < 1, & j \in J(\bar{x}) \setminus D.
\end{cases}
\]

Corollary 4.3.7 Let \( g(x) \) to be the pointwise maximum of smooth functions as defined above. Then

\[
\bigcup_{D \in \mathcal{D}(\bar{x})} \text{conv} \{\nabla g_{j}(\bar{x}), j \in D\} = \bigcup_{p \in S, g'(\bar{x}, p) > 0} \text{Arg max}_{v \in \partial g(\bar{x})} \langle v, p \rangle \subseteq \limsup_{x \to \bar{x}} \partial g(x),
\]

in other words, in [?, Theorem 3.2] the subsets \( D_{AI}(\bar{x}) \) can be replaced by \( D(\bar{x}) \).

Moreover when all \( \{g_j\}_{j \in J} \) are affine we have an identity (instead of an inclusion) in (4.3.9).

Proof We begin by showing the following identity:

\[
\bigcup_{D \in \mathcal{D}(\bar{x})} \text{conv} \{\nabla g_{i}(\bar{x}), i \in D\} = \bigcup_{p \in S, g'(\bar{x}, p) > 0} \text{Arg max}_{v \in \partial g(\bar{x})} \langle v, p \rangle.
\]

Observe that explicitly the equality below holds for all \( p \) (see Proposition 4.2.6)

\[
\partial g(x) = \text{conv} \{\nabla g_i(x) \mid i \in J(x)\}, \quad g'(x; p) = \max_{v \in \partial g(x)} \langle p, v \rangle,
\]
Figure 4.4: On the left: the Fréchet subdifferentials at zero, $\partial f_1(0)$ and $\partial f_2(0)$; on the right: the left hand side union in (4.3.7) is shown in bold solid lines.

hence, for each direction $p$ that features in the union on the right hand side of (4.3.10) the relevant $\text{Arg max}$ gives the support face of the subdifferential

$$\partial g(\bar{x}) = \text{conv}\{\nabla g_i(\bar{x}) \mid i \in J(x)\}. $$

Explicitly, fix $p \in S$ and let

$$s(p) := g'(\bar{x}; p) = \max_{v \in \partial g(\bar{x})} \langle v, p \rangle = \max_{j \in J(\bar{x})} \langle \nabla g_j(\bar{x}), p \rangle,$$

then

$$\bigcup_{p \in S} \text{Arg max} \langle v, p \rangle = \bigcup_{p \in \text{conv}_{j \in J(\bar{x})} \{\nabla g_j(\bar{x})\}} \text{Arg max} \langle v, p \rangle.$$

We now get back to the definition of our index subsets $D$. The system

$$\begin{cases} 
\langle \nabla g_j(\bar{x}), d \rangle = 1, & j \in D, \\
\langle \nabla g_j(\bar{x}), d \rangle < 1, & j \in J(\bar{x}) \setminus D.
\end{cases}$$

is consistent for some nonempty $D \subset J(\bar{x})$ and $d \in \mathbb{R}^n$ if and only if for $p = d/\|d\|

$$g'(\bar{x}; p) = \frac{1}{\|d\|} g'(\bar{x}; d) = \frac{1}{\|d\|} \max_{j \in J(\bar{x})} \langle \nabla g_j(\bar{x}), d \rangle = \frac{1}{\|d\|} > 0,$$
and
\[ \text{Arg max}_{v \in \text{conv}\{\nabla g_i(\bar{x}) \mid i \in J(\bar{x})\}} \langle v, p \rangle = \{ \nabla g_i(\bar{x}) \mid i \in D \}. \]
hence we get (4.3.10). The last inclusion of (4.3.9) follows from Corollary 4.3.5.

Finally, to show that in the affine case an equality holds in (4.3.10), observe that there is a sufficiently small neighbourhood \( N(\bar{x}) \) of \( \bar{x} \) on which the affine function \( g \) coincides with the sum \( g(\bar{x}) + \sigma \partial g(\bar{x})(x - \bar{x}) \), where \( \sigma \partial g(\bar{x})(\cdot) \) is the support of the subdifferential, and hence for any \( x \) in this neighbourhood we have \( \partial g(x) = \partial \sigma \partial g(\bar{x})(x - \bar{x}) \). Since the number of different subdifferentials of points in this neighbourhood is finite, the right hand side is in fact the union
\[ \limsup_{g(x) > g(\bar{x})} \partial g(x) = \bigcup_{x \in N(\bar{x}) : g(x) > g(\bar{x})} \partial \sigma \partial g(\bar{x})(x) = \bigcup_{v \in \partial g(\bar{x})} \text{Arg max}_{v \in \partial g(\bar{x})} \langle v, p \rangle, \]
where the last equality follows from Proposition 4.2.2. \( \square \)

In the expression (4.3.10), the \( \text{Arg max} \) construction gives the support faces of the subdifferential, while the positivity constraint on the directional derivative means that the union of the support faces that has zero strictly on the same side as the whole subdifferential. A geometric interpretation of this construction is shown in Fig. 4.5.

**Example 4.3.8** Consider the function
\[ \phi(x) = \max\{f_1(x), f_2(x)\}, \]
where
- \( f_1(x) = x_1^2 + x_2^2 + \frac{1}{2}(x_1 + x_2) = (x_1 + \frac{1}{4})^2 + (x_2 + \frac{1}{4})^2 - \frac{1}{8} \)
- \( f_1 > 0 \) if and only if it is outside of the circle \( (x_1 + \frac{1}{4})^2 + (x_2 + \frac{1}{4})^2 = \frac{1}{8} \)
- \( f_2(x) = x_1 + x_2 \)
- \( f_2 > 0 \) on right hand side of the line \( x_1 + x_2 = 0 \)
Consider that $f_1 < f_2$ if and only if

$$x_1^2 + x_2^2 + \frac{1}{2}(x_1 + x_2) < x_1 + x_2$$

Rearranging the expression, we get

$$x_1^2 + x_2^2 - \frac{1}{2} < 0$$

Then complete the square,

$$(x_1 - \frac{1}{4})^2 + (x_2 - \frac{1}{4})^2 - \frac{1}{8} < 0$$

Which means $f_1 < f_2$ inside of the circle

$$(x_1 - \frac{1}{4})^2 + (x_2 - \frac{1}{4})^2 = \frac{1}{8}$$

When $f = f_1$, we have

$$\nabla f = (2x_1 + \frac{1}{2}, 2x_2 + \frac{1}{2}) \rightarrow (\frac{1}{2}, \frac{1}{2}) \text{ as } x \rightarrow 0$$
When $f = f_2$, we have $\nabla f = (1, 1)$.

Therefore,
\[
\partial f(0) = \text{conv}\{\left(\frac{1}{2}, \frac{1}{2}\right), (1, 1)\}
\]
\[
\partial^> f(0) = \text{conv}\{\left(\frac{1}{2}, \frac{1}{2}\right), (1, 1)\}
\]

Also, $f'(x, d) > 0$ if and only if $d_1 + d_2 > 0$.

Therefore,
\[
\bigcup_{d \in \mathcal{S}, f'(0, d) > 0} \arg\max_{v \in \partial f(0)} \langle v, d \rangle = \{(1, 1)\}
\]
\[
d(0, \partial^< f(0)) = \frac{\sqrt{2}}{2}
\]
\[
= \frac{1}{\sqrt{2}}
\]
\[
= \text{Er } f(0)
\]
\[
< d(0, \partial^< \sigma_{\partial f(0)}(0))
\]
\[
= \sqrt{2}
\]

### 4.4 Exact representations for piecewise affine functions

We are now ready to generalise Theorem 3.1 from [11]. We first prove that for positively homogeneous functions the inclusion (4.3.7) can be replaced by an equality.

**Lemma 4.4.1** Let $h : \mathbb{R}^n \to \mathbb{R}$ be a pointwise minimum of a finite number of sublinear functions, i.e.
\[
h(x) = \min_{i \in I} h_i(x), \quad h_i(x) = \max_{v \in C_i} \langle v, x \rangle \quad \forall i \in I,
\]
where $C_i$ is a compact convex set for each $i \in I$. Then
\[
\text{cl } \bigcup_{v \in \mathcal{S} \atop h(v) > 0} \bigcap_{i \in I(x)} \arg\max_{v \in C_i} \langle v, x \rangle = \limsup_{x \to 0} \partial h(x), \quad (4.4.1)
\]
where $I(x)$ is the active index set.
Proof Observe that the inclusion \( \subseteq \) in (4.4.1) follows directly from Theorem 4.3.2 substituting \( \bar{x} = 0 \), observing that \( h'(0; p) = h(p) \), \( h(0) = 0 \) and that the right hand side is a closed set. It remains to show the reverse inclusion. Choose any \( y \in \limsup_{x \to 0 \atop h(x) > 0} \partial h(x) \).

There exist sequences \( \{x_k\} \) and \( \{y_k\} \) such that \( x_k \to 0 \), \( y_k \to y \) and \( y_k \in \partial h(x_k) \). We have by Proposition 4.2.5

\[
\partial h(x_k) = \bigcap_{i \in I(x_k)} \partial h_i(x_k);
\]

furthermore, Proposition 4.2.2 yields

\[
\partial h_i(x_k) = \text{Arg max}_{v \in C_i} \langle v, x_k \rangle,
\]

and hence

\[
y_k \in \partial h(x_k) = \bigcap_{i \in I(x_k)} \text{Arg max}_{v \in C_i} \langle v, x_k \rangle,
\]

and so we have shown that \( y \) indeed belongs to the left hand side of (4.4.1). \( \square \)

We are now ready to obtain a generalisation of Theorem 3.2 in [11].

Theorem 4.4.2 Let \( f : X \to \mathbb{R} \) be as in (4.2.9), and in addition assume that for every \( i \in I \) the function \( f_i \) is piecewise affine, i.e.

\[
f_i(x) = \max_{j \in J_i} (a_{ij}, x) + b_{ij} \quad \forall i \in I,
\]

where \( J_i \)'s are finite index sets for each \( i \in I \). Then

\[
\bigcup_{p \in S} \bigcap_{f'(\bar{x}; p) > 0} \text{Arg max}_{v \in \partial f_i(\bar{x})} \langle v, p \rangle = \limsup_{x \to \bar{x} \atop f(x) > f(\bar{x})} \partial f(x), \tag{4.4.2}
\]

where, as before,

\[
I(x, p) = \left\{ i_0 \in I(x) \mid \max_{v \in \partial f_{i_0}(x)} \langle v, p \rangle = \min_{i \in I(x)} \max_{v \in \partial f_i(x)} \langle v, p \rangle \right\}.
\]
Proof Observe that when the sets $C_i$ in Lemma 4.4.1 are polyhedral, there is no need for the closure operation in (4.4.1), since there are finitely many different faces of each subdifferential, and we therefore have a finite union of closed convex sets which is always closed.

To finish the proof it remains to note that convex polyhedral functions are locally positively homogeneous and coincide with their first order approximations in a sufficiently small neighbourhood of each point. Since outer limits of subdifferentials are local notions, it is clear that the application of Lemma 4.4.1 to the directional derivatives of the active functions yields the required result. □

The diagram in Fig. 4.6 gives a geometric demonstration of constructing the union of the minimal support faces as in Theorem 4.4.2.

Figure 4.6: The geometric construction of the active faces for the min-max function

Note that in the case of a min-max type function the piecewise affine assumption is essential for the equality in (4.4.2) to hold. Consider the following classic semialgebraic
Example 4.4.3 Let $f = \min\{f_1, f_2\}$, where

$$f_1(x, y) = 1 - \frac{(x - 2)^2 + y^2}{4}, \quad f_2(x, y) = -1 + ((x - 1)^2 + y^2).$$

At the point $(x, y) = 0_2 = (0, 0)$ we have

$$f'(0_2, l) = \min\{f'_1(0_2, l), f'_2(0_2, l)\} = \min\{\langle \nabla f_1(0_2), l \rangle, \langle \nabla f_2(0_2), l \rangle\} = \min\{l_x, -2l_x\} \leq 0 \quad \forall l,$$

therefore,

$$\{l \in \mathbb{R}^2 \mid f'(0_2; l) > 0\} = \emptyset,$$

and so the expression on the left hand side of (4.3.7) produces the empty set.

We now compute the outer limit directly. We have for the Fréchet subdifferential

$$\partial f(x) = \begin{cases} \nabla f_1(x), & f_1(x) < f_2(x), \\ \nabla f_2(x), & f_1(x) > f_2(x), \\ \emptyset, & f_1(x) = f_2(x), \nabla f_1(x) \neq \nabla f_2(x), \\ \nabla f_1(x), & f_1(x) = f_2(x), \nabla f_1(x) = \nabla f_2(x). \end{cases}$$
Substituting the values and the gradients we have

\[ \partial f(x) = \begin{cases} 
(1 - \frac{x}{2}, \frac{y}{2})^T, & 12x < 5(x^2 + y^2) \\
(2x - 2, 2y)^T, & 12x > 5(x^2 + y^2) \\
\emptyset, & 12x = 5(x^2 + y^2).
\end{cases} \]

It is not difficult to observe that

\[ \limsup_{(x,y) \to 02} \partial f(x,y) \subseteq \limsup_{f(x,y) \to 02} \partial f(x,y) = \{(1,0)^T, (-2,0)^T\}. \]

On the other hand, observe that we can construct sequences of points converging to zero along the curves in the regions that correspond to \( f_2 > f_1 > 0 \) and \( 0 < f_2 < f_1 \) respectively. This works for points on the two curves

\[ 3x - x^2 - y^2 = 0 \quad 11x - 5x^2 - 5y^2 = 0 \]

shown in dashed and dotted lines in the last plot of Fig. 4.8.

Observe that in the case of a min-max function it is impossible to obtain a complete characterisation of the outer limiting subdifferential in terms of the individual gradients.
CHAPTER 4. OUTER LIMITS OF SUBDIFFERENTIALS

and related indices: the intersections of support faces of the polyhedra that represent the subdifferentials may not be representable as convex hulls of gradients that feature in the representation of the functions.

4.5 Conclusion

In this chapter, we have generalised the outer limiting subdifferential construction by Cánovas, Henrion, López and Parra for max type functions to pointwise minima of regular Lipschitz functions. The main result in [11] concerned about the outer limits of subdifferentials for max-functions, and an upper bound for the calmness modulus of nonlinear systems was obtained, apart from that, the lower bound on this modulus was also obtained for the convex case. For our result in this chapter, we were able to provide an exact description for the outer limits of subdifferentials in the case of polyhedral functions, and sharp bounds for a more general case was obtained. The affine independence assumption was also relaxed in our new results compare to in [11].
Chapter 5

Facial Structure for Convex Sets

While faces of a polytope form a well structured lattice, in which faces of each possible dimension are present, this is not true for general compact convex sets. We address the question of what dimensional patterns are possible for the faces of general closed convex sets. We show that for any finite sequence of positive integers there exist compact convex sets which only have extreme points and faces with dimensions from this prescribed sequence. We also discuss another approach to dimensionality, considering the dimension of the union of all faces of the same dimension. We show that the questions arising from this approach are highly nontrivial and give examples of convex sets for which the sets of extreme points have fractal dimension.

5.1 Introduction

It is well known that faces of polyhedral sets have a well-defined structure (see [61, Chap. 2]). In particular, every face of a polyhedral set is a polyhedron, and there are no ‘gaps’ in the dimensions of their faces. On the other hand, a simple reformulation of [25, Corollary 3.7] asserts that in the compact convex set of all positive semidefinite $n \times n$ matrices with trace 1, every proper face has dimension $\frac{k(k+1)}{2} - 1$ for some $k < n$. Thus there are naturally occurring examples with serious gaps in the dimensions of their faces. For other descriptions of this phenomenon, see Theorem 2.25 and the explanation that follows it in [57] (for the cone $\mathbb{S}_+^n$ of positive semidefinite $n \times n$ matrices), or [4, Theorem 5.36] (for the state space of...
a $C^*$-algebra). This raises the question, what are the possible patterns for the dimensions of faces of compact convex sets?

Recall that a face $F$ of a closed convex set $C \subset \mathbb{R}^n$ is a closed convex subset of $C$ such that for any point $x \in F$ and for any line segment $[a, b] \subset C$ such that $x \in (a, b)$, we have $a, b \in F$. The fact that $F$ is a face of $C$ is expressed as $F \preceq C$.

The difference between this definition and the definition of faces of polyhedral sets as intersections with supporting hyperplanes is due to the fact that for nonpolyhedral convex sets faces are not necessarily exposed: it may happen that a face cannot be represented as the intersection of a supporting hyperplane with the set. Some classic examples are shown in Figs. 5.1 (see [47]) and 5.2 (see [46]).

Figure 5.1: The convex hull of a torus is not facially exposed (the dashed line shows the unexposed extreme points).

Figure 5.2: An example of a two dimensional set and a three dimensional cone that have an unexposed face.

The dimension of a convex set is the dimension of its affine hull. We refer the reader to the classic textbooks [27, 47]. We also would like to mention that some problems related to dimensions of convex sets were studied in the literature. For instance, [19] focusses on the
dimensions of convex sets coming from optimisation problems with inequality constraints, and [22] deals with the results related to the dimensions of intersections of convex sets. However, we were unable to identify references that would address the existence of convex sets with prescribed facial dimensions.

The total number of possible face patterns in $n$ dimensional space is the cardinality of the powerset of $n$ elements. This is because every set contains zero-dimensional faces (because of the Krein-Milman theorem). We can write down face patterns either as an increasing sequence of positive numbers $(d_1, d_2, \ldots, d_k)$, which encode all possible dimensions of faces of positive dimension present in a set, or as a binary sequence $(b_1, b_2, \ldots, b_n)$, where $b_i = 1$ if a face of dimension $i$ is present in the set, and $b_i = 0$ otherwise. For example, the dimensional pattern of a tetrahedron is either $(1, 2, 3)$ in the $d$-notation or $(1, 1, 1)$ in the binary notation, and the pattern of a closed Euclidean ball is either $(n)$ or $(0, 0, \ldots, 1)$, as it does not have any faces except for zero- and $n$-dimensional ones. We will use the first encoding style via an increasing sequence of positive numbers in what follows.

The easiest cases to classify are the ones that we can visualise, i.e. the convex compact sets in zero-, one-, two- and three-dimensional spaces. In dimension zero we have singletons \{x\} for any real $x$ with pattern (), in one-dimensional space there is no freedom: the only fully dimensional convex compact sets are line segments, with the only possible pattern (1). On the plane the two-dimensional possibilities are exhausted by a circle and a triangle, with patterns (2) and (1, 2) respectively (see Fig. 5.3). Therefore for the at most two dimensional case we have four possibilities: (), (1), (1, 2) and (2), which coincides with the cardinality of the powerset of two: $2^2 = 4$.

![Figure 5.3: All possible face patterns of full dimensional sets in two dimensional case are given by a disk and a triangle.](image-url)
In three dimensions the possibilities for full dimensional sets are exhausted by the unit ball (3), the tetrahedron (1,2,3), the unit ball intersected with a closed half-space (2,3), and the convex hull of a circle in the plane and two points on opposite sides of the plane (1,3) (see Fig. 5.4), together with the lower dimensional examples we have in total \(2^3 = 8\) possibilities.

![All possible facial patterns for the three dimensional sets](image)

**Figure 5.4: All possible facial patterns for the three dimensional sets**

### 5.2 Main Result

We show that all patterns of facial dimensions can be realised by compact convex sets.

**Theorem 5.2.1** For any increasing sequence of positive integers

\[
d = (d_1, d_2, \ldots, d_k)
\]

there exists a compact convex set in \(d_k\)-dimensional space such that the vector \(d\) describes the pattern of facial dimensions for this set.

To prove this, we need the following technical lemma, which is surely known, but we were not able to identify it in the literature. We hence provide a short proof here as well.

**Lemma 5.2.2** Let \(P, Q \subseteq \mathbb{R}^n\) be nonempty convex compact sets, and let \(C = P + Q\). Then every face of \(C\) is the Minkowski sum of faces of \(P\) and \(Q\). More precisely,

\[
\forall F \triangleleft C \quad \exists F_P \triangleleft P, F_Q \triangleleft Q \text{ such that } F = F_P + F_Q.
\]
Proof Let $F$ be a nonempty face of $C$. We construct two sets

$$F_P := \{ x \in P \mid \exists y \in Q, x + y \in F \}, \quad F_Q := \{ y \in Q \mid \exists x \in P, x + y \in F \}.$$ 

Both $F_P$ and $F_Q$ are nonempty since $F$ is nonempty.

First we show that $F = F_P + F_Q$. It is obvious that $F \subset F_P + F_Q$, and it remains to show the reverse inclusion. For that, pick an arbitrary $x \in F_P$, $y \in F_Q$. We will next show that $z = x + y \in F$.

By the definition of $F_P$ and $F_Q$ there exist $u \in P$ and $v \in Q$ such that $x + v \in F$ and $y + u \in F$. If $x = u$ or $y = v$, there is nothing to prove, as in this case $z = u + v \in F$. Otherwise, by the convexity of $F$ we have

$$z' = \frac{x + v}{2} + \frac{y + u}{2} \in F.$$ 

At the same time, notice that $x + y \in P + Q \subset C$; likewise, $u + v \in P + Q \subset C$, and $z' \in (x + y, u + v)$. Since $F$ is a face of $C$, this yields $z = x + y \in F$.

It remains to show that both $F_P$ and $F_Q$ are faces of $P$ and $Q$ respectively. First note that both are convex compact sets, and that $F_Q \subset Q$ and $F_P \subset P$.

Let $x \in F_P$, and pick any interval $[a, b] \subset P$ such that $x \in (a, b)$. By the definition of $C$, for an arbitrary $y \in F_Q$ we have $a + y, b + y \in C$. At the same time, $x + y \in F_P + F_Q = F$ and $x + y \in (a + y, b + y)$. From $F \subset C$ we have $[a + y, b + y] \subset F$, hence, $a + y, b + y \in F$, and therefore $a, b \in F_P$. This shows that $F_P$ is a face of $P$. The proof for $F_Q$ is identical.

In the proof of Theorem 5.2.1 presented next we use an inductive argument to explicitly construct a compact convex set with a given facial pattern from a lower dimensional example for a truncated sequence. The key observation is that the Minkowski sum of an arbitrary compact convex set with a unit ball does not generate faces of any new dimensions (compared to the original set) other than possibly the fully dimensional face that coincides with the sum, which follows directly from Lemma 5.2.2. We sketched the Minkowski sum of two simple compact convex sets with a Euclidean ball in Fig. 5.5 to illustrate this argument.

**Proof of Theorem 5.2.1** We use induction on $d_k$ to demonstrate the result. Our induction base is lower dimensional examples discussed earlier. For all increasing sequences of positive
numbers \((d_1, \ldots, d_k)\) with \(d_k \leq 2\) we have found the relevant examples. They are realised by a point, line segment, disk and triangle.

Assume that our assertion is proven for all sequences \((d_1, \ldots, d_k)\) with \(d_k \leq m\). We will show that the statement is true for \(d_k = m+1\). Choose an arbitrary sequence \(d = (d_1, \ldots, d_k)\), where \(d_k = m+1\). If \(d = (d_k)\), the sequence is realised by the Euclidean unit ball in \(\mathbb{R}^{m+1}\). If the sequence contains more than one number, consider the truncated sequence \(d' = (d_1, d_2, \ldots, d_{k-1})\). Since \(d_{k-1} < d_k\), we have \(l := d_{k-1} \leq m\), and there exists a compact convex set \(Q \subset \mathbb{R}^l\) that realises the sequence \(d'\) in \(l = d_{k-1}\)-dimensional space. We embed the set \(Q\) in the \(m+1\)-dimensional space by letting \(Q' := Q \times \{0_{m+1-l}\}\). Observe that since the definition of the face is algebraic, the facial pattern of the set \(Q'\) is identical to the one of \(Q\). Let \(B\) be the unit ball in \(\mathbb{R}^{m+1}\). We let

\[
C := B + Q'
\]

and claim that \(d\) is the facial pattern of \(C\).

From Lemma 5.2.2 every face of \(C\) can be represented as the sum of faces of \(Q'\) and \(B\). Since the only faces of \(B\) are the set itself and the singletons on the boundary, the only possible dimensions of the faces of the set \(C\) can come from the sequence \((d_1, \ldots, d_k)\). To show that no facial dimensions are lost, observe that if \(e\) denotes the unit vector \((0, 0, \ldots, 1)\) \(\in B\), then the set \(\{e\} + Q'\) is a face of \(C\) (hence all its faces are also faces of \(C\)). Indeed, for the hyperplane \(H = \{x \mid \langle e, x \rangle = x_{m+1} = 1\}\) supports \(C\) (notice that for every \(x = q + b \in C\)
with \( q \in Q' \) and \( b \in B \) we have \( x_{m+1} = 0 + q_{m+1} \leq 1 \), moreover,

\[
H \cap C = \{ q + b \mid q \in Q', b \in B, q_{m+1} + b_{m+1} = 1 \}
\]

\[
= \{ q + b \mid q \in Q', b \in B, b_{m+1} = 1 \}
\]

\[
= \{ e \} + Q'.
\]

It is not difficult to observe (e.g., see [47, Section 18]) that any supporting hyperplane slices off a face from a convex set, hence, \( F = \{ e \} + Q' \triangleleft C \). This face is linearly isomorphic to \( Q \), and hence the facial structure of \( F \) conicides with the facial structure of \( Q \), giving all possible dimensions of faces from the sequence \( d' \). The face of the maximal dimension \( m + 1 \) is given by the set \( C \) itself, as it has a nonempty interior (take any point from \( Q' \) and sum it with an open ball).

### 5.3 Fractal convex sets

Observe that polytopes not only possess faces of all possible dimensions, but their faces are also arranged in a very regular fashion: the union of the edges of a polytope is a one-dimensional set (here we refer to a general notion of Hausdorff dimension, rather than the dimension of the affine hull that is useful for convex sets), the union of all two dimensional faces is two dimensional, and so on. More generally, the union of all faces of a polytope of a given dimension is a set of the same dimension. This is not the case for a more general setting: for instance, the dimension of the union of all extreme points of a Euclidean ball in \( \mathbb{R}^n \) is \( n - 1 \), a stark contrast with the polyhedral case. Hence it is natural to study the dimension of the unions of equidimensional faces. The purpose of this section is to present some examples of nontrivial sets with fractal facial structure and hence noninteger dimensions of the said unions; these form the foundation for our ongoing research on this topic.

Some work on fractals and convexity has been done before (see the recent work [59] and references therein), but we are not aware of any references studying the particular problems that we propose here. We focus on two examples of convex sets that are generated in a natural way by spherical fractals. Finite root systems and Coxeter systems are fundamental
concepts in Lie algebras, which is very important in many branches of mathematics. Given a finite root system, there is a natural associated finite Coxeter group, which is the Weyl group. People in the field of geometric group theory consider finite Coxeter groups are well-studied and explained in literature, see [30]. Therefore, we are more interested in the behaviours of infinite Coxeter groups. One such fractal comes from a recent work [53] on study of infinite Coxeter groups, another one is constructed via projecting the Sierpinski triangle onto the unit ball.

We first consider an arbitrary fractal set on a sphere and then take its convex hull, hence generating a convex set. Our first example is constructed in a similar way to the Apollonian gasket: we take the unit sphere and construct a tetrahedron whose edges touch the sphere (see Fig. 5.6), then consider the intersection of the sphere with the tetrahedron. After that, we continue slicing off spherical caps in such a way that they are tangential to the existing slices (see Fig. 5.7).

In more details, observe the Figure 5.8, the orange coloured set is formed by infinite number of spherical caps, and we would like to take the union of all these spherical caps, then the complement is our desired fractal set, which is the dark blue coloured set. The resulting body is a spherical fractal, which is also a convex set. If we now take its convex hull, the extreme points of this convex set would be exactly the points on the fractal set, with remaining proper faces disks that result from the sliced off spherical caps.
Algebraically, this particular fractal set is generated by the infinite Coxeter group with following group presentation:

\[ G = \langle s_1, s_2, s_3, s_4 \mid (s_i)^2 = (s_i s_j)^\infty = 1 \rangle \]

The fractal sets are generated by limit roots, see [53]. Limit roots exhibit peculiar geometric behaviour. Even though Coxeter groups are generated by affine reflections across hyperplanes, when we compute the roots of the group and project them down to a lower
dimensional affine hyperplane, the set of limit roots behaves like a fractal set, giving self-similar patterns that cannot be obtained by reflecting across any hyperplanes. For details on the computation process of limit roots, see [53].

This approach can be applied to constructing other spherical fractals. For instance, one can generalise the Sierpinski carpet by cutting out triangular pieces of the sphere in a similar fashion. The convex set obtained after taking the convex hull of this spherical fractal will have faces of all possible dimensions, as in this set contains vertices with dimension 0, edges with dimension 1, two dimensional triangle shaped faces, and so on.

The Hausdorff dimension of the union of the extreme points is non-integer in both cases, and coincides with the dimension of the relevant two-dimensional objects. It would be interesting to study the conditions that can be imposed on the facial dimensions to define good or regular convex sets.

### 5.4 Future work

The results that are presented in this chapter apply to general convex sets. There are some interesting future work we can study. For example, the fractal sets in Section 5.3 have nontrivial Hausdorff dimension of the union of the extreme points, which are non-integer in both cases. What nontrivial constraints on convex sets lead to nontrivial constraints on the facial dimensions? For example, spectrahedra is important in optimization as semidefinite programming problems maximise a linear function over a spectrahedron. So we might be interested in what happens for spectrahedra? Is it possible to explicitly understand these constraints?

### 5.5 Conclusion

In this chapter, we have introduced the difference between patterns for the dimensions of faces of polyhedral sets and compact convex sets. Then, we have proved that given facial dimension sequence, there always exist a compact convex set with this pattern. In the end, we discussed the method to approach dimensionality by considering the dimension of union
of the faces with the same dimension. In particular, we looked at some examples constructed from fractal sets on the sphere. We have also posed some questions for future work.
Chapter 6

Conclusion

The main body of the thesis comes from three published papers.

In Chapter 3, we have proved that, under an affinely independence condition, that is, when we restrict the number of vertices of polytopes in the collection to $n + 1$ affinely independent points for an $n$ dimensional space, the Demyanov-Ryabova conjecture is true. The main idea of the proof of this result comes from by taking the convex hull $C$ which contains all convex sets in the collection, and then considering the points that are in $C$ which will recur after two iterations.

We have also obtained a combinatorial reformulation of the conjecture by ordering vertices in the collection. Essentially, we were able to encode direction vectors in the space using the ordering of vertices. This combinatorial formulation allows us to work on the conjecture using more algebraic approaches. After we obtain the set of orderings on the vertex set that correspond to the set of restricted directions, we are able to forget about the geometry of the sets. We have shown that this combinatorial formulation is equivalent to the original version of the conjecture. This means that, we may try to apply algebraic and combinatorial tools to this problem. We have also provided some observations about symmetry on the conjecture. This should advance insight for the future work on conditions under which the general conjecture holds, using this algebraic approach.

Some potential work can be done for this problem in the future. Firstly, we have provided an example of two circles in Section 3.6, which motivated Conjecture 3.6.2 on general convex
sets. Also, with the counterexample that was constructed in [51], which has disproved
the Demyanov-Ryabova conjecture, we wonder if more interesting counter examples can be
constructed.

For our study of outer limits of subdifferentials, we focused on the constructive evaluation
of the error bound modulus for structured continuous functions. Error bound plays crucial
role in optimisation, the error bound modulus measures whether a given function is steep
enough outside of its level set and gives a lower bound for the relevant slope.

In Chapter 4, We have generalised the outer subdifferential construction suggested by
Cánovas, Henrion, López and Parra for max type functions to pointwise minima of regular
Lipschitz functions. We have provided an exact description for the outer limits of subdif-
ferentials in the case of polyhedral functions and sharp bounds for a more general case. In
particular, the outer limits of subdifferential can be expressed by taking the union of the
intersections of support faces.

We have also answered an open question about the relation between the outer subdiffer-
ential of the support of a regular function and the end set of its subdifferential as posed by
Li, Meng and Yang.

Our work in Chapter 5 focuses on the facial structure for convex sets. There are little
information regard to facial structure of general convex sets that can be found in classic
convex analysis books and earlier literature. Therefore, studies on the facial structure of
convex sets provide good insights into finding minima in optimisation problems.

We address the question of which kind of dimension patterns are possible for the faces
of a general closed convex set. Consider faces of a polytope, we know that faces of every
possible dimension are present, which is not true for a general compact convex set. We have
shown that given a finite sequence of positive integers, there always exist a compact convex
set which consists of faces with dimensions from this prescribed sequence. We also discuss a
different approach to dimensionality, which is by considering the dimension of the union of
all faces of the same dimension. We have provided some examples of convex sets by taking
convex hull of fractal sets, which have shown that the outcome of this approach can be highly
nontrivial, as the union of all extreme points will have non-integer dimension.
Bibliography


