DEPARTMENT OF MATHEMATICS

GRADUATE PROJECT REPORT

ON

Difference of Convex(D.C.) Functions and Their Minimal Representations

SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENT FOR THE DEGREE OF MASTER OF SCIENCE IN MATHEMATICS

By: **Gebretsadik Gidey**

Advisor:**Semu Mitiku(PhD)**

June 27, 2013

Addis Ababa, Ethiopia

This is to certify that this project is compiled by Mr. **Gebretsadik Gidey** in the Department of Mathematics Addis Ababa university under my supervision.

Gebretsadik G. ________________ Date:________________

ADDIS ABABA UNIVERSITY
DEPARTEMENT OF MATHEMATICS

The undersigned hereby certify that they have read and recommended to the Department of Mathematics for acceptance of a project entitled **Difference of Convex(D.C.) Functions and Their Minimal Representations** by **Gebretsadik Gidey** in partial fulfilment of the requirements for the degree of Master of Science.

Date:__________________

Advisor: Dr. Semu Mitiku

Signature: _________________________

Examining Committee: 1. ____________________________________

Signature: _________________________

2.____________________________________

Signature: _________________________

Contents

Acknowledgement

I would like to express my genuine and sincere gratitude to my advisor Dr.Semu Mitiku for his valuable comments, suggestions, remarks, follow up and encouragement while I was working my project.
And also I would like to to thank to my parents at all for their support on all around my life. I want also to say thanks for all my friends.

Abstract

A function f defined on a given convex set X which can be expressed as a difference of two convex (continuous) functions is called d.c function or δ-convex function. The functions which are Lipschitz and bounded variation are expressible as a d.c. function and since those family of d.c. functions form a linear space as well as a lattice, it admits many operations.

The decomposition of a given function f as a d.c. functions is not unique. Choosing the better (minimal) decomposition is useful in describing the optimality conditions for d.c. optimization.

Notations

∇	Gradient
$<,>$	Inner product
$\triangle$	Laplacian operator
$\|\|.\|\|$	Normed space
C^2	Twice continuously differentiable
D.C.	Difference of convex
D.S.F.	Difference of Sabharmonic function
$\mathbb{R}$	The set of all real numbers
l.s.c.	Lower semi-continuous
NSO	Non smooth optimization
QD	Quasidifferentiable
u.s.c.	Upper semi-continuous

Introduction

Many problems in application could not be modelled as a linear programming problem where as they could be modelled as continuous optimization problem. Such problems are called nonlinear. Where one had to minimize (or to maximize) a convex function subject to convex inequality and affine equality constraints, such problems are termed as convex optimization problems. To solve Nonlinear optimization problems, one uses the analogy called convex optimization. For the problem with smooth functions, we use the classical methods that most of the methods depend on the directional derivative (or its generalization) of the functions.

But if f is nonsmooth, the directional derivative (or its generalization) $f'(x; d)$ of f is nonlinear. Then the next attempt would be to employ the tools of convex analysis to treat the nonlinearity of the directional derivative. This may fail because the map $d \mapsto f'(x; d)$ is not convex for general non-smooth function f.

To convexify the map $d \mapsto f'(x; d)$, the generalized gradient(Clarke's generalized gradient) which is found by regularizing the directional derivative could be the first to be mentioned here. Clarke's generalized gradient,$\partial_{cl} f(x)$, is sometimes too large a set whose elements are also difficult to calculate, even though the Clarke generalized directional derivative, $\mathring{f}(x; d)$, is convex. This problem can be eased in some sense if the function f is directionally differentiable and the directional derivative, $f'(x; d)$, can be written as a difference of two positively homogeneous convex (and hence sublinear) functions. However, the dual of $f'(x; d)$ is no longer a single set as in the convex case but a pair of two compact convex sets, denoted by ($\underline{\partial} f(x), \overline{\partial} f(x)$), and is called quasidifferential. This pair is not uniquely defined but each pair can be considered as an equivalence class after defining a certain equivalence relation. The non uniqueness of such pairs much contributed to the study of the problem of finding minimal representatives of each class.

A pair of compact convex sets ($\underline{\partial} f(x), \overline{\partial} f(x)$) is equivalent (in vertue of Minkowski duality) to a pair of two sublinear functions, say (g, h), i.e

$$f'(x; d) = g(d) - h(d) \quad \text{for sublinear functions } g \text{ and } h.$$

Therefore finding a minimal representative of a class of quasidifferentials is related to finding a minimal representative of the two convex (called d.c. for short) functions.
One of the advantages of studying about d.c. functions is, beyond convex functions many nonconvex functions can be written as a d.c. functions. For example, all concave functions are d.c. functions, and polynomial functions in $\mathbb{R}^n$ can decompose as a d.c. functions. Generally almost all types of continuous functions can be represented or approximated by d.c. functions.
D.C. functions widely exist in optimization problems. Some examples of d.c. functions arising naturally are;

- Polynomials in several variables
- Variational analysis; Paraconvex and paraconcave functions are d.c. functions.
- Non-cooperative game theory; Nash equilibria
- Spectral theory in finite dimensions
- Operator theory; each symmetric bounded linear operator defined on a Banach space generates a quadratic form and this quadratic form is d.c. on that set.

Furthermore, the family of d.c. functions is a vector space, an Algebra (closed under multiplication) and a Lattice (closed under finite maxima/minima). So that this family is stable under the operations defined on vector space, an Algebra and a Lattice.

For a given function f, its decomposition as a d.c. functions is not unique. And from these non-unique decompositions, finding the better (minimal) one is useful in describing the optimality conditions of a given d.c. optimization. So finding this better (minimal) decomposition is another task that we discuss about.

Chapter 1

Definitions and Preliminaries

Definition 1.1. *A subset K of a metric space (S, d) is said to be compact if every open covering of K contains a finite subcovering.*

Definition 1.2. *Let a, b be two points of $\mathbb{R}^n$. The set of all $x \in \mathbb{R}^n$ of the form*

$$x = (1 - \lambda)a + \lambda b = a + \lambda(b - a), \quad \lambda \in \mathbb{R} \tag{1.1}$$

is called the line through a and b.
*A subset M of $\mathbb{R}^n$ is called an **affine set (or affine manifold)** if it contains every line through any two points of it, i.e. if $(1 - \lambda)a + \lambda b \in M$ for every $a, b \in M$ and every $\lambda \in \mathbb{R}$.*

Definition 1.3. *A set $C \subset \mathbb{R}^n$ is called **Convex Set** if*

$$(1 - \lambda)a + \lambda b \in C \quad whenever \quad a, b \in C, \quad 0 \leq \lambda \leq 1. \tag{1.2}$$

Definition 1.4. *Given any set $E \subset \mathbb{R}^n$, there exists a convex set containing E namely $\mathbb{R}^n$.The intersection of all convex sets containing E is called **Convex hull** of E and denoted by $convE$.*
i.e. $convE$ is the smallest set containing E.

Given a function $f : S \to [-\infty, \infty]$ on a set $S \subset \mathbb{R}^n$, then the sets

$$domf = \{x \in S | f(x) < +\infty\}$$
$$epif = \{(x, \alpha) \in S \times \mathbb{R} | f(x) \leq \alpha\}$$

are called the effective domain and the epigraph of $f(x)$ respectively.

Definition 1.5. *A function $f : S \to [-\infty, \infty]$ is called convex if its epigraph is a convex set in $\mathbb{R}^n \times \mathbb{R}$.*
Equivalently f is convex if S is a convex set in $\mathbb{R}^n$ and for any $x_1, x_2 \in S$ and $\lambda \in [0, 1]$, we have

$$f((1-\lambda)x_1 + \lambda x_2) \leq (1-\lambda)f(x_1) + \lambda f(x_2) \tag{1.3}$$

Whenever the right side is defined.

We can also define whether a differentiable function is Convex or not as follows.

Corollary 1.1. *A* ***differentiable*** *real valued function $f(x)$ on an open interval is* ***convex*** *if and only if its derivative f' is non decreasing function. A twice* ***differentiable*** *real valued function $f(x)$ on an open interval is* ***convex*** *if and only if its second derivative f'' is nonnegative throughout this interval.*

Proposition 1.1. *A* ***twice differentiable*** *real valued function $f(x)$ on an open convex set $C \subset \mathbb{R}^n$ is convex if and only if $\forall x \in C$ its Hessian matrix*

$$Q_x = (q_{ij}(x)), \quad q_{ij}(x) = \frac{\partial f(x_1, x_2, \ldots, x_n)}{\partial x_i \partial xj}$$

is positive semi definite. i.e.$\langle U, Q_x U\rangle \geq 0 \quad \forall U \in \mathbb{R}^n$

Proof. The function f is convex on C if and only if for each $a \in C$ and $U \in \mathbb{R}^n$,the function $\varphi_{a,U} = f(a + tU)$ is convex on the open real interval $\{t|a + tU \in C\}$. Then this Proposition follows from Corollary 1.1 since an easy computation yields $\varphi''_{a,U} = \langle U, Q_x U\rangle$ with $x = a + tU$. □

Let S be non empty subset of real normed space $(\mathscr{L}, \|.\|)$. Let $f : S \to \mathbb{R}$ be given function. Let $x_0 \in S$, then f is said to be Lipschitz continuous at x_0 if there is a constant $L \geq 0$ called Lipschitz constant and some $\varepsilon > 0$ such that

$$|f(x) - f(y)| \leq L\|x - y\|, \quad \forall x, y \in S \cap \beta(x_0, \varepsilon) \tag{1.4}$$

Where $\beta(x_0, \varepsilon) = \{x \in \mathscr{L} / \|x - x_0\| < \varepsilon\}$.

Let $f : \mathbb{R}^n \to [-\infty, +\infty]$ be any function. And let x_0 be a point where f is finite. If for some $d \neq 0$, the limit (finite or infinite)

$$f'(x_0; d) = \lim_{t \to 0^+} \frac{f(x_0 + td) - f(x_0)}{t} \tag{1.5}$$

exists, then $f'(x_0; d)$ is called the **directional derivative** of f at x_0 in the direction of d.

If the limit in equation (1.5) exists for all $d \in X$, then f is called directionally differentiable at x_0. Nonsmooth optimization (NSO) refers to the general problem of minimizing (or maximizing) functions that are typically not differentiable at their minimizers (maximizers). Since the classical theory of optimization presumes certain differentiability and strong regularity assumptions upon the functions to be optimized, it can not be directly utilized. However, due to the complexity of the real world, functions involved in practical applications are often nonsmooth. That is, they are not necessarily differentiable. In what follows, we briefly introduce the basic concepts of nonsmooth analysis and optimization.

Let us consider the NSO problem of the form

$$\begin{aligned} &\max f(x) \\ &s.t \quad x \in X \end{aligned}$$

where the objective function $f : \mathbb{R}^n \to \mathbb{R}$ is supposed to be locally Lipschitz continuous on the feasible set $G \subseteq \mathbb{R}^n$. Note that NSO techniques can be successfully applied to smooth problems but not vice versa and thus, we can say that NSO deals with a broader class of problems than smooth optimization. Although using a smooth method may be desirable when all the functions involved are known to be smooth, it is often hard to confirm the smoothness in practical applications (e.g. if function values are calculated via simulation). Moreover, as already mentioned, the problem may be analytically smooth but still behave numerically nonsmoothly, in which case an NSO method is needed. The theory of nonsmooth analysis is based on convex analysis. Thus, we start by giving some definitions and results for convex (not necessarily differentiable) functions. We define the subgradient and the subdifferential of a convex function. Then we generalize these results to nonconvex locally Lipschitz continuous functions.

Definition 1.6. *Given a proper function f on $\mathbb{R}^n$, a vector $p \in \mathbb{R}^n$ is called a* ***subgradient*** *of f at a point x_0 if*

$$\langle p, x - x_0 \rangle + f(x_0) \leq f(x) \quad \forall x \tag{1.6}$$

The set of all subgradients of f at x_0 is called the ***subdifferential*** *of f at x_0 and is denoted by $\partial f(x_0)$.*
The function f is said to be subdifferentiable at x_0 if $\partial f(x_0) \neq \emptyset$.

Example 1.1. *:Absolute-value function*
A function $f(x) = |x|$ is clearly convex and differentiable when $x \neq 0$. By the definition of subdifferential

$$\begin{aligned} \xi \in \partial f(0) &\Leftrightarrow |y| \geq |0| + \xi \cdot (y - 0) \textit{ for all } y \in \mathbb{R} \\ &\Leftrightarrow |y| \geq \xi \cdot y \textit{ for all } y \in \mathbb{R} \\ &\Leftrightarrow \xi \leq 1 \textit{ and } \xi \geq -1. \end{aligned}$$

Thus, $\partial f(0) = [-1, 1]$.
The next theorem shows the relationship between the subdifferential and the directional derivative. It turns out that knowing $f'(x; d)$ is equivalent to knowing $\partial f(x)$.

Theorem 1.1. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function. Then for all $x \in \mathbb{R}^n$*

1. $f'(x; d) = \max\{\xi^T d \mid \xi \in \partial f(x)\}$ *for all $d \in \mathbb{R}^n$ and*
2. $\partial f(x) = \{\xi \in \mathbb{R}^n \mid f'(x; d) \geq \xi^T d$ *for all* $d \in \mathbb{R}^n\}$

Example 1.2. *By the previous theorem we have*

$$\xi \in \partial f(0) \Longleftrightarrow f'(0; d) \geq \xi \cdot d \textit{ for all } d \in \mathbb{R}.$$

Now

$$f'(0, d) = \lim_{t \downarrow 0} \tfrac{|0 + td| - |0|}{t} = \lim_{t \downarrow 0} \tfrac{t|d|}{t} = |d|$$

and, thus,

$$\begin{aligned} \xi \in \partial f(0) &\Longleftrightarrow |d| \geq \xi \cdot d \textit{ for all } d \in \mathbb{R} \\ &\Longleftrightarrow \xi \in [-1, 1]. \end{aligned}$$

Since there does not necessarily exist any classical directional derivatives for locally Lipschitz continuous functions, we first define a generalized directional derivative. We then generalize the subdifferential for nonconvex locally Lipschitz continuous functions.

Definition 1.7. *Let $S \subseteq X, S \neq \emptyset$, let $f : S \to \mathbb{R}$ be a given function, $d \in X$ be any given vector, $x \in S$ be any given point. Then if the limit*

$$f^\circ(x; d) = \limsup_{\substack{y \to x \\ t \searrow 0}} \frac{f(y + td) - f(y)}{t}, \quad \text{where } y \in X \text{ and } t \in \mathbb{R} \tag{1.7}$$

*exists, then $f^\circ(x; d)$ is called the **Clarke derivative** of f at x in the direction of d. If the limit in equation (1.7) exists for all $d \in X$, then f is called Clarke differentiable at x.*

Note that this generalized directional derivative always exists for locally Lipschitz continuous functions and, as a function of d , it is sublinear. Therefore, we can now define the subdifferential for nonconvex locally Lipschitz continuous functions as analogous to Theorem 1.1 with the directional derivative replaced by the generalized directional derivative.

Definition 1.8. *(Clarke) Let $f : \mathbb{R}^n \to \mathbb{R}$ be a locally Lipschitz continuous function at a point $x \in \mathbb{R}^n$. Then the subdifferential of f at x is the set $\partial f(x)$ of vectors $\xi \in \mathbb{R}^n$ such that*

$$\partial f(x) = \{\, \xi \in \mathbb{R}^n \mid f^\circ(x; d) \geq \xi^T d \text{ for all } d \in \mathbb{R}^n \,\}.$$

Theorem 1.2. *Let $S \subset \mathbb{R}^n$ be an open set. A function $f : S \to \mathbb{R}$ that is locally Lipschitz continuous on S is differentiable almost everywhere on S.*

Theorem 1.3. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a locally Lipschitz continuous function at a point $x \in \mathbb{R}^n$. Then*

$$\partial f(x) = conv\{\xi \in \mathbb{R}^n \mid \nabla f(x_i) \to \xi,\ x_i \to x \text{ and } f \text{ is differentiable at } x_i\},$$

where $\mathrm{conv} S$ *denotes the convex hull of set S.*

Example 1.3.

$$f(x) = |x|.$$

The subdifferential of the absolute-value function f at $x = 0$ is given by

$$\partial f(0) = conv\{-1, 1\} = [-1, 1].$$

Definition 1.9. *Let S be a subset of a real normed space $(X, \|.\|)$ with nonempty interior, and let $f : S \to \mathbb{R}$ be a functional which is Lipschitz continuous at some $x_0 \in int(S)$. Then the set $\partial_{cl} f(x_0)$ of all continuous linear functional ℓ on X with $f^\circ(x_0; d) \geq \ell(d) \quad \forall d \in X$ is called the Generalized gradient of f at x_0*

Definition 1.10. *Let S be a subset of a real normed space $(X, \|.\|)$ A functional $f : S \to \mathbb{R}$ is said to be regular at a point $x \in S$ if*

(a) f is directionally differentiable at x.

(b) $\forall d \in X, \quad f'(x; d) = f^\circ(x; d)$.

Definition 1.11. *Let S be non-empty open subset of a real normed space $(X, \|.\|)$, Let $f : S \to \mathbb{R}, x_0 \in S$ be given. Then f is said to be quasidifferentiable at x_0 if f is directionally differentiable at x_0 and if there are two non-empty convex weak$^*-$compact subsets A and B of the topological dual space X^* such that*

$$f'(x_0; d) = \max_{p \in A} p(d) - \max_{q \in (-B)} q(d) \quad \forall d \in X. \tag{1.8}$$

A functional $\ell : X^* \to \mathbb{R}$ is said to be *superlinear* if $-\ell$ is sublinear. i.e. ℓ is superlinear if it is

(i) *posetivelly homogeneous*: $\ell(\alpha d) = \alpha \ell(d) \quad \forall \alpha \geq 0, \quad d \in X$

(ii) *Superadditive*: $\ell(d_1 + d_2) \geq \ell(d_1) + \ell(d_2) \quad \forall d_1, d_2 \in X$.

Equation (1.8) can be rewritten as

$$f'(x_0; d) = \max_{p \in A} p(d) + \max_{q \in B} q(d) \quad \forall d \in X.$$

For a quasidifferentiable function f, the pair of sets (A,B) in equation (1.8) is not unique. Because equation (1.8) can be rewritten as

$$\begin{aligned} f'(x_0; d) &= \max_{\ell \in A} \ell(d) + \max_{\ell \in B} \ell(d) \\ &= \max_{\ell \in [A+a]} \ell(d) + \max_{\ell \in [B-a]} \ell(d) \end{aligned}$$

for any Singleton set a.

Definition 1.12. 1. *Two pairs of Convex weak***-compact subsets* $[A_1, B_1]$ *and* $[A_2, B_2]$ *of* X^* *are said to be equivalent (similar) if* $A_1 + B_2 = A_2 + B_1 \Leftrightarrow A_1 - B_1 = A_2 - B_2$*. And since this relation is reflexive, symmetric and transitive, this relation is an equivalence relation.*

2. *Let* f *be quasidifferentiable at* $x_0 \in X$*.Then the class of equivalent pairs of sets (A,B) such that*

$$f'(x_0; d) = \max_{\ell \in A} \ell(d) - \max_{\ell \in (-B)} \ell(d) \quad \forall x_0 \in X$$

is called the quasidifferential (QD) of f *at the point* x_0 *and we denote it by* $Df(x_0)$*. If* $Df(x_0) = [A, B]$*, then the set*

- A *is called a subdifferential of* f *at* x_0 *and denoted by* $\underline{\partial} f(x_0)$*.*
- B *is called a superdifferential of* f *at* x_0 *and denoted by* $\overline{\partial} f(x_0)$*.*

$$\Rightarrow f'(x_0; d) = \max_{\ell \in \underline{\partial} f(x_0)} \ell(d) + \max_{\ell \in \overline{\partial} f(x_0)} \ell(d)$$

Hence the two sets $\underline{\partial} f(x_0)$ and $\overline{\partial} f(x_0)$ are considered always in pair. Therefore, QD of f at x_0 written as $Df(x_0) = [\underline{\partial} f(x_0), \overline{\partial} f(x_0)]$ is uniquely defined only upto the equivalence relation which is defined in Definition 1.12 of number 1.

Chapter 2

D.C. Functions and D.C. Sets

2.1 D.C. functions

Many functions which represent the real world application problems are nonconvex or non concave functions. But most of them can be written as a difference of two convex functions. To illustrate this let us see the following simple example.

Example 2.1. *Suppose we want to maximize the area of a rectangle whose perimeter P is given by $P = 1$. If the length of the rectangle is x_1 and its width is x_2. The problem is to maximize the area $A = x_1x_2$ subject to the perimeter $2x_1 + 2x_2 = 1$ where $x_1, x_2 > 0$. Thus, we have the optimization problem*

$$(P) \qquad \begin{aligned} &\max f(x) = x_1x_2 \\ &s.t \quad 2x_1 + 2x_2 = 1, x_1 > 0, x_2 > 0, \quad x \in \mathbb{R}^2_+ \end{aligned}$$

Here f is neither convex nor concave. So (P) is nonconvex optimization problem. However the objective function can be written as

$$\begin{aligned} f(x) =& \frac{1}{2}(x_1 + x_2)^2 - \frac{1}{2}(x_1^2 + x_2^2) \\ =& h(x) - g(x) \end{aligned}$$

where $g(x) = \frac{1}{2}(x_1^2 + x_2^2)$ and $h(x) = \frac{1}{2}(x_1 + x_2)^2$ are both convex functions. Here the objective function f is expressed as a difference of two convex functions g and h.

And the optimization problems which contains such problems are called d.c. optimization problems.
So in this chapter we discuss about d.c. functions and their properties.

Definition 2.1. *Let X be a convex set in $\mathbb{R}^n$. We say that a function is d.c. on X if it can be expressed as a difference of two convex functions on X. i.e. if $f(x) = f_1(x) - f_2(x)$,where f_1, f_2 are convex functions on X.*

There are many groups of functions which are representable as a d.c.(delta convex) functions. For example

- ▶ Convex functions
- ▶ Concave functions
- ▶ Quadratic functions and
- ▶ Polynomial functions

are the main group of functions which are so interesting in the applied science.
Here *" can every function be expressed as a difference of two convex functions over a given convex set X?"* The answer is negative. There are many functions (even Lipschitz continuous functions) which are not expressed as a difference of two convex functions. So to restrict the space of of d.c. functions, we need to state the following definitions, theorems and propositions.

Definition 2.2. *Let $[a, b] \subset \mathbb{R}$ with $a \leq b$. A partition of $[a, b]$ is finite ordered set $p = \{a = x_0 \leq \cdots \leq x_n = b\}$. Let $\mathfrak{B}_{a,b}$ be the collection of all partitions of $[a, b]$. For a real valued function f on $[a, b]$, we define the **variation** of f on $[a, b]$ corresponding to a partition p by*

$$\bigvee_a^b (f, p) = \sum_{k=1}^{n} |f(x_k) - f(x_{k-1})| \in [0, \infty) \tag{2.1}$$

*We define the **total variation** of f on $[a, b]$ by*

$$\bigvee_a^b f = \sup_{p \in \mathfrak{B}_{a,b}} \bigvee_a^b (f, p) = \sup\{\sum_{k=1}^{n} |f(x_k) - f(x_{k-1})| \in [0, \infty] \tag{2.2}$$

If $\bigvee_a^b(f) < \infty$ then, f is said to be a ***function of (or finite) variation*** *over $[a, b]$. Or simply a* ***Bounded variation (BV)*** *function on $[a, b]$. We write $BV([a, b])$ for the collection of all BV functions on $[a, b]$.*

Proposition 2.1. *(a) If $p, p' \in \mathfrak{B}_{a,b}$ and $p \subset p'$, then $\bigvee_a^b(f, p) \leq \bigvee_a^b(f, p')$.*

(b) If f is a real-valued monotone function on $[a, b]$, then f is bounded variation on $[a, b]$, and in this case we have $\bigvee_a^b f = |f(b) - f(a)|$.

Proof. (a)Let $p, p' \in \mathfrak{B}_{\mathfrak{a},\mathfrak{b}}$ and $p \subset p'$. Let n be the number of partition points in p and m be that of p'. We have $n \leq m$. If $n = m$, then $p = p'$ and $\bigvee_a^b(f, p) = \bigvee_a^b(f, p')$. Consider the case $n < m$.

Then $m = n + r, \quad r \in \mathbb{N}$. Let $\xi_1, \ldots, \xi_r$ be the r partition points in p' that are not in p. Let $p_0 = p$. Let p_1 be the partition of $[a, b]$ obtained by adding ξ_1 to p_0. Let p_2 be the partition of $[a, b]$ obtained by adding ξ_2 to p_1 and so on. So that $p_r = p'$ is the partition of $[a, b]$ obtained by adding ξ_r to p_{r-1}. Thus we have a chain of partitions of $[a, b]$, $p = p_0 \subset p_1 \subset \cdots \subset p_{r-1} \subset p_r = p'$, each having one more partition point than its predecessor.

Let $p = p_0 = \{a = x_0 \leq x_1 \leq \cdots \leq x_n = b\}$. Then $\xi_r \in (x_{k_0-1}, x_{k_0})$ for some $k_0 \leq n$. Now we have $\bigvee_a^b(f, p_0) = \sum_{k=1}^n |f(x_k) - f(x_{k-1})|$ and $\bigvee_a^b(f, p_1)$ is obtained by replacing the summand $|f(x_{k_0}) - f(x_{k_0-1})|$ in $\bigvee_a^b(f, p_0)$ with $|f(\xi_1) - f(x_{k_0-1})| + |f(x_{k_0}) - f(\xi_1)|$ which is greater than or equal to $|f(x_{k_0}) - f(x_{k_0-1})|$ by the triangle inequality in $\mathbb{R}$. Thus we have $\bigvee_a^b(f, p_0) \leq \bigvee_a^b(f, p_1)$. By the same argument

$\bigvee_a^b(f, p_1) \leq \bigvee_a^b(f, p_2) \leq \bigvee_a^b(f, p_3) \leq \cdots \leq \bigvee_a^b(f, p_r)$.

Thus $\bigvee_a^b(f, p) \leq \bigvee_a^b(f, p')$.

(b) Let f be real-valued monotone function on $[a, b]$. Then for an arbitrary partition $P = \{a = x_0 \leq \cdots \leq x_n = b$ of $[a, b]$, the difference $f(x_k) - f(x_{k-1})$ for $k = 1, ..., n$ are all non negative if f is an increasing function and all non-positive if f is decreasing function. Therefore we have

$$\bigvee_a^b(f, p) = \sum_{k=1}^n |f(x_k) - f(x_{k-1})| = |\sum_{k=1}^n [f(x_k) - f(x_{k-1})]| = |f(b) - f(a)|$$

and then $\bigvee_a^b f = \sup_{P\in\mathfrak{B}_{a,b}} \bigvee_a^b (f,p) = \sup \sum_{k=1}^n |f(x_k) - f(x_{k-1})| = |f(b) - f(a)|$

□

Lemma 2.1. *Let f_1, f_2 be functions of bounded variation on $[a,b]$ and $a, b \in \mathbb{R}$. then $cf_1 + df_2$ is a function of bounded variation on $[a,b]$ and $\bigvee_a^b(cf_1 + df_2) \leq |c| \bigvee_a^b(f_1) + |d| \bigvee_a^b(f_2)$*

Proof. For an arbitrary partition $p = \{a = x_0 \leq \cdots \leq x_n = b\}$ of $[a,b]$, we have

$$\begin{aligned}
\bigvee_a^b (cf_1 + df_2, p) &= \sum_{k=1}^n |(cf_1 + df_2)(x_k) - (cf_1 + df_2)(x_{k-1})| \\
&\leq |c| \sum_{k=1}^n |f_1(x_k) - f_1(x_{k-1})| + |d| \sum_{k=1}^n |f_2(x_k) - f_2(x_{k-1})| \\
&= |c| \bigvee_a^b (f_1, p) + |d| \bigvee_a^b (f_2, p) \\
&\leq |c| \bigvee_a^b (f_1) + |d| \bigvee_a^b (f_2)
\end{aligned}$$

Thus we have

$$\bigvee_a^b (cf_1 + df_2) = \sup_P \bigvee_a^b (cf_1 + df_2, p) \leq |c| \bigvee_a^b (f_1) + |d| \bigvee_a^b (f_2)$$

□

Lemma 2.2. *(a) If $f \in BV([a,b])$, then for every closed subinterval $[a_0, b_0]$ of $[a,b]$, we have $f \in BV([a_0,b_0])$ and $\bigvee_{a_0}^{b_0}(f) \leq \bigvee_a^b(f)$.*

(b) Let $c \in (a,b)$. If $f \in BV([a,c])$ and $f \in BV([c,b])$,then $f \in BV([a,b])$ and moreover $\bigvee_a^b(f) = \bigvee_a^c(f) + \bigvee_c^b(f)$.

Proof. (a) Let $p = \{a_0 = x_0 \leq \cdots \leq x_n = b_0\}$ be an arbitrary partition of $[a_0, b_0]$.

Consider the partition $p' = \{a \le a_0 = x_0 \le \cdots \le x_n = b_0 \le b\}$ of $[a, b]$. Then

$$\begin{aligned}\bigvee_{a_0}^{b_0}(f,p) &= \sum_{k=1}^{n}|f(x_k) - f(x_{k-1})| \\ &\le |f(x_0) - f(a)| + \sum_{k=1}^{n}|f(x_k) - f(x_{k-1})| + |f(b) - f(x_n)| \\ &= \bigvee_{a}^{b}(f,p') \le \bigvee_{a}^{b}(f) < \infty\end{aligned}$$

Thus we have

$$\bigvee_{a_0}^{b_0}(f) = \sup_{p\in\mathfrak{B}_{a_0,b_0}} \bigvee_{a_0}^{b_0}(f,p) \le \bigvee_{a}^{b}(f) < \infty$$

So that $f \in BV([a_0, b_0])$

(b) Let $c \in [a, b]$ and suppose $f \in BV([a, c])$ and $f \in BV([c, b])$. Let p be an arbitrary partition of $[a, b]$ and let p' be the refinement of p by adding c as a partition point to p by (a) of proposition 2.1, we have

$$\bigvee_{a}^{b}(f,p) \le \bigvee_{a}^{b}(f,p') \tag{2.3}$$

Let p_1 and p_2 be the restriction of p' to $[a, c]$ and $[c, b]$ respectively. Then we have

$$\bigvee_{a}^{b}(f,p') = \bigvee_{a}^{c}(f,p_1) + \bigvee_{c}^{b}(f,p_2) \le \bigvee_{a}^{c}(f) + \bigvee_{c}^{b}(f) \tag{2.4}$$

by equations 2.3 and 2.4 we have

$$\bigvee_{a}^{b}(f) = \sup_{p\in\mathfrak{B}_{a,b}} \bigvee_{a}^{b}(f,p) \le \bigvee_{a}^{c}(f) + \bigvee_{c}^{b}(f) < \infty \tag{2.5}$$

This shows that $f \in BV([a, b])$.

Let p_1 and p_2 be arbitrary partitions of $[a, c]$ and $[c, b]$ respectively. Then $p_1 \cup p_2$ is a partition of $[a, b]$ and

$$\bigvee_{a}^{b}(f,p_1 \cup p_2) = \bigvee_{a}^{c}(f,p_1) + \bigvee_{c}^{b}(f,p_2).$$ Thus we have

$$\bigvee_{a}^{b}(f) = \sup_{p\in\mathfrak{B}_{a,b}} \bigvee_{a}^{b}(f,p) \ge \bigvee_{a}^{b}(f,p_1 \cup p_2) = \bigvee_{a}^{c}(f,p_1) + \bigvee_{c}^{b}(f,p_2).$$

Since this holds for arbitrary partitions p_1 and p_2 of $[a, c]$ and $[c, b]$ respectively, we have

$$\bigvee_a^b (f) \geq \bigvee_a^c (f) + \bigvee_c^b (f) \tag{2.6}$$

Thus from equations 2.5 and 2.6

$$\bigvee_a^b (f) = \bigvee_a^c (f) + \bigvee_c^b (f). \tag{2.7}$$

□

Definition 2.3. *Let f be a function of bounded variation. The total variation of f on $[a, b]$ is defined by $V_f(x) = \bigvee_a^x f$ for $x \in [a, b]$.*

Proposition 2.2. *Let $f \in BV([a, b])$. The total variation function V_f of f is a non-negative real-valued increasing function on $[a, b]$ with $V_f(a) = 0$. Indeed for $a \leq x' < x'' \leq b$, we have*

$$V_f(x'') - V_f(x') = \bigvee_{x'}^{x''} (f) \geq 0$$

Proof. If $f \in BV([a, b])$, then for any $x \in [a, b]$, we have $f \in BV([a, x])$ with $\bigvee_a^x (f) \leq \bigvee_a^b (f)$ (by (a) of Lemma 2.2).
Thus V_f is a non-negative real-valued function on $[a, b]$.
To show that V_f is an increasing on $[a, b]$, Let $a \leq x' < x'' \leq b$. By (b) of Lemma 2.2 we have

$$\bigvee_a^{x'} (f) + \bigvee_{x'}^{x''} (f) = \bigvee_a^{x''} (f) \text{ that is}$$

$$V_f(x') + \bigvee_{x'}^{x''} (f) = V_f(x'')$$

$$\text{so that } V_f(x'') - V_f(x') = \bigvee_{x'}^{x''} (f) \geq 0$$

□

Theorem 2.1. *[7, 9](**Jordan Decomposition of Functions of Bounded Variation**)*
Let f be a real valued function on $[a, b]$. Then $f \in BV([a, b])$ if and only if there exist two real valued increasing functions f_1 and f_2 on $[a, b]$ such that $f = f_1 - f_2$ on $[a, b]$.

Proof. ($\Leftarrow$) Suppose there exist two real-valued increasing functions f_1 and f_2 on $[a, b]$ such that $f = f_1 - f_2$ on $[a, b]$. Then f_1 and f_2 are functions of bounded variation (by Proposition 2.1 b) and thus f is function of bounded variation (by Lemma 2.1)
($\Rightarrow$) Suppose f is a function of bounded variation.
Consider the total variation function V_f of f on [a,b]. Let $f_1 = V_f$ and $f_2 = V_f - f$. Then $f = f_1 - f_2$. Now f_1 is a real-valued increasing function on $[a, b]$(by Proposition 2.2). It remains to show that f_2 is a real-valued increasing function on $[a, b]$.
Thus let $a \leq x' < x'' \leq b$ and let $P = \{x', x''\}$, a partition of $[x', x'']$. Then we have

$$\begin{aligned} f_2(x'') - f_2(x') &= \{\bigvee_a^b f(x'') - \bigvee_a^b f(x')\} - \{f(x'') - f(x')\} \\ &\geq \bigvee_{x'}^{x''} f - |f(x'') - f(x')| \\ &= \bigvee_{x'}^{x''} f - \bigvee_{x'}^{x''}(f, p) \geq 0 \end{aligned}$$

This shows that f_2 is an increasing function on [a,b]. □

Here since every increasing function is convex, the Jordan decomposition theorem states that every BV function can be written as a difference of two real valued increasing functions (which are convex functions) on a closed and bounded interval $[a, b]$ and vice versa.

Definition 2.4. *A real-valued function f on a closed, bounded interval $[a, b]$ is said to be (absolutely continuous) on $[a, b]$ provided for each $\varepsilon > 0$, there is a $\delta > 0$ such that for every finite disjoint collection $\{(a_k, b_k)\}_{k=1}^n$ of open intervals in (a, b),*

$$if \quad \sum_{k=1}^{n}[b_k - a_k] < \delta, \quad then \quad \sum_{k=1}^{n}|f(b_k) - f(a_k)| < \varepsilon.$$

The criterion for absolute continuity in the case the finite collection of intervals consists of a single interval is the criterion for the **uniform continuity** of f on $[a, b]$. Linear combinations of absolutely continuous functions are absolutely continuous. However, the composition of absolutely continuous functions may fail to be absolutely continuous.

Proposition 2.3. *If the function f is* ***Lipschitz*** *on a closed, bounded interval $[a, b]$, then it is* ***absolutely continuous*** *on $[a, b]$.*

Proof. Let $c > 0$ be a Lipschitz constant for f on $[a, b]$, that is,
$|f(u) - f(v)| < c|u - v|$ for all $u, v \in [a, b]$.
Then, regarding the criterion for the absolute continuity of f, it is clear that $\delta = \varepsilon/c$ responds to any $\varepsilon > 0$ challenge. □

Theorem 2.2. *[7] Let the function f be* ***absolutely continuous*** *on the closed, bounded interval $[a, b]$. Then f is the difference of increasing absolutely continuous functions and, in particular, is of* ***bounded variation****.*

Proof. We first prove that f is of bounded variation. Indeed, let S respond to the $\varepsilon = 1$ challenge regarding the criterion for the absolute continuity of f. Let P be a partition of $[a, b]$ into N closed intervals $\{[c_k, d_k]\}_{k=1}^{N}$, each of length less than δ. Then, by the definition of δ in relation to the absolute continuity of f, it is clear that $TV(f[c_k, d_k]) < 1$, for $1 < k < n$. The additivity formula $V(f_{[a,b]}, P) = V(f_{[a,c]}, P1) + V(f_{[c,b]}, P2)$ extends to finite sums. Hence

$$TV(f) = \sum_{k=1}^{N} TV(f_{[c_k,d_k]}) \leq N.$$

Therefore f is of bounded variation. In view of $f(x) = [f(x)+TV(f_{[a,x])})]-TV(f_{[a,x]})$ for all $x \in [a, b]$, and the absolute continuity of sums of absolutely continuous functions, to show that f is the difference of increasing absolutely continuous functions it suffices to show that the total variation function for f is absolutely continuous. Let $\varepsilon > 0$. Choose δ as a response to the $\varepsilon/2$ challenge regarding the criterion for the absolute continuity of f on $[a, b]$. Let $\{(c_k, d_k)\}_{k=1}^{n}$ be a disjoint collection of open subintervals of (a, b) for which$\sum_{k=1}^{n}[d_k - c_k] < \delta$. For $1 < k < n$, let P_k be a partition of $[c_k, d_k]$. By the choice of δ in relation to the absolute continuity of f on $[a, b]$,

$$\sum_{k=1}^{n} TV(f[c_k, d_k]p_k) < \varepsilon/2.$$

Take the supremum as, for $1 < k < n$, P_k vary among partitions of $[c_k, d_k]$, to obtain

$$\sum_{k=1}^{n} TV(f[c_k, d_k]) < \varepsilon/2 < \varepsilon.$$

We infer from $TV(f_{[a,v]}) - TV(f_{[a,u])} = TV(f_{[u,v]}) > O$ for all $a < u < v < b$ that, for $1 < k < n$, $TV(f_{[c_k,d_k]}) = TV(f_{[a,d_k]}) - TV(f[a, c_k])$. Hence

$$if \quad \sum_{k=1}^{n}[d_k - c_k] < \delta, \quad then \sum_{k=1}^{n} |TV(f_{[a,d_k]}) - TV(f[a, c_k])| < \varepsilon.$$

Therefore the total variation function for f is absolutely continuous on $[a, b]$. □

Theorem 2.3. *A locally Lipschitz function $f : \mathbb{R}^n \to \mathbb{R}$ is d.c. if and only if the subdifferential ∂f of f is of bounded variation at all points $x_0 \in \mathbb{R}^n$.*

Proof. The proof of this theorem comes directly from Theorems 2.2 and 2.1 □

Then since every locally Lipschitz function is differentiable almost everywhere, we can state "a function f is d.c. on $\mathbb{R}$ if and only if f'(defined a.e) is of bounded variation on compact intervals of $\mathbb{R}$".
For functions of several variables a characterization of this kind is not easy to find as each definition of bounded variation for such functions has been introduced by different authors for particular purposes.

Definition 2.5. *A multi-function $F : \mathbb{R}^n \to \mathbb{R}^n$ is said to be of bounded variation at $x_0 \in \mathbb{R}^n$ if there exists maximal cyclically monotone multi-function $M : \mathbb{R}^n \to \mathbb{R}^n$ containing x_0 in the intrior of its domain,such that for some posetive k and neighborhood V of x_0*
$-k\langle x - y, M(x) - M(y)\rangle \leq \langle x - y, F(x) - F(y)\rangle \leq k\langle x - y, M(x) - M(y)\rangle \ \forall x, y \in V$

From the assumptions on M, M is subdifferential of a convex function, finite in the neighborhood of x_0.
The theory of d.c. functions provides us the passage from convex optimization to a more general class of problems including non-differentiable and non-convex ones.

Proposition 2.4. *Every function $f \in C^2(\mathbb{R}^n)$, (a twice continuously differentiable function on $\mathbb{R}^n$), is d.c. on any compact convex set$X$.*

Proof. We show that, given any compact convex set $X \subset \mathbb{R}^n$, the function $g(x) = f(x) + \rho \parallel x \parallel^2$ becomes convex on X when ρ is sufficiently large (then$f(x) = g(x) - \rho \parallel x \parallel^2$ yields a d.c. representation of f). Indeed, since

$$\langle U, \nabla^2 g(x)U\rangle = \langle U, \nabla^2 f(x)U\rangle + \rho \parallel U \parallel^2 \text{ , if } \rho \text{ is so large that}$$
$$-\min\{\langle U, \nabla^2 f(x)U\rangle : x \in X, \|U\| = 1\} \leq \rho$$

then $\langle U, \nabla^2 g(x)U\rangle \geq 0$ for all U, hence $g(x)$ is convex (by Proposition 1.1) □

Proposition 2.5. *For any continuous function f on a compact convex set $D \subset \mathbb{R}^n$ and for any $\varepsilon > 0$ there exists a d.c. function g such that*

$$max_{x\in D}|f(x) - g(x)| \leq \varepsilon.$$

Proof. By Weierstrass theorem there exist a polynomial $g(x)$ satisfying the required condition, and obviously $g(x) \in C^2$. □

Here, from the above proposition one concludes that every continuous function on a convex set, X is the limit of a sequence of uniformly converging DC functions.

Definition 2.6. *A function $f : X \to \mathbb{R}$ defined on $X \subset \mathbb{R}^n$ is said to be* ***locally d.c.*** *if for every $x \in X$ there exist a convex open neighbourhood U of x and a pair of convex functions g and h on U such that $f(x) = g(x) - h(x)$ for all $x \in U$.*

Theorem 2.4. *[14, **?**] A locally d.c. function on a convex, open or closed set X is d.c. on X.*

Proof. We restrict ourselves to the case when X is a compact set (a proof of the general case can be found in Hartman (1959)[6]). From the hypothesis and the compactness of X one can find a finite set $\{x_1, \ldots, x_k\} \subset X$ together with open convex neighbourhoods $X_1, \ldots, X_k$ of these points covering X, and convex functions $h_i : \mathbb{R}^n \to \mathbb{R}(i = 1, 2, ..., k)$ such that $(f + h_i)|X_i$ is convex. let $h = \sum_{i=1}^{k} h_i$ and consider the function $g = f + h$. For each $i, (f + h_i)|X_i$ is convex. Hence $g|X_i = (f + h_i)|X_i + (\sum_{j \neq i} h_j)|X_i$ is convex. g is convex on X. i.e $f = g - h$ with g, h convex. □

From this Theorem, it is evident that local properties lead to global one. That means to check the d.c. property of a function, it is enough to consider a small neighbourhood of a point in its domain.

2.1.1 properties of d.c. functions

Proposition 2.6. *Let $\Omega_1 \subset \mathbb{R}^n, \Omega_2 \subset \mathbb{R}^m$ be convex sets such that Ω_1 is open or closed,Ω_2 is open. If $F : \Omega_1 \to \Omega_2, \quad g : \Omega_2 \to \mathbb{R}^m$ are d.c. mappings, then $g \circ F : \Omega_1 \to \mathbb{R}^m$ is also a d.c. mapping.*

Proof. It suffice to show that $F = (f_1 \ldots f_m) : \Omega_1 \mapsto \Omega_2$ is d.c. and $g : \Omega_2 \mapsto \mathbb{R}^m$ is convex. Then $g(f_1 \ldots f_m) \in DC(\Omega_1)$.
Let $x \in \Omega_1$ and $F(x) = y = (f_1, \ldots, f_m) \in \Omega_2$. Then the convex function $g(y)$ can be represented in a neighbourhood $U_2 \subset \mathbb{R}^m$ of y as pointwise supremum of a family of affine functions: $g(y) = \sup_t l_t(y)$, where $l_t(y) = a_{0t} + a_{1t}y_1 + a_{2t}y_2 \cdots + a_{mt}y_m, y = (y_1, \ldots, y_m)$ are coordinates of y in $\in \mathbb{R}^m$ and $M = sup_{i,t}|a_{it}| < +\infty$. Let $f_i(x) = f_i^+ - f_i^-$, where f_i^+, f_i^- are convex functions in a neighbourhood $U_1 \subset X$ of x such that $F(U_1) \subset U_2$ where $F = (f_1, \ldots, f_m)$ is a mapping.

Then

$$
\begin{aligned}
l_t(f_1, \ldots, f_m) &=_{0t} + a_{1t}f_1 + a_{2t}f_2 \cdots + a_{mt}f_m \\
&= a_{0t} + \sum_{i=1}^{m} a_{it}f_i^+ - \sum_{i=1}^{m} a_{it}f_i^- \\
&= a_{0t} + M\sum_{i=1}^{m} f_i^+ + \sum_{i=1}^{m} a_{it}f_i^+ + M\sum_{i=1}^{m} f_i^- - \sum_{i=1}^{m} a_{it}f_i^- - M\sum_{i=1}^{m} f_i^+ - M\sum_{i=1}^{m} f_i^- \\
&\text{(Adding zero number)} \\
&= [a_{0t} + \sum_{i=1}^{m}(M + a_{it})f_i^+ + \sum_{i=1}^{m}(M - a_{it})f_i^-] - M\sum_{i=1}^{m}(f_i^+ + f_i^-) \\
&= p_t - q,
\end{aligned}
$$

with p_t and q convex and q is independent of t.
Then $g(f_1, \ldots, f_m) = sup_t l_t(f_1, \ldots, f_m) = sup_t(p_t - q) = sup_t p_t - q = p - q$. That is $g(f_1, \ldots, f_m)$ is locally d.c. on Ω_1. Hence by Theorem 2.4, $g \circ F$ is d.c. function.

□

Not that: If the assumptions in proposition 2.6 and theorem 2.4 are violated, a composition of two d.c. functions need not be convex.

Example 2.2. *Let $f : (0, 1) \to [0, 1) : x \mapsto |x - 1/2|$ and $g : [0, 1) \to \mathbb{R} : y \mapsto 1 - \sqrt{y}$ Then gof is not d.c. as it has both left and right derivatives infinite at $1/2$.*
***Note that**: the assumption of openness of [0,1) was not fulfilled in this case, and g is not Lipschitz at zero.*

Proposition 2.7. *If $f_i(x), i = 1, \ldots, m$ are d.c. functions on a convex set X, then each of the following are also d.c. functions on X.*

(i) $\sum_{i=1}^{m} \alpha_i f_i(x)$ *for any real numbers* α_i*;*

(ii) $g(x) = max\{f_1(x), \ldots, f_m(x)\}$*;*

(iii) $h(x) = min\{f_1(x), \ldots, f_m(x)\}$*;*

(iv) The product $f_1.f_2$ *,for d.c. functions* f_1 *and* f_2 *.*

Proof. (i) Since $f_i(x)$ is a d.c. function for all $i = 1, \dots, m$,

$f_i(x) = g_i(x) - h_i(x)$, where g, h are convex functions on X. So, for any constant α_i

$\sum_{i=1}^m \alpha_i f_i(x) = \sum_{i=1}^m |\alpha_i|[g_i(x) - h_i(x)] = \sum_{i=1}^m |\alpha_i| g_i(x) - \sum_{i=1}^m |\alpha_i| h_i(x)$

Hence $\sum_{i=1}^m \alpha_i f_i(x)$ is a d.c. function.

(ii) For $i = 1, \dots m$, let $f_i(x) = g_i(x) - h_i(x)$ with $g_i(x), h_i(x)$ convex on X. Then

$$f_i(x) = g_i(x) + \sum_{i \neq j} h_j(x) - \sum_{j=1}^m h_j(x) \text{ for any } x \in X.$$

Then $\max\{f_1(x), \dots, f_m(x)\} = \max_i \{g_i(x) + \sum_{i \neq j} h_j(x)\} - \sum_{j=1}^m h_i(x)$

which is a difference of two convex functions, as the pointwise maximum and the sum of finitely many convex functions are convex.

(iii) Let $f_i(x) = g_i(x) - h_i(x)$ with $g_i(x), h_i(x)$ convex on X. Then from the equality

$$f_i(x) = \sum_{j=1}^m g_j(x) - \sum_{i \neq j} g_j(x) + h_i(x) \text{ for any } x \in X.$$

Then we have $\min\{f_1(x), \dots, f_m(x)\} = \sum_{j=1}^m g_j(x) - \min_i \{\sum_{i \neq j} g_j(x)\} + h_i(x)\}$

(iv) let $f_1 = g_1 - h_1$ and $f_2 = g_2 - h_2$ are d.c. functions.

Then

$$\begin{aligned} f_1.f_2 &= (g_1 - h_1).(g_2 - h_2) \\ &= (g_1 - h_1).(g_2 - h_2) + 2[(h_1^2 + g_1^2) + (h_2^2 + g_2^2)] \\ &- 2[(h_1^2 + g_1^2) + (h_2^2 + g_2^2)] \text{ (adding zero number)} \\ &= \frac{1}{2}(h_1 + h_2)^2 + (g_1 + g_2)^2 - \frac{1}{2}[(h_1 + g_2)^2 + (h_2 + g_1)^2] \end{aligned}$$

Hence $f_1.f_2$ is a d.c. function if f_1 and f_2 are d.c. functions.

(v) □

Therefore, the set of d.c. functions on X forms a linear space (denoted by $DC(X)$) stable under pointwise maximum and pointwise minimum.

Corollary 2.1. *Let $f(y)$ be a real valued d.c. function defined on a subset of $\mathbb{R}^n$ and let $g_i(x), i = 1, 2, \ldots, n$, be real valued d.c. functions defined on a subset X of $\mathbb{R}^m$. Then the composed function $f(g_1(x), \ldots, g_n(x))$ is also d.c. on X.*

This Corollary follows directly from Proposition 2.6 and Theorem 2.4 .
If f and g are d.c. functions on X then $|f(x)|^{1/m}$ is also d.c. on X. It is followed directly from Corollary 2.1 .
As it is well known, convexity and concavity are not preserved under some simple algebraic operations, i.e. for any strictly convex (or concave) function $f, (-1f)$ is not convex (or concave). Unlike convex and concave functions, DC functions are closed with respect to many operations frequently used in optimization such as scalar multiplication, lower and upper envelope, function composition, product and quotient.
Consider a d.c. function $f \in DC(\mathbb{R}^n)$ defined on $\mathbb{R}^n$. Then since the component functions are convex, f is locally Lipschitz on $\mathbb{R}^n$ and is differentiable almost everywhere. Moreover, f is directionally differentiable on $\mathbb{R}^n$ and

$$f'(x; d) = g'(x; d) - h'(x; d) \text{ for all x and d in } \mathbb{R}^n$$

Therefore the directional derivative of f is also a difference of two positively homogeneous convex functions for whatever the decomposition of f as a difference of two convex functions, g and h. Hence, every d.c. function f is quasidifferentiable on $\mathbb{R}^n$.But there are quasidifferentiable functions which are not d.c. functions. For example, every differentiable function is quasidifferentiable while a little more is required on the derivative of the function to be d.c. as it is mentioned in Jordan decomposition Theorem.

2.2 D.C. Sets

This section provides the definitions of d.c. set and d.c. inequality, then shows that all sets defined by a d.c. inequality in $\mathbb{R}^n$ can be seen as the projection of a d.c. set in $\mathbb{R}^{n+1}$. In particular, this section provides several ways to define closed sets by d.c. inequalities.

Definition 2.7. *A set $M \subseteq \mathbb{R}^n$ is called a d.c. set if it can be described in the form $M = D \backslash C$ where D and C are two convex sets in $\mathbb{R}^n$.*

A d.c. set can be described as a difference of two convex sets,while a d.c. function can be described as a difference of two convex functions. Then a question arises: given any real-valued d.c. function f, is there any relationship between the set $\{x \in \mathbb{R}^n | f(x) \leq 0\}$ and d.c. sets? The answer is positive. Then it is natural to introduce d.c. inequalities.

Definition 2.8. *An inequality of the form $f(x) \leq 0$ is convex when f is convex. If f is concave, then this inequality is called reverse convex. If f is d.c., then this inequality is called d.c. inequality.*

It should be noted that, when f is d.c., $f(x) \geq 0$ is still a d.c. inequality since $-f$ is also d.c. Proposition 2.7 guarantees that d.c. functions are closed under lower and upper envelopes. Thus any finite system of d.c. inequalities, whether conjunctive or disjunctive, is equivalent to a single d.c inequality. Furthermore, by introducing an additional variable t, the d.c. inequality $p(x) - q(x) \leq 0$ where p and q are convex can be split into two inequalities: $p(x) - t \leq 0, \quad t - q(x) \leq 0$, where the first is a convex inequality and the second is a reverse convex inequality. By this property, we see that the set of constraints in the nonsmooth optimization problem can written as a set which is d.c. inequality and any DC optimization problem can be reduced to a canonical form.

Chapter 3

Difference of Subharmonic Functions

3.1 Subharmonic Functions

A function f is said to be convex if the affine lines joining any two points of the graph of f lie above the graph of f between these two points. If the affine lines say u replaces by a more general property: $\Delta u = \sum_{i=1}^{n} \frac{\partial^2 u}{\partial x_i^2} = 0$, (Note that all affine functions satisfy this property trivially) then we get a more broader class of functions than that of convex.

Definition 3.1. *A function $f(x)$ defined in a set E in $\mathbb{R}^m$ is said to be upper semi-continuous(u.s.c) on E if*

1. $-\infty \leq f(x) < \infty, \quad x \in E$
2. *The sets* $\{x | x \in E, \quad f(x) < a\}$ *are open in E for* $-\infty < a < +\infty$.

A function $f(x)$ is said to be lower semi-continuous (l.s.c) in E if $-f(x)$ is u.s.c. in E. If $f(X)$ is both u.s.c. and l.s.c., then $f(x)$ is said to be continuous.

Definition 3.2. *Let $D \subset \mathbb{R}^n$ be a connected subset of (or equivalently a domain in) $\mathbb{R}^n$. A function u, from D to $\mathbb{R}$ is called **harmonic** if and only if*

i. *u is continuous.*

ii. *For any compact set, $K \subset D$, there is an $\varepsilon > 0$ such that for all $x \in K$, $B_x^\varepsilon = \{y/|y-x| < \varepsilon\} \subset D$ and*

$$u(x) = \int_{S^{n-1}} u(x + r\omega) d\omega \quad \forall r \in (0, \varepsilon). \tag{3.1}$$

Here S^{n-1} is the unit sphere in $\mathbb{R}^n, S^{n-1} = \{x/|x| = 1\}$ and $d\omega$ is the normalized spherical measure, that is $\omega(S^{n-1}) = 1$ and ω is the unique rotation invariant measure.

Equivalently, a $C^2(\mathbb{R}^n)$ function $h : D \mapsto \mathbb{R}$ is called harmonic in D if $\Delta h = 0$, where Δ denotes the Laplacian operator.

Definition 3.3. *A function $u(x)$ defined in a domain $D \subset \mathbb{R}^n$ is said to be* ***subharmonic(s.h)*** *function in D if*

i. *u is u.s.c.*

ii. *$-\infty \leq u(x) < +\infty$ in D.*

iii. *If x_0 is any point of D, then there exist arbitrary small positive values of r such that*

$$u(x_0) \leq \frac{1}{c_n r^{n-1}} \int_{s(x_0,r)} u(x) d\sigma(x),$$

where $d\sigma(x)$ denotes surface area on $s(x_0, r)$.

On $\mathbb{R}^1$ subharmonicity is identical with convexity. But on $\mathbb{R}^n, n \geq 2$ every linear function is harmonic and every convex function defined on a convex domain D is subharmonic in D. There are also some nonconvex and nonconcave functions defined on $\mathbb{R}^n, n \geq 2$ which are subharmonic. [14, 17]

3.1.1 Some properties of subharmonic functions

Proposition 3.1. 1. *If $u_i : D \mapsto \mathbb{R} \cup \{-\infty\}$ be subharmonic for all $i = 1, \ldots, k$ and $\lambda_i \geq 0$ be numbers. Then $u_0 = \sum_{i=1}^k \lambda_i u_i$ is also subharmonic on D.*

2. *The suprimum of a finite collection of subharmonic functions is subharmonic on D.*

3. *If $u \in C^2(D)$ denotes the set of functions in D where up to second order derivative are continuous, then u is subharmonic in D if and only if $\triangle u \geq 0$ in D.*

4. *u is continuous subharmonic function in D if and only if u is the uniform limit of functions u_i with $\triangle u_i \geq 0$ in D.*

5. *Let c be a constant and u is a subharmonic function in D. If $u \leq c$ in D and there exists $x_0 \in D$ such that $u(x_0) = c$ then $u \equiv c$ in the whole of D.*

Proof. See [17] □

Theorem 3.1. *(Riesz theorem) Suppose that u is subharmonic and not identically $-\infty$ in a domain $D \subset \mathbb{R}^n$. Then there exists a unique (non-negative) Borel-measure μ in D such that for any compact subset E of D*

$$u(x) = \int_E K(x-\zeta)d\mu(\zeta) + h(x),$$

$$where \quad K(x) = \begin{cases} \log|x|, & \text{if } n = 2, \\ -|x|^{2-n} & \text{if } n > 2 \end{cases}$$

and $h(x)$ is a harmonic function in the interior of E.

Proof. See [17] on page 112-113. □

3.2 Difference of Subharmonic Functions

Definition 3.4. *A function w is called a **difference of two subharmonic functions (d.s.h.)** u and v on an open connected set $d \subset \mathbb{R}^n$ if either u or v is finite in D and $w = u - v$ holds in D.*

Since this representation is not unique and depends on the particular choice of functions u and v, we denote the representation $w = u - v$ by an ordered pair of two subharmonic functions(u, v).

such functions are called δ-subharmonic. As an application to Riesz-Representation theorem we can consider a quasidifferebtiable function f. Then the directional derivative of f as a function of direction is a **difference of two sublinear functions.** Recall that sublinear functions are **convex** and hence **subharmonic.** Thus, there are two Borel measures μ_1 and μ_2 on $\mathbb{R}^n$ such that

$$f'(x_0, v) = \int_E K(v-\zeta)d\mu_1(\zeta) - \int_E K(v-\zeta)d\mu_2(\zeta) + h(v, E),$$

Where h is a harmonic function in the interior of E for any compact subset E of $\mathbb{R}^n$. That means, the directional derivative of a quasidifferentioable function can be considered as a d.s.h. function as a function of direction. It turns out that it is worthly to investigate the features of d.s.h. functions in light of its uniqueness and minimal representation. The following theorem is fundamental in defining a 'better' representation from the class of representations of d.s.h. functions. It states that every d.s.h. function induces a unique signed Borel-measure and conversely.

Theorem 3.2. *Let $w : D \mapsto \mathbb{R}$ be a d.s.h. function defined on a domain $D \subset \mathbb{R}^n$. then w induces a unique signed Borel-measure μ. Conversely, if μ is a signed Borel-measure on D, then there corresponds a function $w : D \mapsto \mathbb{R}$ which is representable as a difference of two functions u and v which are subharmonic on D.*

Proof. See [2] □

Definition 3.5. *Let $w : \mathbb{R}^n \mapsto \mathbb{R}$ be a d.s.h. function on $\mathbb{R}^n$. A pair (u^*, v^*) of two functions, subharmonic on $\mathbb{R}^n$ is called a* ***canonical representation*** *for w if $w = u^* - v^*$ and for the Jordan-decomposition μ^+ and μ^- of μ, u^* is induced by μ^+ and v^* is induced by μ^-.*

The unique signed Borel-measure μ of w has a positive part and a negative part, denoted by μ^+ and μ^- respectively. To find these two positive measures, we proceed as follows.

Let $w(x) = u(x) - v(x)$ where u and v are subharmonic in a domain $D \subset \mathbb{R}^n$. Then from Riesz representation theorem, we have that for E a compact subset of D,

$$w(x) = h(x) + \int_E K(x-\zeta)d\mu(\zeta),$$

where h is harmonic in the interior of E and μ is a difference of two positive measures on subsets of D whose closures are compact in D.

Note: The sum and difference of two harmonic functions is again harmonic. Now define $\mu^+(E) = \sup \mu(E)$, where the supremum is taken over all Borel subsets of E. Then clearly $\mu^+(E) \geq 0$, That is μ^+ is positive measure. Next to describe μ as a difference of two positive measures on D we write

$$\mu^-(E) = \mu^+(E) - \mu(E) \tag{3.2}$$

Ones again $\mu^-(E) \geq 0$.

Equation (3.2) is uniquely determined by μ and is independent of the choice of the representation $w = u - v$ of w.

If w is a C^2 function in a domain $D \subset \mathbb{R}^n$, then w is a d.s.h. function determined by the unique signed Borel-measure

$$\mu(E) = \frac{1}{e_n} \int_E \Delta w dx,$$

where $e_2 = 2\pi$ and $e_n = \frac{4\pi^{n/2}}{\Gamma(n/2-1)}$, $n \geq 3$, dx denotes a volume in $\mathbb{R}^n$ and $\Delta w = \sum_{i=1}^n \frac{\partial^2 w}{\partial x_i^2}$ is a Laplacian in which case

$$\mu^+(E) = \frac{1}{e_n} \int_E (\Delta w)^+ dx \text{ and } \mu^-(E) = \frac{1}{e_n} \int_E (\Delta w)^- dx,$$

where $(\Delta w)^+ = \max(\Delta w, 0)$ and $(\Delta w)^- = \max(-\Delta w, 0)$.

Thus the existence of a canonical representation of any d.s.h. function w follows immediately.

Moreover, following Riesz representation theorem to subharmonic functions, in a canonical representation (u^*, v^*) of w, u^* and v^* are unique to within a common additive harmonic function. Hence we have

Corollary 3.1. *Every d.s.h. function has a canonical representation and this representation is unique up to a common additive harmonic function.*

Proof. See Theorem 4. on [2] □

Now we characterize canonical representations using subharmonic functions instead of signed Borel measure.

Theorem 3.3. *Let $w : D \mapsto \mathbb{R}$ be a d.s.h. function on $D \subset \mathbb{R}^n$. A representation (u^*, v^*) of w is canonical if and only if to every representation (u, v) of w there corresponds a function s subharmonic on D such that $u = u^* + s$ and $v = v^* + s$.*

Proof. See Theorem 5. on [2] □

Remark 3.1.

1. *There exist bounded subharmonic functions in every dimension $n \geq 2$.*

2. *The only functions which are subharmonic and bounded from above in the plane are constants. However if $n > 2$, the unique subharmonic function with Riesz measure μ (satisfying a certain condition) in $\mathbb{R}^n$ and least upper bound C is given by*

$$u(x) = C - \int \frac{1}{|x - \zeta|^{n-2}} d\mu(\zeta)$$

Thus if u is a function which is subharmonic, positively homogeneous and bounded above, then u is a zero function.

Chapter 4

Minimal(Better) representation of D.C. functions

4.1 Minimal of D.C. Functions

Given a d.c. function f, it has many ways of representations in which it can be written as a difference of two convex functions. From these decompositions, choosing a *better* one is useful in describing optimality conditions for d.c. optimization.
Let us consider a decomposition of f as

$$f = g - h = (g + \varphi) - (h + \varphi,)$$

where g and h are convex and φ is strictly convex.
This non-unique way of representing f has its own advantage in that one can describe g and h by adding more structure to them. On the other hand in determining the optimality conditions, the particular decomposition of f as a difference of two convex functions plays a great role. So we have to define what to mean by "better" decomposition of d.c. function f as difference of two convex functions.

Definition 4.1. *A pair of convex functions g_1 and h_1,written (g_1, h_1) on a convex domain $D \subset \mathbb{R}^n$ is* ***equivalent*** *to a convex pair (g_2, h_2) of functions defined on the same domain D (denoted by $(g_1, h_1) \sim (g_2, h_2)$) if and only if there are decompositions of the* *same d.c. function, say f in D.*

$$i.e. \quad iff \quad f = g_1 - h_1 = g_2 - h_2 \quad on \quad D$$

Clearly the above relation defines an equivalence relation on the set of pairs of convex functions on a convex domain D. Denote by $[g, h]$ the equivalence class determined by the pair (g, h) of convex functions. To define minimality we need to compare two functions. Therefore, all the inequalities between any two functions are related to the point-wise ordering.

Definition 4.2. *[14]*

(i) A representation $g_0 - h_0$ of a d.c. function f is said to be ***minimal*** *at a point x_0 if for any pair $(g, h) \in [g_0, h_0]$, the inequalities $g(x_0) \leq g_0(x_0)$ and $h(x_0) \leq h_0(x_0)$ imply $g(x_0) = g_0(x_0)$ and $h(x_0) = h_0(x_0)$.*

(ii) A representation $g_0 - h_0$ of a d.c. function f is said to be ***a minimal representation of a d.c. function f*** *on the domain D if and only if for any pair (g,h) the relations $g(x) \leq g_0(x), h(x) \leq h_0(x), \quad \forall x \in D$ and $g_0 - h_0 = g - h$ imply $g_0 = g, \quad h_0 = h$, where the inequalities involved are taken as point-wise ordering.*

Proposition 4.1. *if $(g, h) \sim (g_0, h_0)$ and $g_0 \leq g$ (where the ordering $\leq$ is point-wise ordering), then there exists a non-negative d.c. function s such that $g = g_0 + s$ and $h = h_0 + s$.*

Proof. Let $(g, h) \sim (g_0, h_0)$. Then by definition $g_0 - h_0 = g - h$.
Since $g_0 \leq g$ implies there exists a non-negative d.c. function s such that $g_0 + s = g$. thus we have $g_0 - h_0 = g_0 + s - h$ and from this we get $h = h_0 + s$. □

From proposition 4.1 one can drive the following corollary.

Corollary 4.1. *Suppose (g_0, h_0) is a minimal representation of a d.c. function f. then for any $(g, h) \sim (g_0, h_0)$ if $g_0 + s = g$ or $h_0 + s = h$ with s non-negative d.c. function, then s is the zero function.*

Theorem 4.1. *Let the sequence of convex functions $g_k : \mathbb{R}^n \mapsto \mathbb{R}$ converge pointwise for $k \to +\infty$ to $g : \mathbb{R}^n \mapsto \mathbb{R}$. Then g is convex and for each compact set S, the convergence of g_k to g is uniform on S.*

Proof. $[Step-1]$ First the function $h := \sup_k g_k$ is convex and $g(x) < +\infty$ for all x and because the convergent sequence $(g_k(x))_k$ is certainly bounded. Hence g is continuous and therefore bounded, say by M, on the compact set $B(0, 2r)$:

$g_k(x) \leq h(x) \leq M$ for all k and all $x \in B(0, 2r)$.

Second, the convergence sequence $(g_k(0))_k$ is bounded from below.

$\mu \leq g_k(0)$ for all k.

Then, for $x \in B(0, 2r)$ and all k, use convexity on $[-x, x] \subset B(0, 2r)$:

$$2\mu \leq 2g_k(0) \leq g_k(x) + g_k(-x) \leq g_k(x) + M$$

i.e the $g_k's$ are bounded from below independently of k. Thus, there is some L (independent of k) such that

$$|g_k(y) - f_k(y')| \leq \|y - y'\| \quad \text{for all } k \text{ and all } y, y' \text{ in } B(0, 2r).$$

Naturally, the same Lipschitz property is transmitted to the limiting function g.

$[Step-2]$ Now fix $\varepsilon > 0$. Cover S by the balls $B(x, \varepsilon)$ for x describing S, and extract a finite covering $S \subset B(x_1, \varepsilon) \cup \cdots \cup b(x_m, \varepsilon)$. With x arbitrary in S,take an x_i such that $x \in B(x_i, \varepsilon)$. There is $k_{i,\varepsilon}$ such that for all $k \geq k_{i,\varepsilon}$,

$$|g_k(x) - g(x)| \leq |g_k(x) - g_k(x_i) - g(x_i)| + |g(x_i) - g(x)| \leq (2L + 1)\varepsilon$$

Knowing that x and x_i in $S \subset B(0, r)$, the above inequality is then valid uniformly in x, providing that

$k \geq \max\{k_{1,\varepsilon}, \ldots, k_{m,\varepsilon}\} := k_\varepsilon$ □

For detail proof of this theorem see on [8] page 105 Theorem 3.1.4

Theorem 4.2. *[14] let f be a d.c. function whose component functions belong to Ω (the set of all proper, convex and lower semi-continuous functions). Then f has a minimum representation on some convex domain.*

Proof. Suppose $f = u - v$ where $u, v \in \Omega$. If (u, v) is minimal, then there is nothing to proof. Assume (u, v) is not minimal. Then there are $u_1, v_1 \in \Omega$ such that

$$u - v = u_1 - v_1 \text{ with } u_1 \leq u \text{ and } v_1 \leq v$$

If again (u_1, v_1) is not minimal, we can find $u_2, v_2 \in \Omega$ such that $u_1 - v_1 = u_2 - v_2$ with $u_2 \leq u_1 \leq u$ and $v_2 \leq v_1 \leq v$. Continuing this process we can find sequences of functions $\{u_i\}$ and $\{v_i\}$ satisfying the following properties.

1. $u_i, v_i \in \Omega$ for all i, j.
2. $u_i \leq u_k$ for $i \geq k$ and $v_i \leq u_l$ for $i \geq l$.
3. $u_i - v_i = u_j - v_j$ for all i, j.

Since by the assumption both the sequences $\{u_i\}$ and $\{v_i\}$ are pointwise bounded from below they converge pointwise, say to u_0 and v_0 respectively. Then by Lemma 4.1 above u_0 and v_0 are convex. Moreover, since $u - v = u_i - v_i \quad \forall i$ we can pass to the limit to get that $u - v = u_0 - v_0$. Hence $(u, v) \sim (u_0, v_0)$.
Therefore, (u_0, v_0) is a minimal representation of the d.c.function f.[14] □

On the other side another researcher V.A. Zalgaller [16] defines and find minimal representation of a function f as a d.c. functions as follows.

Definition 4.3. *Let D be aconvex compact set in $\mathbb{R}$ having interior points. We say that a function $f : D \mapsto \mathbb{R}$ admits a minimal representation if there exist convex functions h and g on D such that $f = g - h$ and h is bounded from above.*

Denote the class of functions having the property described in the above definition by ϕ. For a function $f \in \phi$, there exists infinitely many representations $g_\xi - h_\xi$. By adding the same constant to both the functions g_ξ and h_ξ, we may assume that in any representation of this form we have $h_\xi \leq 0$ and $\sup_{a \in D} h_\xi(a) = 0$.
Now introduce the functions g and h as follows.

$$g(a) = \sup_\xi g_\xi(a), \quad h(a) = \sup_\xi h_\xi(a), \quad a \in D$$

and the representation $f = g - h$ obtained is called the minimal representation of f by difference of convex functions.
The above-mentioned process gives no information on inner properties of the function f that are necessary and sufficient for f to belong to ϕ.
Let $D \subset \mathbb{R}$ be a convex compact set having interior points. If $f : D \mapsto \mathbb{R}$ is a continuous function, then f is bounded from below and from above and the following process can be realized. We set $f_1 = f$ and construct two sequence of functions:

$$r_i := \overline{f}_i - f, \quad f_{i+1} := \overline{r}_i + f$$

for any fixed $a \in D$, the values of f_i and $\overline{f}_i$ form a monotone sequence:

$$f_1(a) \geq \overline{f}_1 \geq f_2(a) \geq \overline{f}_2 \geq f_3(a) \geq \overline{f}_3 \geq \dots$$

Hence for any $a \in D$ the following value (finite or infinite) is defined.

$$g^*(a) = \lim_{i \to \infty} \overline{f}_i(a) \tag{4.1}$$

Theorem 4.3. *If there exists a point at which the function $g^* : D \mapsto \mathbb{R} \cup (-\infty)$ defined by (4.1) takes the value $-\infty$, then $f \notin \phi$. If the function g^* is finite at every point, then $f \in \phi$, and the function $g \to g^*$ determines the minimal representation $f = g - h$ by a difference of convex functions.*

Proof. [16] □

These two different approaches by different researchers tend to the same idea which is to characterize the d.c.functions which admits a minimal decomposition. And these approaches defined have their own limitations for all class of d.c. functions to admit minimal decomposition. But the following section strengthen those limitations.

4.2 Application of D.S.H.F to Minimal representations of D.C. functions

Returning back to the minimal representation of d.c. functions, we can see that since every convex function is subharmonic, every d.c. function is a d.s.h. function. therefore,

we can apply the results of d.s.h functions to investigate minimal representations of d.c. functions on $\mathbb{R}^n$.Again on this section we use the pointwise ordering of two functions to order the functions themselves.
The following corollary follows as a consequence of Theorem 3.3 and remark 3.1.

Corollary 4.2. *for any representation (u, v) of a d.c. function f, if both u and v are sublinear, then a canonical representation (p, q) of f is minimal.*

Proof. Let f be a d.c. function. Since every d.c. function is d.s.h. function, f has a canonical representation, say (p, q). Suppose (u, v) be any representation of f such that $u \leq p$ and $v \leq q$. Then by Theorem 3.3 there exists a uniquely defined subharmonic function s such that

$$u = p + s \text{ and } v = q + s \text{ with } s \leq 0$$

Since both u and p (resp. v and q) are sublinear by the assumption, it follows that s is positive homogeneous. Moreover, s is a subharmonic function which is bounded from above. Hence by (2) of Remark 3.1 s is a zero function. i.e. (p, q) is minimal. □

To sharpen the above corollary to a more general class of d.c. functions, we need the following observation first.
Let (u, v) be a canonical representation of a d.c. function f (the existence of which assumed by corollary 3.1). Then for any d.c. representation (g, h) of f there exists a subharmonic function s such that

$$g = u + s \text{ and } h = v + s$$

Since both g and h are convex,one can convince that each of the functions u, v and s are d.c. . Hence canonical representations of d.c. functions and their d.c. representations are not always comparable. But if the components of every canonical representation of a d.c. function is convex, then we get a relationship between minimality of a d.c. and canonical representations. But first we need to proof the following Lemma.

Lemma 4.1. *If (g, h) is a minimal representation of a d.c. function and $g = u + s$ with u and s convex functions, then (u, h) is also minimal.*

Proof. Let $(g_1, h_1) \sim (u, h)$ be such that $g_1 \leq u$ and $h_1 \leq h$. Then by definition $g_1 - h_1 = u - h$ or $g_1 + h = u + h_1$. Adding g to both sides, we get $g + g_1 + h = g + u + h_1$. But since $g = u + s$, we have $u + s + g_1 + h = g + u + h_1$ and this is equivalent to $(s + g_1, h_1) \sim (g, h)$. But since $g_1 \leq u$ implies $g_1 + s \leq u + s = g, \quad h_1 \leq h$ and (g, h) is minimal, we have $h_1 = h$. Hence from the relation $g_1 - h_1 = u - h$ it follows that $g_1 = u$, that is, (u, h) is minimal. □

Theorem 4.4. *[14] Let the components of every canonical representation of the function f be convex functions. Let (u, v) be a canonical representation of a d.c. function f and $(g, h) \sim (u, v)$ be any d.c. representation of f. If in the expression*

$$g = u + s, \text{ and } h = v + s$$

the subharmonic function s is convex for all $(g, h) \sim (u, v)$ then (u, v) is minimal.

Proof. Suppose (g, h) is a minimal representation of f. Then since $(g, h) \sim (u, v)$ and (u, v) is canonical, there is a subharmonic function s such that $g = u + s$ and $h = v + s$. Then since (g, h) is minimal and $g = u + s$,invoking Lemma 4.1 gives us (u, h) is minimal. Ones again the minimality of (u, h) and the relation $h = v + s$ imply the minimality of (u, v).
Therefore (u, v) is minimal. □

Some times it may be easier to check minimality at a point as it deals with only a set of pairs of real numbers. Therefore, the following Theorem characterizes minimality of a canonical representation on the whole domain with its minimality at a point in the domain.

Theorem 4.5. *[14] A canonical representation of any d.c. function is minimal if it is minimal at least at one point in its domain.*

Proof. Suppose (u, v) is a canonical representation of any d.c. function f and it is minimal at x_0. Then by definition for any $(g, h) \sim (u, v)$, the inequalities $g(x_0) \leq u(x_0)$ and $h(x_0) \leq v(x_0)$ imply $g(x_0) = u(x_0)$ and $h(x_0) = v(x_0)$.

Suppose $(g, h) \sim (u, v)$, $\quad g(x) \leq u(x)$ and $h(x) \leq v(x)$ for any $x \in \mathbb{R}^n$. Then there exists a subharmonic function s such that

$$g = u + s, \text{ and } h = v + s$$

But since $g(x) \leq u(x)$ and $h(x) \leq v(x)$ for any $x \in \mathbb{R}^n$, we have $s \leq 0$. again since (u, v) is minimal at a point x_0 and $u(x_0) = g(x_0) = u(x_0) + s(x_0)$ we have $s(x_0) = 0$. Then $s \equiv 0$ follows from (5) of Proposition 3.1.
Therefore, the pair (u, v) is minimal. □

Here since every convex function is subharmonic function, every d.c. function can be written as a d.s.h. functions. But there is no guaranty that every subharmonic functions are convex. But Theorem 4.4 and Theorem 4.5 concludes that every canonical representation of any d.c. function is minimal at least at one point in its domain.

Conclusion

Many optimization problems in the real world are modelled as a continuous nomsmooth functions. And such functions are so difficult to solve using the classical methods in convex analysis as well as to convexify these nonconvex functions with nonsmooth data. But these functions with nonsmooth data can be expressed as a family of d.c. functions and admits all the operations in convex analysis and other operations like product, finite min/max which are not functional in the convex analysis.

So studying the properties and nature of the family of d.c. functions play a critical role in the study of global optimization.

Especially the researchers which are interested in studying problems related with polynomials in several variables, variational analysis, non-cooperative game theory, spectral theory as well as operator theory and related problems must be study to ease their tasks.

Bibliography

[1] Andrew Eberhard, Nicolas Hadjisavvas and Dinh The Luc, *Nonconvex Optimization and its Applications: Proceedings of the International Symposium on Generalized Convexity and Generalized Monotonicity*, Springer Volume 77, 2005.

[2] Arsove, M.G., *Functions Representable as Differences of Subharmonic Functions,* Trans.Amer.Math.Soc.75(1953),327-365.

[3] Demyanov, V.F and Rubinov, A.M., *Constructive Nonsmooth Analysis,* Verlage Peter Lang,1995.

[4] Demyanov, V.F and Rubinov, A.M., *On Quasidifferentiable Mappings,* Math. Operationsforschung and Statistick, Ser.Optim.14(1983).3-29.

[5] Eligius M.T. Hendrix and Boglarka G.-Toth, *Introduction to Nonlinear and Global Optimization*, Springer, Volume 37, 2010.

[6] Hartman,P., *On Functions Representable as a Difference of Convex functions*, Pacific J.Math.9 (1959),707-713.

[7] H.L.Royeden, P.M. Fitzpatrick, *Real Analysis*, China Machine press, Fourth edition, 2010.

[8] Jean-Baptiste Hiriart Urruty, Claude Lemaréchal, *Fundamentals of Convex Analysis*, Springer-Verlage Berlin Heldberg 2001.

[9] J Yeh, *Real Analysis: Theory of Measure and Integration*, Second Edition, World Scientific Publishing co.pte.ltd., 2006.

[10] Le Thi Hoai An and Pham Dinh Tao, *DC Programming and DCA for Nonconvex Optimization: Theory, Algorithms and Applications*, MAMERN09: 3rd International Conference on Approximation Methods and Numerical Modelling in Environment and Natural Resources Pau (France), June 8-11, 2009.

[11] Nicolas Hadjisavvas ,Sandor K. and Siegfried S., *Nonconvex Optimization and Its Applications: Handbook of Generalized Convexity and Generalized Monotonicity*, Springer,Volume 76 , 2005.

[12] Pham Dinh Tao and Thi Hoai An, *Convex Analysis Approach to D.c. programming; Theory, Algorithms and Applications*, Acta Mathematica Vietnamica Volume 22, Number 1, 1997, p. 289-355.

[13] Qinghua Zhang, *Outer Approximation Algorithms for DC programs and beyond*,Ph.D. Thesis,Università di Pisa Dipartimento di Matematica L.Tonelli Scuola di Dottorato.

[14] Semu M. *On minimal Pairs of Compact Convex sets and of Convex functions*, PhD Thesis,Addis Ababa University Departement of mathematics, Karlsruhe 2002

[15] Tuy, H.,*Convex analysis and Global Optimization*, Kluwer Academic publishers, Dordrecht,1998.

[16] V. A. Zalgaller, *Representation of Functions of Several Variables by Difference of Convex Functions*, Journal of Mathematical Sciences, Vol. 100, No. 3, 2000.

[17] W.K.Hayman, F.R.S and P.B.Kennedy, *Subharmonic Functions*, Academic Press Volume 1, 1976.