University of Alberta

The Quadratic Sub-Lagrangian of Prox-Regular Functions

by

Warren L. Hare ©

A thesis
submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of

Master of Science

in

Mathematics

Department of Mathematics

Edmonton, Alberta

Spring 2000

0-612-60128-5

University of Alberta

Library Release Form

Name of Author: Warren L. Hare

Title of Thesis: The Quadratic Sub-Lagrangian of Prox-Regular Functions

Degree: Master of Science

Year this Degree Granted: 2000

Warren L. Hare
7350 Bendigo Circle
Mississauga, Ontario
L5N 1Z5

DATE: 2000-4-4

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled

The Quadratic Sub-Lagrangian of Prox-Regular Functions

submitted by **Warren L. Hare** in partial fulfillment of the requirements for the degree of **Master of Science** in **Mathematics**.

Dr. R. Poliquin, Supervisor

Dr. H. Van Roessel

Dr. J. Tuszynski

DATE: 2000 - 3 - 30

Abstract

Given a finite valued convex function, f, the U-Lagrangian, as defined by Lemaréchal, Oustry, and Sagastizábal, provides an envelope of f that maintains these properties. By selecting the appropriate UV decomposition on $\Re^n$ one can also ensure that the U-Lagrangian is not only continuous at the origin, but differentiable there. This thesis extends these properties to prox-regular functions by creating a new envelope which we call the Quadratic Sub-Lagrangian. In further exploration of this envelope it will be demonstrated how to use a quadratic expansion of the Quadratic Sub-Lagrangian to create a quadratic expansion for the original function. Finally some properties that guarantee such an expansion for the Quadratic Sub-Lagrangian will be developed.

Acknowledgements and Dedication

To **Dr. H. Bell**,
for making me think it would be possible;

To **Dr. R. Kerman**,
for making me think it would be interesting;

and, To **Dr. R. Poliquin**,
for making both these things true.

I would also like to express my deep gratitude to the Natural Sciences and Engineering Research Council of Canada and the University of Alberta for the financial support they have given me throughout my studies. Finally I would like to thank Dr. Tuszynski and Dr. Van Roessel, who dedicated much of their time and energy to the improvement of this thesis.

Contents

Chapter 1

Introduction and Background

Si quid calumnietur levinus esse quam decet theologum,
aut mordacius quam deceat Christianum –
non Ego, sed Dionysos dixit. [1]

Laurence Sterne

1.1 Introduction

The theory of Optimization may be said to have begun in late 1620's, when Fermat solved the problem of finding the maximum of xy given the constraint condition that $x + y = 10$. Amazingly he accomplished this twenty years before either Newton or Leibniz were born (1642 and 1646 respectively), and sixty years before Calculus was first presented to the public view. In fact Fermat died in 1665, almost twenty years before Leibniz first published his new methods for solving maxima and minima problems in 1684 ([9], pp. 396, 430, 461, 473, 477).

Unlike in Fermat's time, the study of optimization today has a fully developed field of Calculus to support it. Therefore the question of optimizing a differentiable function is mote; so, we turn our attention to non-differentiable functions. Since the area of differentiable functions is so well developed one of the primary goals in non-differentiable optimization is to determine ways of estimating non-differentiable functions via differentiable ones. We call

[1] If anyone falsely accuses that this is more light-hearted than becomes a theologian, or more biting than becomes a Christian — not I, but Dionysos wrote it.

such estimations envelope functions and in this thesis we will focus on the development of a new one.

Until recently, much of the research in optimization was directed towards convex functions. The reasons for this are plentiful, but largely based on the fact that convex functions usually obtain their minimum. The drawback of convex functions is that they are not dense in the space of measurable functions. This prompted the development of a new set of functions, known as prox-regular (see Section 1.4 below), which are dense in the measurable function space. In this thesis we shall generalize some the results of envelopes of convex functions to the broader set of functions known as prox-regular.

1.2 Basic Definitions

Before approaching the subject of this thesis it is prudent to ensure that some basic background material is covered.

1.2.1 Lower Semi-Continuity, and Proper Functions

We assume that the reader is familiar with the greatest lower bound axiom of the extended real numbers, and the definition of $\inf_{x \epsilon X} f(x)$. From here we begin by defining the liminf and limsup of a function f at the point $\bar{x}$ as:

$$\liminf_{x \to \bar{x}} f(x) := \lim_{\tau \searrow 0} [\inf_{x \epsilon B_\tau(\bar{x})} f(x)],$$

$$\limsup_{x \to \bar{x}} f(x) := \lim_{\tau \searrow 0} [\sup_{x \epsilon B_\tau(\bar{x})} f(x)].$$

Although we shall deal mostly with liminf it will be important in several circumstances to know limsup as well. Most importantly we note that $\liminf f = -\limsup(-f)$, so most statements on liminf can be easily inverted to statement on limsup.

It is clear that the liminf of f at $\bar{x}$ is always less than or equal to $f(\bar{x})$ since $\bar{x} \epsilon B_\tau(\bar{x})$ for all $\tau > 0$. Thus we are lead to our first important definition, lower semi-continuity (or lsc). We call a function lower semi-continuous at $\bar{x}$ if

$$\liminf_{x \to \bar{x}} f(x) = f(\bar{x}),$$

and say the function is lower semi-continuous if this holds for all $\bar{x} \epsilon \Re^n$. Conversely upper semi-continuity (usc) corresponds to $\limsup_{x \to \bar{x}} f(x) = f(\bar{x})$.

It can be shown that a function is continuous if and only if it is both upper and lower semi-continuous ([8] Example, 1.12). Since the study of optimization generally focuses on the achievement of minimums, lower semi-continuity will play a much larger role than that of upper semi-continuity.

Next we call a function, f, proper if f is not constantly infinity and $f(x) \neq -\infty$ for all $x \in \Re^n$. That is $\inf f$ is a real number when examined over any compact set. The study of Optimization often focuses on proper lower semi-continuous functions because of the remarkable fact that these functions always obtain their infimum when examined over any compact set which intersects their domain ([8], Corollary 1.10).

Justifying the examination of proper functions is not difficult, as if the function is not proper one finds $\inf_{|x| \leq \rho} f(x)$ is not a real number whenever ρ becomes sufficently large. To justify the focus on lower semi-continuous functions, the concept of lower semi-continuity shall be examined a little more closely.

For any function, f, the epi-graph of f is the set of all points lying on or above the graph of f. More rigorously we define the epigraph of f by

$$epi(f) := \{(x, \alpha) \in \Re^n \times \Re : \alpha \geq f(x)\}.$$

The importance of this set is made clear by the fact that f is lower semi-continuous if and only if the epigraph of f is a closed subset of $\Re^n \times \Re$ ([8], Theorem 1.6). Thus given any proper function one can create it's lower semi-continuous regularization (or closure) by defining $\bar{f}$ to be the function associated with the closure of the epigraph of f. This new function is not only lower semi-continuous, it is maximal in the set of lower semi-continuous functions, g, such that $g \leq f$ ([8], p. 14).

1.2.2 The Geometry of Functions

In the previous section we showed that by viewing function geometrically new properties of the function could be established. This thesis will make strong use of the association between a function, its graph, and its epigraph. To understand these relations we shall need some background in Variational Geometry.

We will assume that the reader is familiar with the concept of an interior, and use that to build the definition of a relative interior. Given C, a nonempty subset of $\Re^n$, there is no guarantee that the interior of C is nonempty; so, we seek a more general concept. To begin we define an affine subset of $\Re^n$ to be any translation of a subspace of $\Re^n$. Thus every affine subset has associated with it a well defined dimension. (For example any line in $\Re^n$ is an affine set with dimension equal to 1, while $\Re^n$ is an affine set with dimension equal to n). Returning to C we note that every subset of $\Re^n$ is contained at least one affine set (namely $\Re^n$ itself). We call the smallest affine set containing C the affine hull of C. Lastly we define the relative interior of C to be the interior of C relative to it's affine hull, and denote it rint$C$. To clarify the definition let us consider the case $C = \{56\} \times [714, 755] \subseteq \Re^2$. The affine hull of C is the line $\{(x, y) : x = 56, y \in \Re\}$, which has dimension 1. Thus the relative interior of C is the set $\{56\} \times (714, 755) \subseteq \Re^2$.

Before establishing some basic geometry we must briefly return our attention to functions. We call a map, $S : \Re^n \rightarrow \Re^m$, set valued if $S(x)$ is a set in $\Re^m$ for each $x \in \Re^n$. By relating points to the singleton sets and $\pm\infty$ to the empty set we note that all functions are set valued. Thus in seeking to create the idea of continuity for a set valued map we should extend the logic used previously. Therefore, once again we begin by creating a limsup and liminf functions, this time with regards to set valued mappings. Specifically,

$$\limsup_{x \to \bar{x}} S(x) := \{u : \exists x_k \to \bar{x}, u_k \to u \; with \; u_k \in S(x_k)\},$$

$$\liminf_{x \to \bar{x}} S(x) := \{u : \forall x_k \to \bar{x}, \exists u_k \to u \; with \; u_k \in S(x_k)\}.$$

Similar to the case of functions we find $\limsup S \supseteq S$, and $\liminf S \subseteq S$ always hold true, thus prompting definitions for outer semi-continuous (osc) and inner semi-continuous (isc). Specifically S is outer semi-continuous if $\limsup S \subseteq S$, and inner semi-continuous if $\liminf S \supseteq S$. Inspiration from the single valued case leads us to define a mapping as continuous if it is both outer semi-continuous and inner semi-continuous.

With this background we can now define the remainder of our geometrical definitions. The first of these will be the concept of Normal and Tangent cones to a set. Intuitively to say w is a tangent vector to C at $\bar{x}$ means that some sequence in C converges to $\bar{x}$ from the direction of w. It is from this intuition we provide the definition of a tangent cone; if C is a subset of $\Re^n$ and $\bar{x} \in C$ then we define the tangent cone of C at $\bar{x}$ as

$$T_C(\bar{x}) := \limsup_{\tau \searrow 0} \frac{C - \bar{x}}{\tau}.$$

Continuing with this intuition we would like to provide a concept of a vector normal to the set by taking the vectors normal to the tangent cone. This leads to the following two definitions. If C is a subset of $\Re^n$, and $\bar{x} \in C$, then we say $w \in \Re^n$ is normal to C at $\bar{x}$ in the regular sense if for any $x \to \bar{x}$ with $x \in C$ one has

$$\langle w, x - \bar{x} \rangle \leq o(|x - \bar{x}|),$$

where $o(|x - \bar{x}|)$ refers to a term such that $\frac{o(|x-\bar{x}|)}{|x-\bar{x}|} \to 0$ as $x \to \bar{x}$. We call the set of all such vectors the regular normal cone to C at $\bar{x}$, and denote it $\hat{N}_C(\bar{x})$. We define w as normal to C at $\bar{x}$ in the general sense if there exists sequences $x_k \to \bar{x}$, and $w_k \to w$ with $x_k \in C$, and $w_k \in \hat{N}_C(x_k)$. We denote the set of all such vectors $N_C(\bar{x})$, and call $N_C(\bar{x})$ the normal cone of C at $\bar{x}$. In all three of these definitions we note the term cone is not misused, as multiplying a given vector by a positive constant will not affect any of these properties.

The last geometrical notion we shall require for this thesis is that of epi-convergence. Suppose f_k is a sequence of functions on $\Re^n$, then we define the lower epi-limit ($e - \liminf$) as the function whose epigraph corresponds to $\limsup_{k\to\infty} epi(f_k)$. Similarly the upper epi-limit (e-limsup) is defined such that

$$epi(e - \limsup_{k\to\infty} f_k) = \liminf_{k\to\infty} epi(f_k).$$

If these two functions agree and equal f then we say f_k epi-converges to f, and write $e - \lim_{k\to\infty} f_k = f$.

The main use of epi-convergence will be in the examination of epi-derivatives. We will call a function epi-differentiable at $\bar{x}$ for $\bar{w}$ if the epi-limit of $\frac{f(\bar{x}+\tau\bar{w})-f(\bar{x})}{\tau}$ exists as τ decreases to 0. We call a function twice epi-differentiable at $\bar{x}$ for $\bar{w}$ if $\frac{f(\bar{x}+\tau z)-f(\bar{x})-\tau\langle\bar{w},z\rangle}{\frac{1}{2}\tau^2}$ epi-converges to

$$f''_{\bar{x},\bar{w}}(z) := \liminf_{\substack{\delta \searrow 0 \\ z' \to z}} \frac{f(\bar{x} + \delta z') - f(\bar{x}) - \delta\langle\bar{w}, z'\rangle}{\frac{1}{2}\delta^2}$$

as τ converges to 0 from above, in which case we call $f''_{\bar{x},\bar{w}}$ the second epi-derivative of f. The uses of second epi-derivatives will be explored further in Chapter 4.

1.2.3 Subgradients, Subderivatives, and Strict Differentiability

We have seen how it is reasonable to assume a given function is proper and lower semi-continuous; however, it would be too large of an assumption to presume that a given function was differentiable. To provide generalizations of differentiability we introduce the notions of subgradients, and subdifferentiability.

Consider a function, $f : \Re^n \to \bar{\Re}$, and a point $\bar{x}$ at which $f(\bar{x})$ is finite. Then we say w is a regular subgradient of f at $\bar{x}$ if

$$f(x) \geq f(\bar{x}) + \langle w, x - \bar{x} \rangle + o(|x - \bar{x}|),$$

and denote the set of all regular subgradients of f at $\bar{x}$ by $\hat{\partial} f(\bar{x})$. Alternately if one wanted to avoid the 'little o' notation, one could write $w \in \hat{\partial} f(\bar{x})$ if and only if

$$\liminf_{\substack{x \to \bar{x} \\ x \neq \bar{x}}} \frac{f(x) - f(\bar{x}) - \langle w, x - \bar{x} \rangle}{|x - \bar{x}|} \geq 0.$$

Having established this we denote the set of all subgradients of f at $\bar{x}$ as $\partial f(\bar{x})$, and define them by $w \in \partial f(\bar{x})$ if and only if there exists $x_k \to \bar{x}$, and $w_k \in \hat{\partial} f(x_k)$ with $w_k \to w$, and $f(x_k) \to f(\bar{x})$. The final condition here ($f(x_k) \to f(\bar{x})$) is usually referred to as f-attentive convergence. To simplify notion we shall henceforth write $x \to_f \bar{x}$ to mean $x \to \bar{x}$ and $f(x) \to f(\bar{x})$. This concept will play a role in the next definition.

A related, but different notion of subdifferentiable is that of a subderivative. Once again we begin with a function f, and a point $\bar{x}$ at which $f(\bar{x})$ is finite. Given these we define regular subderivatives,

$$\hat{d} f(\bar{x}, \cdot) := e - \limsup_{\substack{\tau \searrow 0 \\ x \to_f \bar{x}}} \frac{f(\bar{x} + \tau \cdot) - f(\bar{x})}{\tau},$$

where $e - \limsup$ refers to the epi-graphical convergence described in subsection 1.2.2. Fortunately this definition will be of much less use than that of simpler formula for general subderivatives (henceforth called subderivatives):

$$df(\bar{x}, \bar{w}) := \liminf_{\substack{\tau \searrow 0 \\ w \to \bar{w}}} \frac{f(\bar{x} + \tau w) - f(\bar{x})}{\tau}.$$

By the definition it is clear that subderivatives can be viewed as directional derivatives for a function at a point. Indeed it is true that a function is differentiable at $\bar{x}$ if and only if the subderivative function at $\bar{x}$ is linear. That is $df(\bar{x},\cdot)$ is of the form $df(\bar{x},\cdot) = \langle A, \cdot \rangle$, in which case $A = \nabla f(\bar{x})$ ([8], Exercise 8.20).

Much of this thesis relies on the interaction between subgradients, subderivatives and their regular counterparts. To facilitate this interaction we shall define a function to be subdifferentially regular at $\bar{x}$ if it is locally lower semi-continuous and the normal cone of epi$f$ at $(\bar{x}, f(\bar{x}))$ is equal to the regular normal cone there. Two important results on subdifferentially regular functions are that a lower semi–continuous function is subdifferentially regular at $\bar{x}$ if and only if $df(\bar{x},\cdot) = \hat{d}f(\bar{x},\cdot)$ which is true if and only if $\hat{\partial} f(\bar{x}) = \partial f(\bar{x})$ ([8], Theorems 8.9 and 8.19). As a result if f is subdifferentially regular at $\bar{x}$ then $\partial f(\bar{x})$ is a convex subset of $\Re^n$ ([8], Theorem 8.6). Thus requiring a function to be subdifferentially regular is equivalent to asking that the subderivatives and subgradients are well behaved.

In one final notion regarding the differentiability of a function we examine strict differentiability. Recall a function is differentiable at $\bar{x}$ with gradient $\nabla f(\bar{x}) = \bar{w}$ if

$$\lim_{x \to \bar{x}} \frac{f(x) - f(\bar{x}) - \langle \bar{w}, x - \bar{x} \rangle}{|x - \bar{x}|}$$

exists and is equal to 0. We call the function strictly differentiable at $\bar{x}$ with gradient $\nabla f(\bar{x}) = \bar{w}$ if this limit can be loosened to

$$\lim_{\substack{x, x' \to \bar{x} \\ x \neq x'}} \frac{f(x) - f(x') - \langle \bar{w}, x - x' \rangle}{|x - x'|}.$$

One can immediately see how this is a much stronger notion than ordinary differentiability. This will be examined further in Fact 3.6 where a list of equivalent conditions is given.

1.3 Moreau Envelopes and U-Lagrangians

The major focus of this thesis will be the construction and properties of a new envelope called the Quadratic Sub-Lagrangian. The inspiration of this envelope came from a combination of Moreau envelopes and U-Lagrangians. Therefore it will be useful to discuss these two envelopes and some of their properties before we continue.

1.3.1 Moreau Envelopes

We begin by examining the first of the two concepts, Moreau envelopes. Moreau first developed the idea of this envelope in 1963 in the paper entitled "Propriétés des applications "prox"" [3]. However, most of the results attributed to Moreau come from his paper "Proximité et dualité dans un espace hilbertien" [4] where he generalized most of his results to Hilbert spaces, and developed many new ones. Before we discuss these results, we need the definition of a Moreau envelope; if $f : \Re^n \to \bar{\Re}$ is a proper lower semi-continuous function, and $\lambda > 0$, then we define the Moreau envelop $e_\lambda f$, and it's related Proximal mapping $P_\lambda f$ by,

$$e_\lambda f(\bar{x}) := \inf_{x \epsilon \Re^n} \{f(x) + \frac{1}{2\lambda}|x - \bar{x}|^2\}$$

$$P_\lambda f(x) := \arg\min_{x \epsilon \Re^n} \{f(x) + \frac{1}{2\lambda}|x - \bar{x}|^2\}.$$

Where the argmin function refers to the argument of the minimum, and is defined:

$$\arg\min_{x \epsilon C} f(x) := \begin{cases} \{\bar{x} \epsilon C : f(\bar{x}) = \inf_{x \epsilon C} f(x)\} & \inf_{x \epsilon C} f(x) \neq \emptyset \\ \emptyset & otherwise \end{cases}$$

The concept of a Moreau envelope is to use minimization as a means of defining one function in terms of another. Of course the primary desire of such an envelope function is that it is proper. Since our original function, f, was proper and $e_\lambda f(\bar{x}) \leq f(\bar{x})$ everywhere, we know that $e_\lambda f$ is not constantly infinity. It can be shown that if there exists some $\lambda_0 > 0$ such that $e_{\lambda_0} f(\bar{x}) > -\infty$ for some $\bar{x} \epsilon \Re^n$ then $e_\lambda f(x) \nearrow f(x)$ for all x as $\lambda \searrow 0$ ([8], Theorem 1.25). If f has such a λ_0 then we say f is prox-bounded, and call the supremum of the set of all such λ_0 the threshold of prox-boundedness for f. It follows that if f is prox-bounded with threshold λ_f then for any $\lambda \epsilon (0, \lambda_f)$, one has $e_\lambda f$ is not only proper, but finite valued ([8], Theorem 1.25).

One of the more interesting results in the theory of Moreau envelopes is that the Moreau envelope of a proper lower semi-continuous prox-bounded function is actually continuous for sufficiently small λ. In fact if λ_f is the threshold of prox-boundedness then $e_\lambda f$ is continuous for any $\lambda \epsilon (0, \lambda_f)$ ([8], Theorem 1.25). This result sets one of the major themes in the development of new envelope functions, that the envelope should have properties above and beyond that of the original function. In Chapter 4 we shall see that the Quadratic Sub-Lagrangian can satisfies this requirement, in that the envelope can be twice epi-differentiable at a point where the original function was not.

1.3.2 Lagrangians, and the U-Lagrangian

Before tackling the concept of a U-Lagrangian we should examine the construction of a Lagrangian.

In 1788, Joseph Louis Lagrange published his most famous result in "Analytical Mechanics", stating that problems in mechanics can generally be solved by reducing them to the theory of ordinary and partial differential equations. Most of his work centered around this idea, and therefore the study of partial differential equations, however he is also credited with the development of the Lagrangian function ([9], p. 531). In 1973 Rockafellar published "The Multiplier Method of Hestenes and Powell Applied to Convex Programming" in which he combined recent results of Hestenes and Powell to show how the Lagrangian function could be used in the problem of optimizing a convex function [7]. Although many variations of the Lagrangian exist generally the Lagrangian of a function $f : \Re^n \times \Re^m \to \bar{\Re}$ is accepted to be defined by

$$\begin{array}{rcl} l. \ \Re^n \times \Re^m & \to & \bar{\Re} \\ (x, y) & \mapsto & \inf_{u \in \Re^m} \{f(x, u) - \langle u, y \rangle\}. \end{array}$$

In a recently published paper by Lemaréchal, Oustry, and Sagastizábal, entitled "The U-Lagrangian of a Convex Function", the concept of a Lagrangian was modified to create the definition of a U-Lagrangian [2]. There are three major difference in the creation of U-Lagrangians from that of the Lagrangian. The first and greatest difference is instead of being given a function on $\Re^n \times \Re^m$ Lemaréchal, Oustry, and Sagastizábal began with a function $f : \Re^n \to \Re$, and from it created a decomposition on $\Re^n$ into two perpendicular subspaces U, and V ([2], Definition 3.1). The second is that instead of centering the Lagrangian at 0, Lemaréchal, Oustry, and Sagastizábal selected an arbitrary $x \in dom f$. Lastly, in the case of a U-Lagrangian the vector corresponding to the y is the usual Lagrangian is fixed. Thus the U-Lagrangian of $f : \Re^n \to \Re$, at $\bar{x}$ for $\bar{w}$ is defined:

$$L_U(u) := \inf_{v \in V} \{f(\bar{x} + (u + v)) - \langle \bar{w}, v \rangle\}.$$

To explain further let us briefly discuss projections and perpendicular subspaces of $\Re^n$. To begin with we should note that if U is a subspace of $\Re^n$

then there exists an unique subspace $V \subseteq \Re^n$ such that U is perpendicular to V and $U \oplus V = \Re^n$. That is for any $u \in U$ and $v \in V$ one finds $\langle u, v\rangle = 0$, and for each $x \in \Re^n$ there are elements x_1 and x_2 in $\Re^n$ such that $x_1 \in U$, $x_2 \in V$ and $x = x_1 + x_2$. Furthermore for any fixed $x \in \Re^n$, x_1 and x_2 will be unique with $|x|^2 = |x_1|^2 + |x_2|^2$. In this thesis we shall denote this decomposition by $x = x_u + x_v$. Lastly note that since $\langle \bar{w}_v, v\rangle = \langle \bar{w}, v\rangle$, we can (and will) write the former as a reminder that only the V subspace of $\Re^n$ is being considered.

To relate the U-Lagrangian to the construction of an ordinary Lagrangian consider the function f defined on $\Re^n \times \Re^m$. One can easily consider the function to be instead defined on $\Re^{n+m}$ by the of the projection mappings $P_U(x) = x_u$ and $P_V(x) = x_v$. Using these maps we can create $\hat{f}(x) := f(P_U(x), P_V(x))$, where $x \in \Re^{m+n}$, $U = \Re^n \times \{0\}^m$, $V = \{0\}^n \times \Re^m$, and the zero elements of $P_U(x)$ and $P_V(x)$ are ignored. In doing this we immediately note that the opposite is also true. That is if $f : \Re^n \to \bar{\Re}$, where $n = n_1 + n_2$, and $n_1, n_2 \geq 1$, then we can consider f to be a function on $\Re^{n_1} \times \Re^{n_2}$. To show this we consider the function defined by, $\hat{f}(x_1, x_2) = f((x_1, 0, 0, \ldots 0) + (0, 0, \ldots 0, x_2))$, where there are n_1 0's inserted at the end of the x_1 term, and n_2 0's inserted at the beginning of the x_2 term. In fact this technique can be used equally well to consider $f : \Re^n \to \bar{\Re}$ to be a function of $f : U \times V \to \bar{\Re}$ where U and V are any perpendicular decomposition of $\Re^n$.

Taking this into consideration we examine the usual Lagrangian on a function $f : \Re^n \times \Re^m \to \bar{\Re}$,

$$l(x, y) := \inf_{u \in \Re^m} \{f(x, u) - \langle u, y\rangle\}.$$

If we consider the related $\hat{f} : \Re^{n+m} \to \bar{\Re}$, with U and V to be the subspaces corresponding to $\Re^n$ and $\Re^m$ respectively then we can rewrite this as,

$$l(x, y) := \inf_{v \in V} \{\hat{f}(x + v) - \langle v, y\rangle\}.$$

Thus by fixing $y = \bar{w} \in V$ we have the U-Lagrangian of $\hat{f}$ at $\bar{x} = 0$ for $\bar{w}$, and see the direct relation between the two concepts.

In "The U-Lagrangian of a Convex Function", Lemaréchal, Oustry, and Sagastizábal showed that given a proper convex function and selecting a proper choice in the UV decomposition one can guarantee the U-Lagrangian will be well behaved. Specifically they demonstrated that if $\bar{w} \in \operatorname{rint} \partial f(\bar{x})$, and $U = N_{\partial f(\bar{x})}(\bar{w})$ then the U-Lagrangian will also be a proper convex function. Moreover the U-Lagrangian will be differentiable at 0 with $\nabla L_U(0) = \bar{w}_u$. In examining the higher order behaviour of the U-Lagrangian, Lemaréchal, Oustry, and Sagastizábal showed a relationship between second order expansions of the U-Lagrangian and those of the Moreau envelope of f centered at $\bar{x}+\bar{w}$ when examined in the U direction ([2], Proposition 5.3). Lemaréchal, Oustry, and Sagastizábal justified the appeal of this result further by showing a second order expansion of the U-Lagrangian could be used (to a degree) as a second order expansion of the original function ([2], Theorem 3.9). In this thesis we will extend these results to the broader class of functions known as prox-regular (described in the next section), while simultaneously exploring why the subspace $N_{\partial f(\bar{x})}(\bar{w})$was chosen for the UV decomposition.

1.4 Prox-Regularity

The concept of prox-regularity was first introduced by Poliquin and Rockafellar in 1996, by the paper entitled "Prox-Regular Functions in Variational Analysis" [6]. Poliquin and Rockafellar have showed that prox-regularity is in many ways the natural extension of the limited class of functions called convex. To begin with, the very definition is focused around the ability to bound functions from below by quadratics. Specifically they defined a function, $f : \Re^n \to \bar{\Re}$, to be prox-regular at $\bar{x}$ for $\bar{w} \in \partial f(\bar{x})$ if f is a locally lower semi-continuous there, and there exists $\varepsilon > 0$ and $\rho > 0$ such that

$$f(\bar{x}) > f(x) + \langle w, \bar{x} - x \rangle - \rho |\bar{x} - x|^2/2$$

whenever $|\bar{x} - x| < \varepsilon$, $|x - \bar{x}| < \varepsilon$, $x \neq \bar{x}$, $|f(x) - f(\bar{x})| < \varepsilon$, and $|w - \bar{w}| < \varepsilon$ with $w \in \partial f(x)$. We call a function prox-regular at $\bar{x}$ if it is prox-regular at $\bar{x}$ for all $\bar{w} \in \partial f(\bar{x})$, and call a function prox-regular if this holds for all $\bar{x} \in \Re^n$ [6].

Since convex functions can be thought of as those functions which can be bounded below by affine functions this appears to be the natural evolution. In the paper, "Generalized Hessian Properties of Regularized Nonsmooth Functions", Poliquin and Rockafellar showed that a proper lower semi-continuous function, f, is prox-regular if and only if the f-attentive ε-localization of subgradient map is pre-monotone ([5], Theorem 2.2). By this second condition we mean that if I is the identity mapping then

$$S_\varepsilon := \begin{cases} \{w \in \partial f(x) : |w| < \varepsilon\} & if\ |x| < \varepsilon, |f(x) - f(0)| < \varepsilon \\ \emptyset & otherwise \end{cases}$$

has the property that $S_\epsilon + rI$ is monotone for sufficently large r. This is the natural extension of convex functions, which can be shown to be the proper lower semi-continuous functions whose subgradient map is monotone ([8], Theorem 12.17). Having noted this relationship with convex functions, it is not surprising that the set of all prox-regular functions include all convex functions. Moreover, prox-regularity covers two other broad classes of functions, strongly amenable functions and lower-$\mathcal{C}^2$ functions ([6], Proposition 2.5 and Example 2.7).

A function, $f : \Re^n \to \Re$, is strongly amenable at $\bar{x}$ if there exists a neighbourhood of $\bar{x}$ on which f can be written as $f(x) = g(F(x))$, where g is a proper lower semi-continuous function of $\Re^m \to \bar{\Re}$, and $F : \Re^n \to \Re^m$ is $\mathcal{C}^2$ such that if D is the closure of the domain of g then the only vector $y \in N_D(F(\bar{x}))$ with $\nabla F(\bar{x})^* y = 0$ is $y = 0$. We call the function strongly amenable if this property holds for all $\bar{x} \in \Re^n$. Lower-$\mathcal{C}^k$ functions are functions, f, that can be written in the form

$$f(x) := \max_{t \in T} f_t(x)$$

where f_t are $\mathcal{C}^k$ functions, and T is a compact index set. It has been shown that for $k \geq 2$ the class of lower-$\mathcal{C}^k$ functions is indistinguishable from the class of lower-$\mathcal{C}^2$ functions ([8], Corollary 10.34). Since the definition of prox-regularity is based on local properties one also finds that if either property occurs locally then the function is locally prox-regular as well.

Of course since all $\mathcal{C}^2$ functions are lower-$\mathcal{C}^2$ it follows that any smooth function is prox-regular. This demonstrates the powerful fact that any measurable function (i.e. $\mathcal{L}^1$ functions) can be approximated by prox-regular functions, a property that convex functions do not share. This property makes it desirable to generalize the results of Lemaréchal, Oustry, and Sagastizábal to prox-regular functions. In this thesis we shall see how by altering the U-Lagrangian slightly we can create a new envelope, deemed the Quadratic Sub-Lagrangian. We shall begin by showing this envelope is generally well behaved, and a better estimate of the original function than the Moreau envelope. Next we shall show that there is an optimal UV decomposition, and in applying it the Quadratic Sub-Lagrangian will have many of the properties of the U-Lagrangian. Lastly we shall examine the properties of a second order expansion of the Quadratic Sub-Lagrangian; showing that it can act as a second order expansion of the original function, and demonstrate several properties that will lead to its existence.

Chapter 2

The Quadratic Sub-Lagrangian and its Basic Properties

Non enim excursus hic ejus, sed opus ipsum ets. [1]

Pliny the Younger

2.1 The Quadratic Sub-Lagrangian

In chapter 1 we examined the U-Lagrangian of a convex function, as developed by Lemaréchal, C., Oustry, F. and, Sagastizábal, C., in "The U-Lagrangian of a Convex Function" [2]. In this thesis we shall see how many of their results can be applied to the more general class of functions known as prox-regular. To do this we shall create a new envelope function, called the Quadratic Sub-Lagrangian, as follows.

Let $f : \Re^n \to \bar{\Re}$ be a proper function with $\bar{x} \in dom(f)$, and $\bar{w} \in \partial f(\bar{x})$. Let U be a subspaces of $\Re^n$, and $V = U^{\perp}$. We shall denote by Φ_R the Quadratic Sub-Lagrangian of f with respect to $\bar{x}$, $\bar{w}$, R and U; and by W_R it's related proximal map. More specifically:

$$\Phi_R(u) := \inf_{v \in V}\{f(\bar{x} + (u + v)) - \langle \bar{w}_v, v\rangle + \frac{R}{2}|v|^2\} \tag{2.1}$$

$$W_R(u) := \arg\min_{v \in V}\{f(\bar{x} + (u + v)) - \langle \bar{w}_v, v\rangle + \frac{R}{2}|v|^2\} \tag{2.2}$$

[1]For this is not a digression from it, but the work itself.

where x_u and x_v are the projections of x onto the U and V subspaces respectively, for example $\bar{w} = (\bar{w}_u + \bar{w}_v)$. The inspiration for this definition comes from "The U-Lagrangian of a Convex Function", where Lemaréchal, Oustry, and Sagastizábal show that for convex functions, this envelope agreeds with the U-Lagrangian at 0 up to second order ([2], Lemma 5.1). Before we continue we simplify notation by defining,

$$h_R(u,v) := f(\bar{x} + (u+v)) - \langle \bar{w}_v, v\rangle + \frac{R}{2}|v|^2. \tag{2.3}$$

In examining the formation of the Quadratic Sub-Lagrangian one can see aspects of both the U-Lagrangian function and the Moreau envelope. Like the U-Lagrangian the Quadratic Sub-Lagrangian has a linear term $(\langle \bar{w}_v, v\rangle)$ inserted to shift the interior function (h_R) so that $0 \epsilon \partial h_R(0)$. It will be shown later (Corollary 2.13) that this shift will force $\bar{w}_u \epsilon \partial \Phi_R(0)$. The Quadratic term $(\frac{R}{2}|v|^2)$ follows the same purpose as that of the Quadratic term in a Moreau envelope, it helps the envelope to be a proper function. When we apply this term to prox-regular functions that are bounded below by a Quadratic we shall see that for sufficiently large R the Quadratic Sub-Lagrangian is not only a proper function (Theorem 2.6), but whenever $\Phi_R(u) < \infty$ the infimum is actually obtained (Proposition 2.7). When applying the Quadratic Sub-Lagrangian to convex functions this term can be taken to be arbitrarily small, and if f is strictly convex, R can actually be taken to be 0. This will follow from thinking of convex functions as functions that can be bounded below by linear functions instead of Quadratics. Thus, in the case of strictly convex functions, this thesis will provide alternate proofs to many of Lemaréchal, Oustry, and Sagastizábal's theorems.

2.2 Basic Properties of the Quadratic Sub-Lagrangian

The first question one should ask of any envelope function is does the envelope form a good estimate of the original function? In the case of the Quadratic Sub-Lagrangian Φ_R is defined only on the U subspace of $\Re^n$, so the question can be phrased does $\lim_{R\to\infty} \Phi_R(u) = f(\bar{x} + (u+0))$? In showing this is true we shall also see that the Quadratic Sub-Lagrangian forms a better estimate of $f(\bar{x} + (u+0))$ than the Moreau envelope, further justifying its examination.

Proposition 2.1 *Let Φ_R be the Quadratic sub-Lagrangian of f with respect to $\bar{x}$, $\bar{w} \in \partial f(\bar{x})$, R, and a subspace U. Let $e(u)$ be the Moreau envelope of $\tilde{f} := f - \langle \bar{w}, \cdot \rangle$ centered at $\bar{x}$ and restricted to the subspace U with parameter $\frac{1}{R}$. That is,*

$$e(u) := e_{\frac{1}{R}} \tilde{f}(\bar{x} + (u + 0)).$$

Then $\Phi_R(u) \geq e(u) + \langle \bar{w}, \bar{x} + (u + 0) \rangle$ for all $u \in U$.

Proof:

Let $\tilde{f}(x) = f(x) - \langle \bar{w}, x \rangle$, therefore,

$$\begin{aligned}
e(u) &= e_{\frac{1}{R}} \tilde{f}(\bar{x} + (u + 0)) \\
&= \inf_{x \in \Re^n} \{\tilde{f}(x) + \tfrac{R}{2} |\bar{x} + (u + 0) - x|^2\} \\
&= \inf_{\substack{\tilde{u} \in U \\ \tilde{v} \in V}} \{\tilde{f}(\bar{x} + (\tilde{u} + \tilde{v})) + \tfrac{R}{2} |(\tilde{u} + \tilde{v}) - (u + 0)|^2\} \\
&\leq \inf_{\tilde{v} \in V} \{\tilde{f}(\bar{x} + (u + \tilde{v})) + \tfrac{R}{2} |\tilde{v}|^2\} \\
&= \inf_{\tilde{v} \in V} \{f(\bar{x} + (u + \tilde{v})) - \langle \bar{w}, (u + \tilde{v}) \rangle + \tfrac{R}{2} |\tilde{v}|^2\} - \langle \bar{w}, \bar{x} \rangle \\
&= \Phi_R(u) - \langle \bar{w}_u, u \rangle - \langle \bar{w}, \bar{x} \rangle
\end{aligned}$$

♣

Corollary 2.2 *Let $f : \Re^n \to \bar{\Re}$ be a proper lsc function that is prox-regular at $\bar{x}$ for $\bar{w} \in \partial f(\bar{x})$ with respect to ρ and ε, and bounded below by a Quadratic there also with respect to ρ. Let $f_U(u) = f(\bar{x} + (u + 0))$, and Φ_R be the Quadratic Sub-Lagrangian of f with respect to $\bar{x}$, $\bar{w}$, R, and a subspace U. Then as $R \to \infty$ one finds Φ_R converges to f_U pointwise.*

Proof:

Again, let $\tilde{f}(x) = f(x) - \langle \bar{w}, x \rangle$. By the prox-regularity of f at $\bar{x}$ we know that locally

$$f(x) \geq f(\bar{x}) + \langle \bar{w}, x - \bar{x} \rangle - \frac{\rho}{2} |x - \bar{x}|^2.$$

Since we have f bounded below by a Quadratic at $\bar{x}$ we may assume that this inequality holds for all $x \in \Re^n$. Rewriting at $x = \bar{x} + (u + v)$ we note for all $(u + v) \in U \oplus V = \Re^n$,

$$\begin{aligned}
f(\bar{x} + (u + v)) - \langle \bar{w}, \bar{x} + (u + v) \rangle &\geq f(\bar{x}) - \langle \bar{w}, \bar{x} \rangle - \tfrac{\rho}{2} |(u + v)|^2 \\
\tilde{f}(\bar{x} + (u + v)) &\geq f(\bar{x}) - \langle \bar{w}, \bar{x} \rangle - \tfrac{\rho}{2} |(u + v)|^2.
\end{aligned}$$

Therefore,

$$\begin{aligned} e_{\frac{1}{R}}\tilde{f}(\bar{x}+u+0) &= \inf_{x \in \Re^n}\{\tilde{f}(x) + \tfrac{R}{2}|\bar{x}+(u+0)-x|^2\} \\ &= \inf_{\substack{\tilde{u}\in U\\ \tilde{v}\in V}}\{\tilde{f}(\bar{x}+(\tilde{u}+\tilde{v})) + \tfrac{R}{2}|(u-\tilde{u}-\tilde{v})|^2\} \\ &\geq \inf_{\substack{\tilde{u}\in U\\ \tilde{v}\in V}}\{f(\bar{x}) - \langle\bar{w},\bar{x}\rangle - \tfrac{\rho}{2}|(\tilde{u}+\tilde{v})|^2 + \tfrac{R}{2}|(u-\tilde{u}-\tilde{v})|^2\} \\ &\geq \inf_{\substack{\tilde{u}\in U\\ \tilde{v}\in V}}\{\tfrac{R}{2}|u-\tilde{u}|^2 - \tfrac{\rho}{2}|\tilde{u}|^2 + \tfrac{(R-\rho)}{2}|\tilde{v}|^2\} + f(\bar{x}) - \langle\bar{w},\bar{x}\rangle \end{aligned}$$

Provided $R > \rho$, then $(\frac{R}{2}|u-\tilde{u}|^2 - \frac{\rho}{2}|\tilde{u}|^2) + \frac{(R-\rho)}{2}|\tilde{v}|^2$ is strictly convex in $\tilde{u}$, and $\tilde{v}$. Therefore the infimum is obtained, and we have $e_{\frac{1}{R}}\tilde{f}(\bar{x}+(u+0)) > -\infty$. So we find that $\tilde{f}(x) = f(x) - \langle\bar{w},x\rangle$ is prox-bounded at $\bar{x}+(u+0)$ for any $u \in U$ with threshold $\lambda_{\tilde{f}} = \frac{1}{\rho}$. Applying this fact we note that $e_{\frac{1}{R}}\tilde{f} \nearrow \tilde{f}$ pointwise for $\bar{x}+(u+0), u \in U$ as $R \to \infty$. Thus

$$e_{\frac{1}{R}}\tilde{f}(\bar{x}+u+0)+\langle\bar{w},\bar{x}+(u+0)\rangle \nearrow \tilde{f}(\bar{x}+u+0)+\langle\bar{w},\bar{x}+(u+0)\rangle = f(\bar{x}+u+0)$$

as $R \to \infty$. By Proposition 2.1 and the squeeze theorem we have: $\Phi_R(u)$ converges pointwise to $f(\bar{x}+(u+0))$ as $R \to \infty$.

◇

Next we address the question of whether or not this inequality can in fact be strict (with the exception of at $f(\bar{x})$ of course). The answer turns out to be yes, as the next example will demonstrate.

Example 2.3

Consider the function $f(x,y) = -x^2 + |y|$. Then

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

so $\bar{w} = (0,0) \in \partial f(0,0)$. Lastly we select

$$U = \Re \times \{0\}$$

$$V = \{0\} \times \Re.$$

Then,

$$\begin{aligned} e_{\frac{1}{R}}f(u+0) &= \inf_{\substack{\tilde{u}\in U\\ \tilde{v}\in V}}\{f(\tilde{u}+\tilde{v}) + \tfrac{R}{2}|u-\tilde{u}|^2 + \tfrac{R}{2}|\tilde{v}|^2\} \\ &= \inf_{\substack{\tilde{u}\in U\\ \tilde{v}\in V}}\{-\tilde{u}^2 + \tfrac{R}{2}(u-\tilde{u})^2 + |\tilde{v}| + \tfrac{R}{2}\tilde{v}^2\} \\ &= \inf_{\tilde{u}\in U}\{-\tilde{u}^2 + \tfrac{R}{2}(u-\tilde{u})^2\} \end{aligned}$$

Assuming $R > 2$ we find that infimum is unique and obtained at $\tilde{u} = \frac{R}{R-2}u$. So

$$e_{\frac{1}{R}}f(u+0) = -R\left[\frac{u}{(R-2)} + \frac{Ru}{2} - \frac{Ru}{(R-2)}\right]^2 = -\frac{Ru^2}{(R-2)}.$$

Comparing this to Φ_R we note,

$$\begin{aligned}\Phi_R(u) &= \inf_{v \epsilon V}\{f(u+v) + \frac{R}{2}|v|^2\} \\ &= \inf_{v \epsilon \Re}\{-u^2 + |v| + \frac{R}{2}|v|^2\} \\ &= -u^2 \\ &= f(u+0) =: f_U(u)\end{aligned}$$

Thus Φ_R is a strictly better estimate of f_U than $e_{\frac{1}{R}}f(\cdot_u + 0)$.

♡

In the above example we found $\Phi_R(u) = f(\bar{x} + (u+0))$ for any $R \geq 2$ and $u \,\epsilon\, U$. We now see that this is not a coincidence, but will always occur when the equation is variable separable along the UV decomposition.

Proposition 2.4 *Suppose f: $\Re^n \to \Re$ is a proper lsc function that is prox-regular at $\bar{x}$ for $\bar{w}\epsilon\partial f(\bar{x})$ with respect to ρ and ε, and bounded below by a Quadratic also with respect to ρ. Suppose also that f is of the form $f(\bar{x}+(u+v)) = f_U(u) + f_V(v) + f(\bar{x})$, where $f_U : U \to \Re$ and $f_V : V \to \Re$, and U and V are perpendicular subspaces of $\Re^n$. Then for some R sufficiently large, $\Phi_R(u) = f(\bar{x} + (u+0))$.*

Proof:

First note that without loss of generality we may assume $f_U(0) = f_V(0) = 0$. Indeed if this is not true then we may create $\tilde{f}_U(u) = f_U(u) - f_U(0)$, and $\tilde{f}_V(v) = f_V(v) - f_V(0)$. Clearly these satisfy $\tilde{f}_U(0) = \tilde{f}_V(0) = 0$. Moreover by assumption we have

$$f(\bar{x}) = f(\bar{x}) + f_U(0) + f_V(0),$$

so $f_U(0) = -f_V(0)$. Therefore we have

$$\begin{aligned}f(\bar{x}) + \tilde{f}_U(u) + \tilde{f}_V(v) &= f(\bar{x}) + f_U(u) - f_U(0) + f_V(u) - f_V(0) \\ &= f(\bar{x}) + f_U(u) + f_V(v) \\ &= f(\bar{x} + (u+v)),\end{aligned}$$

as desired. Returning to $\Phi_R(u)$ we see:

$$\Phi_R(u) = \inf_{v \epsilon V}\{f(\bar{x} + (u, v)) - \langle \bar{w}_v, v\rangle + \frac{R}{2}|v|^2\}$$

$$\Phi_R(u) = f(\bar{x}) + f_U(u) + \inf_{v \epsilon V}\{f_V(v) - \langle \bar{w}_v, v\rangle + \frac{R}{2}|v|^2\}. \tag{2.4}$$

Since f is bounded below by a Quadratic at $\bar{x}$, and prox-regular there for $\bar{w}$ we may assume,

$$\begin{aligned} f(x) &\geq f(\bar{x}) + \langle \bar{w}, x - \bar{x}\rangle - \tfrac{\rho}{2}|x - \bar{x}|^2 \\ f(\bar{x} + (0 + v)) &\geq f(\bar{x}) + \langle \bar{w}_v, v\rangle - \tfrac{\rho}{2}|v|^2 \\ f_V(v) - \langle \bar{w}_v, v\rangle + \tfrac{\rho}{2}|v|^2 &\geq 0. \end{aligned}$$

So for any $R > \rho$ we have $f_V(v) - \langle \bar{w}_v, v\rangle + \frac{R}{2}|v|^2 \geq 0$, and since $f_V(0) = 0$ we have,

$$\inf_{v \epsilon V}\{f_V(v) - \langle \bar{w}_v, v\rangle + \frac{R}{2}|v|^2\} = 0. \tag{2.5}$$

Combining equation (2.4) and (2.5) we see

$$\Phi_R(u) = f_U(u) + f(\bar{x}) = f(\bar{x} + (u + 0)).$$

♠

Immediately, this raises the question: does the Quadratic Sub-Lagrangian differ from the function $f_U = f(\bar{x} + (\cdot_u + 0))$. Although it would be difficult to believe the two functions to always be the same, Proposition 2.4 showed that for many equations the two are indeed the same. The answer, (fortunately), is that Φ_R and f_U can be almost entirely different as the next example will show.

Example 2.5

Consider the function $f : \Re^2 \to \Re$ defined by $f(x, y) = |x|^{|y|+k} + \frac{1}{2}y^2$, where k is a fixed constant greater than 1.

First we show $\bar{w} := (0, 0) \,\epsilon\, \partial f(0, 0)$. Indeed

$$\liminf_{\substack{(x,y)\to(0,0)\\(x,y)\neq(0,0)}} \frac{f(x, y) - f(0, 0) - \langle \bar{w}, (x, y)\rangle}{|(x, y)|} = \liminf_{\substack{(x,y)\to(0,0)\\(x,y)\neq(0,0)}} \frac{|x|^{|y|+k} + \frac{1}{2}y^2}{|(x, y)|}.$$

Since both the numerator and denominator in the above limit are positive, it follows that

$$\liminf_{\substack{(x,y)\to(0,0)\\(x,y)\neq(0,0)}} \frac{f(x,y)-f(0,0)-\langle\bar{w},(x,y)\rangle}{|(x,y)|} \geq 0,$$

so $\bar{w} := (0,0) \in \hat{\partial} f(0,0) \subseteq \partial f(0,0)$.

Next we note that f is prox-regular at $\bar{x}$ and bounded below by a Quadratic there, in fact $f(x,y) \geq 0$ for all $(x,y) \in \Re^2$, so the quadratic can be take to be the constant function $g \equiv 0$. Lastly we select $U = \Re \times \{0\}$ and $V = \{0\} \times \Re$. Since this is the decomposition under which the function was defined we shall relax the notation considering U and V to simply be $\Re$.

We begin by showing that $\Phi_R(u) \neq f(u,0)$ for all $u \neq 0$:

$$\begin{aligned}\Phi_R(u) &= \inf_{v\in V}\{f(u,v)+\tfrac{R}{2}|v|^2\}\\ &= \inf_{v\in\Re}\{|u|^{|v|+k}+\tfrac{R+1}{2}|v|^2\}\\ &= \inf_{v\geq 0}\{|u|^{v+k}+\tfrac{R+1}{2}v^2\}\end{aligned}$$

First we note that if $u = 0$ then the infimum is obtained at $v = 0$ as expected. Next notice that for any $u > 0$ the function inside the infimum is differentiable, and strictly convex in v in the critical area ($v \geq 0$). Indeed for $u > 0$,

$$\frac{d}{dv}(|u|^{v+k}+\frac{R+1}{2}v^2) = \ln u(u^{v+k})+(R+1)v,$$

and

$$\frac{d^2}{dv^2}(|u|^{v+k}+\frac{R+1}{2}v^2) = (\ln u)^2(u^{v+k})+R+1.$$

The second derivative test therefore shows that $(|u|^{v+k}+\frac{R+1}{2}v^2)$ is strictly convex for any $u > 0$. Therefore for any $u > 0$ we know the minimum will occur at the unique point for which $\frac{d}{dv}(|u|^{v+k}+\frac{R+1}{2}v^2) = 0$. Solving this equality we find

$$\frac{d}{dv}(|u|^{v+k}+\frac{R+1}{2}v^2) = 0,$$

$$(\ln|u|)(|u|^{v+k})+(R+1)v = 0.$$

Thus v is required to satisfy

$$|u|^v = -\frac{R+1}{\ln|u|(u^k)}v,$$

and this is not satisfied at $v = 0$. Since the infimum is obtained at a unique point, and $v = 0$ is not that point, it becomes clear that $\Phi_R(u)$ cannot be equal to $f(u, 0) = u^k$.

♠

Our next desire for the Quadratic Sub-Lagrangian is to show it maintains some of the basic properties of the original function. First we shall show Φ_R is proper, and from there show $W_R(u)$ is nonempty whenever $\Phi_R(u) < \infty$.

Theorem 2.6 *Let f: $\Re^n \to \bar{\Re}$ be a proper lsc function that is subdifferentially regular at $\bar{x}$, prox-regular there for $\bar{w} \in \partial f(\bar{x})$, with respect to ρ and ε, and bounded below by a Quadratic (also with respect to ρ). Let U and V be perpendicular subspaces of $\Re^n$. Then Φ_R as defined in (2.1) is proper for any $R > \rho$. Moreover $\Phi_R(0) = f(\bar{x})$, and $W_R(0) = \{0\}$.*

Proof:

First we note that since f is proper, $\Phi_R(u) \neq \infty$ for at least one value of u. Thus we need only show a lower bound of $h_R(u, v)$ for any given $u \in U$.

By the prox-regularity of f at $\bar{x}$ we know that locally

$$f(x) \geq f(\bar{x}) + \langle \bar{w}, x - \bar{x} \rangle - \frac{\rho}{2}|x - \bar{x}|^2.$$

Since we have f bounded below by a Quadratic at $\bar{x}$ we may assume that this inequality holds for all $x \in \Re^n$. Rewriting with $x = \bar{x} + (u, v)$ we note for all $(u + v) \in U \oplus V = \Re^n$,

$$\begin{aligned} f(\bar{x} + (u + v)) &\geq f(\bar{x}) + \langle \bar{w}, (u + v) \rangle - \tfrac{\rho}{2}|(u + v)|^2 \\ f(\bar{x} + (u + v)) - \langle \bar{w}_v, v) \rangle + \tfrac{\rho}{2}|v|^2 &\geq f(\bar{x}) + \langle \bar{w}_u, u \rangle - \tfrac{\rho}{2}|u|^2. \end{aligned}$$

Thus for any $R > \rho$ and $(u + v) \in U \oplus V$ with $v \neq 0$ we find,

$$f(\bar{x} + (u + v)) - \langle \bar{w}_v, v \rangle + \frac{R}{2}|v|^2 > f(\bar{x}) + \langle \bar{w}_u, u \rangle - \frac{\rho}{2}|u|^2. \tag{2.6}$$

Therefore for any $R > \rho$ and $u \in U$,

$$\inf_{v \in V}\{f(\bar{x} + (u + v)) - \langle \bar{w}_v, v \rangle + \frac{R}{2}|v|^2\} \geq f(\bar{x}) + \langle \bar{w}_u, u \rangle - \frac{\rho}{2}|u|^2$$

$$\Phi_R(u) \geq f(\bar{x}) + \langle \bar{w}_u, u \rangle - \frac{\rho}{2}|u|^2 > -\infty, \tag{2.7}$$

proving Φ_R is proper.

Equation (2.7) above also shows $\Phi_R(0) \geq f(\bar{x})$. Since the definition of Φ_R clearly shows $\Phi_R(0) \leq f(\bar{x})$ equality is established. Lastly by equation (2.6) applied to $u = 0$, we see $W_R(0) = \{0\}$.

♣

Besides showing that whenever $\Phi_R(u) < \infty$ one can take $\Phi_R(u)$ to be $\min_{v \in V}\{f(\bar{x} + (u + v)) - \langle \bar{w}_v, v\rangle + \frac{R}{2}|v|^2\}$, the next proposition provides two very important equations regarding W_R and Φ_R. The first, equation (2.8), will be often referred to in order to show that $|W_R(u)| \to 0$ as $u \to 0$. The second, equation (2.9), regards a technique of shifting f by a linear function so that $0 \in \partial f(\bar{x})$. This shift will often be used to split proofs into two simpler components. The comparison of the original Quadratic Sub-Lagrangian, to the Quadratic Sub-Lagrangian of the shifted function will therefore become very important in many theorems.

Proposition 2.7 *Let f: $\Re^n \to \bar{\Re}$ be a proper lsc function that is subdifferentially regular at $\bar{x}$, prox-regular there for $\bar{w} \in \partial f(\bar{x})$, and bounded below by a Quadratic. Let U and V be perpendicular subspaces of $\Re^n$. Then whenever R is sufficiently large, W_R is nonempty for any u in the domain of Φ_R.*

Proof: Case I, $\bar{w} = 0$.

Let $\delta > 0$ and $u \in \text{dom}(\Phi_R)$ be given, then there exists $v \in V$ such that.

$$f(\bar{x} + (u + v)) + \frac{R}{2}|v|^2 \leq \Phi_R(u) + \delta$$

Since $f(x)$ is bounded below by a Quadratic we have, for appropriate ρ,

$$\begin{aligned} f(\bar{x}) - \tfrac{\rho}{2}|(u+v)|^2 &\leq f(\bar{x} + (u+v)) \\ f(\bar{x}) - \tfrac{\rho}{2}|(u+v)|^2 + \tfrac{R}{2}|v|^2 &\leq \Phi_R(u) + \delta. \end{aligned}$$

Which yields the useful inequality,

$$\frac{(R-\rho)}{2}|v|^2 \leq \Phi_R(u) + \frac{\rho}{2}|u|^2 - f(\bar{x}) + \delta \qquad \forall\, \delta > 0. \tag{2.8}$$

As $\delta \to 0$ we see;

$$\begin{aligned} W_R(u) &= \arg\min_{v \in V}\{f(\bar{x} + (u+v)) + \tfrac{R}{2}|v|^2\} \\ &= \arg\min\{f(\bar{x} + (u+v)) + \tfrac{R}{2}|v|^2 : v \in (V \cap B_\sigma(0))\} \end{aligned}$$

where $\sigma := [\frac{2}{(R-\rho)}(\Phi_R(u) + \frac{\rho}{2}|u|^2 - f(\bar{x}))]^{1/2}$.

Since $V \bigcap B_\sigma(0)$ is a compact in $\Re^n$ and $h_R(u,v)$ is lsc we know the argmin is obtained. Thus $W_R(u)$ is nonempty for any u in the domain of Φ_R as desired.

Case II, general $\bar{w}$.

Let $\tilde{f}(x) := f(x) - \langle x, \bar{w}\rangle$, and let $\tilde{\Phi}_R$ and $\tilde{W}_R$ be the Quadratic Sub-Lagrangian and its related proximal map for $\tilde{f}$ with respect to R and U. Then $\tilde{\Phi}_R$ and $\tilde{W}_R$ satisfy all the conditions required for Case I. Moreover,

$$\begin{aligned}
\tilde{\Phi}_R(u) &= \inf_{v \epsilon V}\{\tilde{f}(\bar{x} + (u+v)) + \frac{R}{2}|v|^2\} \\
&= \inf_{v \epsilon V}\{f(\bar{x} + (u+v)) - \langle \bar{w}, \bar{x} + (u+v)\rangle + \frac{R}{2}|v|^2\} \\
&= \inf_{v \epsilon V}\{f(\bar{x} + (u+v)) - \langle \bar{w}_v, v\rangle + \frac{R}{2}|v|^2\} - \langle \bar{w}, \bar{x}\rangle - \langle \bar{w}_u, u\rangle.
\end{aligned}$$

Yielding the following useful equation:

$$\tilde{\Phi}_R(u) = \Phi_R(u) - \langle \bar{w}, \bar{x}\rangle - \langle \bar{w}_u, u\rangle \tag{2.9}$$

Therefore $\tilde{W}_R(u) \neq \emptyset \Leftrightarrow W_R(u) \neq \emptyset$, and we conclude the result for all $\bar{w}$.

$\diamond$

Having now shown the Quadratic Sub-Lagrangian maintains the most basic property of the original function, we seek to show that it is lower semi-continuous. As a result, we shall also learn that the related proximal mapping is outer semi-continuous with respect to Φ_R attentive convergence. To do this we shall apply the following facts to h_R.

Fact 2.8 *([8], Theorem 1.17 part (a) and Theorem 7.41 parts (a) and (b)) Let $f : \Re^n \times \Re^m \to \bar{\Re}$ be a proper lower semi-continuous function, and*

$$p(u) := \inf_{v \epsilon V} f(u,v), \qquad P(u) := \arg\min_{v \epsilon V} f(u,v).$$

If f is level-bounded in v locally uniformly in u then:

i) p is a lower semi-continuous function of $\Re^n \to \bar{\Re}$.

ii) P is a compact-valued mapping that is outer semi-continuous with respect to p attentive convergence.

iii) If also one has f continuous at $\bar{u} \epsilon \Re^n$ and $P(\bar{u}) \neq \emptyset$, then P is outer semi-continuous (not just in the p-attentive sense) at $\bar{u}$.

By f is level-bounded in v locally uniformly in u we mean that for each $\bar{u} \epsilon \Re^n$ and $\alpha \epsilon \Re$ there is a neighborhood N of $\bar{u}$ with a bounded set $B \subseteq \Re^n$

such that $\{v : f(u,v) \leq \alpha\} \subseteq B$ for all $u \epsilon N$. Instead of using this complicated definition we shall make use of Fact 2.9 in showing h_R is level-bounded in v locally uniformly in u. Before stating this, we first note that since U and V are perpendicular subspaces of $\Re^n$, we may, through a change of basis, apply Fact 2.8 to Φ_R.

Fact 2.9 *([8], Example 5.17 (b)) For a function $f : U \oplus V \to \bar{\Re}$, one has f(u,v) is level-bounded in u locally uniformly in v if and only if for each $\alpha \epsilon \Re$ the mapping*

$$u \mapsto \{v : f(u,v) \leq \alpha\}$$

is locally bounded.

By $u \mapsto \{v : f(u,v) \leq \alpha\}$ locally bounded we mean that given any $u \epsilon U$ there exists some $\delta > 0$ such that $B_\delta(u)$ is mapped to a bounded set. In Lemma 2.10 we shall actually show a stronger result for h_R, that for any $\delta > 0$ one finds $B_\delta(u)$ is mapped to a bounded set. Using this we conclude that h_R is level bounded in v locally uniformly in u, and from there gain the lower semi-continuity and outer semi-continuity results we desire.

Lemma 2.10 *Let f: $\Re^n \to \bar{\Re}$ be a proper lsc function that is subdifferentially regular at $\bar{x}$, prox-regular there for $\bar{w} \epsilon \partial f(\bar{x})$with respect to ρ and ε, and bounded below by a Quadratic also with respect to ρ. Let U and V be perpendicular subspaces of $\Re^n$.*

Then h_R is level bounded in v locally uniformly in u for any $R > \rho$.

Proof:

By the prox-regularity of f we may assume that the Quadratic bounding f from below is of the form

$$f(\bar{x} + (u+v)) \geq f(\bar{x}) + \langle \bar{w}, (u+v)\rangle - \frac{\rho}{2}|(u+v)|^2$$

Let $S_\alpha(u) := \{v \epsilon V : h_R(u,v) \leq \alpha\}$. Therefore for any $u \epsilon U$ we find,

$$\begin{aligned} S_\alpha(u) &= \{v \epsilon V : h_R(u,v) \leq \alpha\} \\ &= \{v \epsilon V : f(\bar{x} + (u+v)) - \langle \bar{w}_v, v\rangle + \tfrac{R}{2}|v|^2 \leq \alpha\} \\ &\subseteq \{v \epsilon V : f(\bar{x}) + \tfrac{(R-\rho)}{2}|v|^2 - \tfrac{\rho}{2}|u|^2 + \langle \bar{w}_u, u\rangle \leq \alpha\} \\ &\subseteq \{v \epsilon V : \tfrac{(R-\rho)}{2}|v|^2 \leq -f(\bar{x}) + \tfrac{\rho}{2}|u|^2 - \langle \bar{w}_u, u\rangle + \alpha\}. \end{aligned}$$

Thus for any $\delta > 0$ we have,

$$S_\alpha(B_\delta(u)) \subseteq \left\{v \epsilon V : \frac{(R-\rho)}{2}|v|^2 \leq \alpha - f(\bar{x}) + \max_{\tilde{u} \epsilon B_\delta(u)} \{-\langle \bar{w}_u, \tilde{u}\rangle + \frac{\rho}{2}|\tilde{u}|^2\}\right\}$$

which is a bounded set.

Thus S_α is locally bounded, so by fact 2.9 we have h_R is level bounded in v locally uniformly in u.

♡

Corollary 2.11 *Let f: $\Re^n \to \bar{\Re}$ be a proper lsc function that is subdifferentially regular at $\bar{x}$, prox-regular there for $\bar{w} \in \partial f(\bar{x})$ with respect to ρ and ε, and bounded below by a Quadratic also with respect to ρ. Let U and V be perpendicular subspaces of $\Re^n$. Then Φ_R is lower semi-continuous, and W_R is outer semi-continuous with respect to Φ_R attentive convergence.*

Proof: Since h_R is proper, lower semi-continuous and level-bounded in v locally uniformly in u, we may apply Fact 2.8 parts (i) and (ii).

♠

2.3 Subgradient Properties

Having now established some of the basic properties of the Quadratic Sub-Lagrangian, we turn our attention to its subgradient mapping. To begin we show that by projecting the subgradients of $f(\bar{x})$ from a neighbourhood of $\bar{w}$ onto U one can find a subset of the regular subgradients of $\Phi_R(0)$. We then approach the question from the other side showing that the subgradients of $\Phi_R(u)$ are the projections of subgradients of $f(\bar{x} + (u + 0))$. This is not sufficient to show equality in the two sets, but it does give an excellent feel for the behaviour of $\partial\Phi_R$.

Proposition 2.12 *Let f: $\Re^n \to \bar{\Re}$ be a proper lsc function that is subdifferentially regular at $\bar{x}$, prox-regular there for $\bar{w} \in \partial f(\bar{x})$, with respect to ρ and ε, and bounded below by a Quadratic. Let U be any subspace of $\Re^n$ and $R \geq \rho$. Let $\Phi_R(u)$ be the Quadratic sub-Lagrangian of f with respect to $\bar{x}$, $\bar{w}$, U and R. Then if $w \in \partial f(\bar{x})$ with $|w - \bar{w}| < \varepsilon$ then $w_u \in \partial\Phi_R(0)$. In fact:*

$$P_U\left(\partial(f(\bar{x}))\bigcap B_\varepsilon(\bar{w})\right) \subseteq \hat{\partial}\Phi_R(0) \tag{2.10}$$

Where P_U is the projection map of $\Re^n$ onto U.

Proof:

By the prox-regularity of f, for $|(u+v)| < \varepsilon$ we have for any $R \geq \rho$,

$$\begin{aligned} f(\bar{x}+(u+v)) &\geq f(\bar{x}) + \langle w, (u+v)\rangle - \tfrac{R}{2}|(u+v)|^2 \\ f(\bar{x}+(u+v)) - \langle w_v, v\rangle + \tfrac{R}{2}|v|^2 &\geq f(\bar{x}) + \langle w_u, u\rangle - \tfrac{R}{2}|u|^2 \\ \Phi_R(u) &\geq \Phi_R(0) + \langle w_u, u\rangle - \tfrac{R}{2}|u|^2 \\ \Phi_R(u) &\geq \Phi_R(0) + \langle w_u, u\rangle + o(|u|) \end{aligned}$$

Therefore $w_u \in \hat{\partial}\Phi_R(0) \subseteq \partial\Phi_R(0)$, as desired.

♣

Corollary 2.13 *Let f: $\Re^n \to \bar{\Re}$ be a proper lsc function that is subdifferentially regular at $\bar{x}$, prox-regular there for $\bar{w} \in \partial f(\bar{x})$, with respect to ρ and ε, and bounded below by a Quadratic. Let U be any subspace of $\Re^n$ and $R > \rho$. Let $\Phi_R(u)$ be the Quadratic sub-Lagrangian of f with respect to $\bar{x}$, $\bar{w}$, U and R. Then,*

$$\bar{w}_u \in \hat{\partial}\Phi_R(0)$$

Proof: This follows directly from equation (2.10).

◇

It is worth noting here that since Lemma 2.12, and Corollary 2.13 both apply to any $R \geq \rho$, they both generalize perfectly to the convex case by letting $R = 0$ and $\varepsilon \to \infty$. This will be important in Chapter 3 where we shall use these results to show the existence of a maximal subspace, U, for which the Quadratic Sub-Lagrangian is differentiable (see Theorem 3.12). Since these results can be generalized perfectly to the convex case with the U-Lagrangian, it will follow that the same subspace is maximal in the set of choices for U for which the U-Lagrangian is differentiable.

Returning our attention to the question of subgradients we require the following fact:

Fact 2.14 *([8], Theorem 10.13) If $f(u,v)$ is a proper lsc function from $U \oplus V = \Re^n$ into $\bar{\Re}$ that is level bounded in v locally uniformly in u, with $p(u) = \inf_{v \in V}\{f(u,v)\}$ and $P(u) = \arg\min_{v \in V}\{f(u,v)\}$ then:*

$$\hat{\partial}p(u) \subseteq \bigcap_{v \in P(u)} \{w_u \in U : w \in \hat{\partial}f(u,v)\} \tag{2.11}$$

$$\partial p(u) \subseteq \bigcap_{v \in P(u)} \{w_u \in U : w \in \partial f(u,v)\} \tag{2.12}$$

Theorem 2.15 *Let f: $\Re^n \to \bar{\Re}$ be a proper lsc function that is subdifferentially regular at $\bar{x}$, prox-regular there for $\bar{w} \in \partial f(\bar{x})$ with respect to ρ and ε, and bounded below by a Quadratic also with respect to ρ. Let U and V be perpendicular subspaces of $\Re^n$, and P_U be the projection map of $\Re^n \to U$. Then:*

$$\hat{\partial}\Phi_R(u) \subseteq P_U(\hat{\partial} f(\bar{x} + (u+v))) \qquad \forall\ v \in W_R(u) \tag{2.13}$$

$$\partial\Phi_R(u) \subseteq P_U(\partial f(\bar{x} + (u+v))) \qquad \forall\ v \in W_R(u) \tag{2.14}$$

Proof:

By Lemma 2.10 we have h_R is level bounded in v locally uniformly in u. Since h_R is a proper lsc function we may apply Fact 2.14.

$$\begin{aligned}
\hat{\partial}(\inf_{v \in V} h_R(u,v)) &\subseteq \bigcap_{v \in W_R(u)} \{w_u : w \in \hat{\partial} h_R(u,v)\} \\
\hat{\partial}\Phi_R(u) &\subseteq \bigcap_{v \in W_R(u)} \{w_u : w \in \hat{\partial}[f(\bar{x} + u +)) - \langle \bar{w}_v, v\rangle + \tfrac{R}{2}|v|^2]\} \\
\hat{\partial}\Phi_R(u) &\subseteq \bigcap_{v \in W_R(u)} \{w_u : w \in \hat{\partial} f(\bar{x} + u + v) - \{\bar{w}_v\} + Rv\} \\
\hat{\partial}\Phi_R(u) &\subseteq \left\{ \begin{array}{r} w_u : (w_u + w_v + \bar{w}_v - Rv) \in \hat{\partial} f(\bar{x} + (u+v)) \\ \forall\ v \in W_R(u) \end{array} \right\} \\
\hat{\partial}\Phi_R(u) &\subseteq P_U(\hat{\partial} f(\bar{x} + (u+v))) \qquad \forall\ v \in W_R(u)
\end{aligned}$$

The only change in the proof of equation (2.14) is the replacement of the regular subgradient map with the subgradient map.

♡

Besides the intrinsic value of this theorem, it provides some insight into one use of the Quadratic Sub-Lagrangian. Specifically, if u minimizes Φ_R then $0 \in \partial\Phi_R(x)$. Theorem 2.15 then shows that $0 \in P_U(\partial f(\bar{x} + (u+v)))$ for all $v \in W_R(u)$. As a result we have if 0 is not in $\partial\Phi_R(u)$ then f is not minimized at $\bar{x} + (u+v)$ for any $v \in W_R(u)$.

Chapter 3

Results for "Good" *UV* Decompositions

Just as Einstein observed that space was not an absolute, but depended on the observer's movement in space, and that time was not an absolute, but depended on the observer's movement in time, so it is now realized that numbers are not absolute, but depend on the observer's movement in restaurants.

Douglas Adams

It Chapter 2 we discussed the properties that the Quadratic Sub-Lagrangian retained regardless of the UV decomposition. We found many strong results, including lower semi-continuity and similarities in the subgradient maps of $f(\bar{x} + (\cdot_u + 0))$ and Φ_R. In this Chapter we shall turn our attention to the question of what is the "best" UV decomposition. In order to define best we must decide what extra properties we desire Φ_R to have. Since Φ_R is lower semi-continuous, the next step would be to determine when Φ_R is continuous at 0. Later we shall find a stronger result, namely that there is a subspace for which the Quadratic Sub-Lagrangian is not only continuous at 0, but differentiable there (see Theorem 3.8). For now we begin with a simpler result, showing that for any subspace, U, for which $f(\bar{x} + (\cdot_u + 0))$ is continuous at 0 we have the Quadratic Sub-Lagrangian prox-regular and continuous at 0 for sufficiently large R.

3.1 UV Decompositions and Smoothness of the Quadratic Sub-Lagrangian

It is not surprising that there exists a subspace for which the Quadratic Sub-Lagrangian of f is continuous and prox-regular. Indeed, one needs only consider the trivial case of $U = \{0\}$ to confirm that such a subspace exists. A more interesting result will be that this subspace is often non-trivial. To begin with we examine any subspace on which $f(\bar{x}+(\cdot_u+0))$ is continuous at 0, later we shall show that this subspace is a superset of $N_{\partial f(\bar{x})}(\bar{w})$ whenever $\bar{w} \in \text{rint}\partial f(\bar{x})$ (see Lemma 3.4, and Theorem 3.7), and can therefore often be taken to be non-trivial. Before we continue with this result we shall need the following fact.

Fact 3.1 *([6], Theorem 3.2) If f is locally lsc at $\bar{x}$, then the following are equivalent.*

i) f is prox-regular at $\bar{x}$ for $\bar{w}$.

ii) The vector $\bar{w}$ is a proximal subgradient to f at $\bar{x}$, and there is an f-attentive δ-localization, T, of ∂f at $(\bar{x}, \bar{w})$ with a constant $r > 0$ such that $T + rI$ is monotone.

Theorem 3.2 *Let f: $\Re^n \to \bar{\Re}$ be a proper lsc function that is subdifferentially regular at $\bar{x}$, prox-regular there for $\bar{w} \in \partial f(\bar{x})$with respect to ρ and ε, and bounded below by a Quadratic also with respect to ρ. Let U be a subspace on which $f_U(u) := f(\bar{x}+(u,0))$ is continuous at 0, and $V = U^{\perp}$. Then whenever $R > \rho$ one finds, Φ_R is continuous at 0, and prox-regular there for $\bar{w}_u$.*

Proof: Case I: $\bar{w} = 0$.

By the prox-regularity of f we may assume without loss of generality that for all $x \in \Re^n$ one finds $f(x) \geq f(\bar{x}) - \frac{\rho}{2}|x-\bar{x}|^2$. Thus for any $u \in U, v \in V$,

$$\begin{aligned} f(\bar{x}+(u+v)) &\geq f(\bar{x}) - \tfrac{\rho}{2}|(u+v)|^2 \\ f(\bar{x}+(u+v)) + \tfrac{R}{2}|v|^2 &\geq f(\bar{x}) - \tfrac{\rho}{2}|u|^2 \\ \Phi_R(u) &\geq f(\bar{x}) - \tfrac{\rho}{2}|u|^2 \end{aligned}$$

Moreover, $\Phi_R(u) = \inf_{v \in V}\{f(\bar{x}+(u+v)) + \frac{R}{2}|v|^2\} \leq f(\bar{x}+(u+0))$.

So we have the estimates; $f(\bar{x}) - \frac{\rho}{2}|u|^2 \leq \Phi_R(u) \leq f_U(u)$. Where $f_U(u) = f(\bar{x}+(u+0))$. By the squeeze theorem we have Φ_R continuous at 0.

In showing the prox-regularity of Φ_R at 0 for $\bar{w}_u$ we note $\bar{w}_u \in \partial\Phi_R(0)$ by Corollary 2.13. Now we shall demonstrate Fact 3.1 can be applied to Φ_R.

Since f is prox-regular at $\bar{x}$ for 0, there exists (by Fact 3.1) $\delta > 0$ and $r > 0$ such that $T + rI$ is monotone, where T is the f-attentive δ-localization of ∂f at $(\bar{x}, 0)$.

Let $S_{\tilde{\delta}}$ be the Φ_R-attentive $\tilde{\delta}$-localization of $\partial \Phi_R$ at $(0,0)$, that is;

$$S_{\tilde{\delta}} := \begin{cases} \{w \in \partial \Phi_R(u) : |w| < \tilde{\delta}\} & if\ |u| < \tilde{\delta}, |\Phi_R(u) - \Phi_R(0)| < \tilde{\delta} \\ \emptyset & otherwise \end{cases}$$

If P_U is the projection map of $\Re^n \to U$, then by showing the existence of some $\tilde{\delta} > 0$ such that $S_{\tilde{\delta}}(u) \subseteq P_U(T(u,v))$ for all u and $v \in W_R(u)$, we will have $S_{\tilde{\delta}} + rI$ monotone. Thus Fact 3.1 will show Φ_R prox-regular at 0 for 0 (recall Φ_R is lower semi-continuous by Corollary 2.11).

So we turn our efforts to showing $\tilde{\delta}$ exists as required. First note;

$$\begin{aligned} |\Phi_R(u) - \Phi_R(0)| < \tilde{\delta} \quad &\Rightarrow \quad |f(\bar{x} + (u+v)) + \tfrac{R}{2}|v|^2 - f(\bar{x})| < \tilde{\delta}, \\ &\qquad\qquad \forall\, v \in W_R(u) \\ &\Rightarrow \quad |f(\bar{x} + (u+v)) - f(\bar{x})| < \tilde{\delta} + \tfrac{R}{2}|v|^2, \\ &\qquad\qquad \forall\, v \in W_R(u) \end{aligned}$$

From Proposition 2.7, (equation (2.8)) and the fact $\Phi_R(u) \leq f(\bar{x}+(u+0))$ we have:

$$\begin{aligned} \tfrac{R}{2}|v|^2 &\leq \tfrac{R}{2}\tfrac{2}{(R-\rho)}(\Phi_R(u) - f(\bar{x}) + \tfrac{\rho}{2}|u|^2) \\ &\leq \tfrac{R}{R-\rho}(f(\bar{x} + (u+0)) - f(\bar{x}) + \tfrac{\rho}{2}|u|^2) \end{aligned}$$

Now f_U is continuous near $\bar{x}$, we have the existence of some $\tilde{\delta} > 0$ such that,

$$|u| < \tilde{\delta} \Rightarrow |f(\bar{x} + (u+0) - f(\bar{x})| < \frac{\delta(R-\rho)}{2R}.$$

Without loss of generality $\tilde{\delta} < 1$, so we also have $\tilde{\delta}^2 < \tilde{\delta}$. Thus we find for any $|u| < \tilde{\delta}$,

$$\frac{R}{2}|v|^2 \leq \frac{R}{R-\rho}\left(\frac{\delta(R-\rho)}{2R} + \frac{\rho}{2}|u|^2\right) = \left(\frac{R}{R-\rho}\right)\frac{\rho}{2}|u|^2 < \frac{\delta}{2} + \left(\frac{R}{R-\rho}\right)\frac{\rho\tilde{\delta}}{2}$$

So,

$$|\Phi_R(u) - \Phi_R(0)| < \tilde{\delta} \Rightarrow |f(\bar{x} - (u+v)) - f(\bar{x})| < \tilde{\delta}\left(1 + \frac{R\rho}{R-\rho}\right) + \frac{\delta}{2}$$

Selecting $\tilde{\delta}$ sufficiently small we can assure that $\tilde{\delta} < \frac{\delta}{2}$ and, $\tilde{\delta}(1+\frac{R\rho}{R-\rho}) < \frac{\delta}{2}$. (In fact the later implies the first).

Therefore,

$$S_{\bar{\delta}} := \begin{cases} \{w \in \partial\Phi_R(u) : |w| < \bar{\delta}\} & |u| < \bar{\delta}, |\Phi_R(u) - \Phi_R(0)| < \bar{\delta} \\ \emptyset & otherwise \end{cases}$$

$$S_{\bar{\delta}} \subseteq \begin{cases} \{w \in \partial\Phi_R(u) : |w| < \bar{\delta}\} & |u| < \delta, |\Phi_R(u) - \Phi_R(0)| < \delta \\ \emptyset & otherwise \end{cases}$$

$$\subseteq \begin{cases} \left\{ \begin{array}{r} w \in P_U(\partial f(\bar{x} + u + v)) : \\ |w| < \delta \end{array} \right\} & \begin{array}{l} |u| < \delta, \\ |\Phi_R(u) - \Phi_R(0)| < \delta, \\ v \in W_R(u) \end{array} \\ \emptyset & otherwise \end{cases}$$

$$\subseteq \begin{cases} \left\{ \begin{array}{r} w \in P_U(\partial f(\bar{x} + u + v)) : \\ |w| < \delta \end{array} \right\} & \begin{array}{l} |u| < \delta, \\ |f(\bar{x} + u + v) - f(\bar{x})| < \delta, \\ v \in W_R(u) \end{array} \\ \emptyset & otherwise \end{cases}$$

$$\subseteq \; P_U(T(u + v))$$

Case II, general $\bar{w}$.

Let $\tilde{f}(x) = f(x) - \langle \bar{w}, x \rangle$. Let $\tilde{\Phi}_R$ be $\tilde{f}$'s corresponding Quadratic sub-Lagrangian. Then $\tilde{f}$ satisfies the requirements of case I, and as shown in equation (2.9),

$$\tilde{\Phi}_R(u) = \Phi_R(u) - \langle \bar{w}, \bar{x} \rangle - \langle \bar{w}_u, u \rangle$$
$$\Phi_R(u) - \langle \bar{w}, \bar{x} \rangle = \tilde{\Phi}_R(u) + \langle \bar{w}_u, u \rangle$$

So $\Phi_R(u)$ is continuous at 0, and is prox-regular there for $\bar{w}_u$.

♣

This yields an interesting corollary that will be of use later.

Corollary 3.3 *Let f: $\Re^n \to \bar{\Re}$ be a proper lsc function that is subdifferentially regular at $\bar{x}$, prox-regular there for $\bar{w} \in \text{rint}\partial f(\bar{x})$with respect to ρ and ε, and bounded below by a Quadratic (also with respect to ρ). Let U be a subspace on which $f(\bar{x} + (u + 0))$ is continuous at 0, and $R > \rho$. Then W_R as defined in equation (2.2) is outer semi-continuous at 0.*

Proof: By Lemma 2.10 and Theorem 3.2 we have h_R level-bounded in v locally uniformly in u, and Φ_R continuous at 0, therefore we apply Fact 2.8 part (iii) and the proof is complete.

◇

3.2 Effects of UV Decomposition on f

Theorem 3.2 shows us that for any subspace, U, along which f is continuous the Quadratic Sub-Lagrangian is continuous and prox-regular; however, it gives us very little idea what these subspaces look like. Our next goal therefore is to examine less vague possibilities for the subspace U. In "The U-Lagrangian of a Convex Function", Lemaréchal, Oustry, and Sagastizábal, used $U = N_{\partial f(\bar{x})}(\bar{w})$, where $\bar{w} \in \text{rint}\partial f(\bar{x})$, from which we draw inspiration. Lemaréchal, Oustry, and Sagastizábal, found that on this subspace the U-Lagrangian was differentiable at 0 ([2], Theorem 3.3 (ii)). They did not mention in their paper how they chose this subspace, but in exploring this subspace further we shall show that it is maximal in the set of all subspaces for which Φ_R is strictly differentiable.

Lemma 3.4 *Let f: $\Re^n \to \bar{\Re}$ be a proper lsc function that is subdifferentially regular at $\bar{x}$. Let $\bar{w} \in \text{rint}\partial f(\bar{x})$. Then the following subspaces are equal to $N_{\partial f(\bar{x})}(\bar{w})$:*

i) U_1 := The subspace perpendicular to the affine plane of $\partial f(\bar{x},)$

ii) $U_2 := \{d \in \Re^n : \sup_{d \in \partial f(\bar{x})} \langle w, d\rangle = \inf_{d \in \partial f(\bar{x})} \langle w, d\rangle\}$,

iii) $U_3 := \{d \in \Re^n : df(\bar{x}, d) = -df(\bar{x}, -d)\}$,

iv) $U_4 := \{d \in \Re^n : \langle w, d\rangle = \langle \bar{w}, d\rangle \;\forall\; w \in \partial f(\bar{x})\}$.

In order to prove lemma 3.4 we will make use of the following fact:

Fact 3.5 *([8], Theorem 8.30) If $f : \Re^n \to \bar{\Re}$ is subdifferentially regular at $\bar{x}$ then one finds,*

$$df(\bar{x}, d) = \sup_{w \in \partial f(\bar{x})} \langle w, d\rangle.$$

We may now proceed with the proof in question.

Proof of Lemma 3.4:

We begin by examining U_1. Since f is subdifferentially regular at $\bar{x}$ we have $N_{\partial f(\bar{x})}(\bar{w})$ is convex. Therefore, since $\bar{w} \in \text{rint}\partial f(\bar{x})$, we have $d \in N_{\partial f(\bar{x})}(\bar{w})$ if and only if d satisfies $\langle w - \bar{w}, d\rangle = 0 \;\forall\; w \in \partial f(\bar{x})$. Which is true if and only if $\langle w_1 - \bar{w}, d\rangle = 0 = \langle w_2 - \bar{w}, d\rangle \;\forall\; w_1, w_2 \in \partial f(\bar{x})$. That is $\langle d, w_1 - w_2\rangle = 0 \;\forall\; w_1, w_2 \in \partial f(\bar{x})$. Therefore $U_1 = N_{\partial f(\bar{x})}(\bar{w})$.

It is clear that $U_2 = U_4$, so next we show that these are equal to $N_{\partial f(\bar{x})}(\bar{w})$. Suppose $d \in U_4$ then $\langle w - \bar{w}, d\rangle = 0 \,\forall\, w \in \partial f(\bar{x})$, hence $d \in N_{\partial f(\bar{x})}(\bar{w})$.

Conversely, suppose $d \in N_{\partial f(\bar{x})}(\bar{w})$ and $w \in \partial f(\bar{x})$ then we shall show $\langle w - \bar{w}, d\rangle = 0$, making $d \in U_4$.

If $w = \bar{w}$ the result is trivial, so we examine the case where $w - \bar{w} \neq 0$. Since f is subdifferentially regular at $\bar{x}$, we have $\hat{\partial} f(\bar{x}) = \partial f(\bar{x})$, so $\partial f(\bar{x})$ is convex. Thus $w - \bar{w}$ is parallel to the affine hull of $\partial f(\bar{x})$. Hence there exists some $\gamma > 0$ such that for all $0 < \eta < \gamma$ we have $v = (w - \bar{w})$ satisfies $\bar{w} + \eta v \in \partial f(\bar{x})$ and $\bar{w} - \eta v \in \partial f(\bar{x})$. So, by definition of $N_{\partial f(\bar{x})}(\bar{w})$ we have,

$$\limsup_{\substack{w \to \bar{w} \\ w \neq \bar{w} \\ w \in \partial f(\bar{x})}} \frac{\langle d, w - \bar{w}\rangle}{|w - \bar{w}|} \leq 0.$$

Thus,

$$\lim_{\eta \searrow 0} \langle d, \bar{w} + \eta v - \bar{w}\rangle \leq 0$$

$$\lim_{\eta \searrow 0} \langle d, \bar{w} - \eta v - \bar{w}\rangle \leq 0.$$

Therefore, for small but positive η we find $\langle d, \eta v\rangle \leq 0$ and $\langle d, -\eta v\rangle \leq 0$. Which shows $\langle d, w - \bar{w}\rangle = 0$.

We now turn our attention to U_3. Since f is subdifferentially regular at $\bar{x}$ we may apply Fact 3.5 yielding for any $d \in U_3$,

$$-df(\bar{x}, d) = df(\bar{x}, -d) = \sup_{w \in \partial f(\bar{x})} \langle w, -d\rangle$$

$$df(\bar{x}, d) = \inf_{w \in \partial f(\bar{x})} \langle w, d\rangle.$$

So,

$$\inf_{w \in \partial f(\bar{x})} \langle w, d\rangle = df(\bar{x}, d) = \sup_{w \in \partial f(\bar{x})} \langle w, d\rangle \Rightarrow U_3 \subseteq U_2.$$

Conversely if $d \in U_2$ then

$$df(\bar{x}, d) = \sup_{w \in \partial f(\bar{x})} \langle w, d\rangle = \inf_{w \in \partial f(\bar{x})} \langle w, d\rangle = -\sup_{w \in \partial f(\bar{x})} \langle w, -d\rangle = -df(\bar{x}, -d)$$

Thus $d \in U_3$, and the proof is complete.

♡

Having now established a greater understanding of $N_{\partial f(\bar{x})}(\bar{w})$ we can begin to examine how $f(\bar{x} + (\cdot_u + 0))$ behaves when $U = N_{\partial f(\bar{x})}(\bar{w})$. Interestingly, we shall find not only is $f(\bar{x} + (\cdot_u + 0))$ continuous at 0 for this subspace, but it is actually strictly differentiable at 0. To achieve this we shall need one more fact.

Fact 3.6 *([8], Theorem 9.18) Let $\bar{x} \in dom(f)$, and f: $\Re^n \to \bar{\Re}$. Then the following are equivalent to f being strictly differentiable at $\bar{x}$:*

i) f is Lipschitz continuous at $\bar{x}$ and has at most one subgradient there,

ii) f is continuous near $\bar{x}$, and f and $-f$ are subdifferentially regular functions,

iii) f is locally lower semi-continuous at $\bar{x}$, and $\hat{d}f(\bar{x}, \cdot)$ is a linear function,

iv) f is locally lower semi-continuous at $\bar{x}$ and $\hat{d}f(\bar{x}, d) = -\hat{d}f(\bar{x}, -d)$ for all $d \in \Re^n$.

It should be noted here that f Lipschitz continuous at $\bar{x}$ refers to the existence of some neighbourhood, N, of $\bar{x}$ and some $K > 0$ such that whenever $x \in N$ one finds $|f(x) - f(\bar{x})| < K|x - \bar{x}|$, we then refer to K as the Lipschitz constant of f.

Theorem 3.7 *Let f: $\Re^n \to \bar{\Re}$ be a proper lsc function that is subdifferentially regular at $\bar{x}$. Let $\bar{w} \in rint\partial f(\bar{x})$and $U = N_{\partial f(\bar{x})}(\bar{w})$. Define $f_U : U \to \Re^n$ as $f_U(u) = f(\bar{x} + (u + 0))$. Then f_U is strictly differentiable at 0, and therefore continuous there. In fact this is true for U any subspace of $N_{\partial f(\bar{x})}(\bar{w})$.*

Proof:

By Lemma 3.4, $U = \{d \in \Re^n : df(\bar{x}, d) = -df(\bar{x}, -d)\}$. Furthermore, since f is subdifferentially regular at $\bar{x}$ we have $df(\bar{x}) = \hat{d}f(\bar{x})$. Therefore we find

$$\hat{d}f_U(0, d) = df_U(0, d) = -df_U(0, -d) = -\hat{d}f_U(0, -d) \qquad \forall\ d \in U$$

Since f is lsc if follows that f_U is, and Fact 3.6 completes the proof.

♠

3.3 *UV* Decomposition and First Order Behaviour of the Quadratic Sub-Lagrangian

Having now discovered a subspace on which $f(\bar{x} + (\cdot_u + 0))$ is continuous at 0, we seek to examine what further properties this subspace has in the construction of Quadratic Sub-Lagrangians. In this section we shall turn our attention to how the Quadratic Sub-Lagrangian behaves under the circumstance of U being a subspace of $N_{\partial f(\bar{x})}(\bar{w})$. We begin with the very interesting result that, (for this choice of subspace U), Φ_R is strictly differentiable at 0.

Theorem 3.8 *Let $f: \Re^n \to \bar{\Re}$ be a proper lsc function that is subdifferentially regular at $\bar{x}$, prox-regular there for $\bar{w} \in \text{rint}\partial f(\bar{x})$ with respect to ρ and ε, and bounded below by a Quadratic. Let U be any subspace of $N_{\partial f(\bar{x})}(\bar{w})$, and $R > \rho$. If Φ_R is the Quadratic Sub-Lagrangian with respect to f, U, and R then Φ_R is Lipschitz continuous at 0, and strictly differentiable there with $\{\nabla\Phi_R(0)\} = \{\bar{w}_u\} = \partial\Phi_R(0)$.*

Proof: Case I, $\bar{w} = 0$.

Let $f_U(u) = f(\bar{x}+(u+0))$. First we note that, just as in Theorem 3.2, we have $f(\bar{x}) - \frac{\rho}{2}|u|^2 \leq \Phi_R(u) \leq f_U(u)$. By Theorem 3.7, and Fact 3.6 we have f_U is Lipschitz continuous at 0. Since $f(\bar{x}) - \frac{\rho}{2}|u|^2$ is also Lipschitz continuous there we may simply take the larger of the two Lipschitz constants to provide a Lipschitz constant for Φ_R. Therefore Φ_R is Lipschitz continuous at 0.

In order to apply Fact 3.6 we now turn our attention to showing $\partial\Phi_R(0)$ is the singleton $\{\bar{w}_u\}$. We know $\bar{w}_u \in \partial\Phi_R(0)$ by Corollary 2.13, so only have to show its uniqueness.

Suppose $\tilde{w} \in \partial\Phi_R(0)$, then by Theorem 2.15, equation (2.14), combined with Theorem 2.6 we have $\tilde{w} \in P_U\partial f(\bar{x})$. Therefore $\tilde{w} = P_U(\tilde{w} + \tilde{w}_v)$ for some $\tilde{w}_v \in V$ such that $\tilde{w} + \tilde{w}_v \in \partial f(\bar{x})$. By Lemma 3.4 we have U is the subspace perpendicular to the affine hull of $\partial f(\bar{x})$, therefore V is the subspace parallel to the affine hull of $\partial f(\bar{x})$. Therefore for any $w \in \partial f(\bar{x})$ we have $w - \bar{w} \in V$, thus $(\tilde{w} + \tilde{w}_v) - \bar{w} \in V$. It follows that,

$$\begin{aligned} P_U((\tilde{w} + \tilde{w}_v) - \bar{w}) &= \{0\} \\ \tilde{w}_u - \bar{w}_u &= \{0\} \\ \tilde{w}_u &= \bar{w}_u, \end{aligned}$$

and we conclude $\partial\Phi_R(0) = \{\bar{w}_u\}$.

Case II, general $\bar{w}$.

Let $\tilde{f}(x) = f(x) - \langle\bar{w}, x\rangle$. We first show that $U = N_{\partial f(\bar{x})}(\bar{w}) = N_{\partial\tilde{f}(\bar{x})}(0)$.

Indeed by Lemma 3.4 $N_{\partial f(\bar{x})}(\bar{w})$ we equivalent to the space normal to the affine plane of $\partial f(\bar{x})$. By the same proof $N_{\partial\tilde{f}(\bar{x})}(0)$ is the space normal to the affine place of $\partial\tilde{f}(0) = \partial(f(x) - \langle\bar{w}, x\rangle) = \partial f(\bar{x}) - \{\bar{w}\}$. Clearly these to subspaces are equivalent.

Next, let $\tilde{\Phi}_R$ be $\tilde{f}$'s Quadratic Sub-Lagrangian with respect to U, then $\tilde{f}$ satisfies the requirements of Case I. Moreover

$$\Phi_R(u) = \tilde{\Phi}_R(u) + \langle\bar{w}, \bar{x}\rangle + \langle\bar{w}_u, u\rangle,$$

as shown in equation (2.9). Therefore we have Φ_R Lipschitz continuous at 0, as the addition of a linear factor does not effect this property.

Furthermore we have,

$$\partial\tilde{\Phi}_R(u) = \partial(\Phi_R(u) - \langle\bar{w},\bar{x}\rangle - \langle\bar{w}_u, u\rangle) = \partial\Phi_R(u) - \{\bar{w}_u\}.$$

Thus, $\partial\Phi_R(0)$ is the singleton $\{\bar{w}_u\}$ and Fact 3.6 completes the proof.

♣

Theorem 3.8 shows immediately that any $U \subseteq N_{\partial f(\bar{x})}(\bar{w})$ forms a good choice for the base of our UV decomposition. With the addition of one previously known Proposition we can further strengthen the appeal of $U \subseteq N_{\partial f(\bar{x})}(\bar{w})$ by showing that for this decomposition Φ_R gains the further properties of being subdifferentially regular at 0, and a lower-$\mathcal{C}^2$ function near 0.

Fact 3.9 *([8], Proposition 13.33) If f is prox-regular at a point $\bar{x}$ where it is strictly differentiable, it must be lower-$\mathcal{C}^2$ around $\bar{x}$.*

Corollary 3.10 *Let f: $\Re^n \to \bar{\Re}$ be a proper lsc function that is subdifferentially regular at $\bar{x}$, prox-regular there for $\bar{w} \in \operatorname{rint}\partial f(\bar{x})$ with respect to ρ and ε, and bounded below by a Quadratic. Let U be any subspace of $N_{\partial f(\bar{x})}(\bar{w})$, and $R > \rho$. If Φ_R is the Quadratic Sub-Lagrangian with respect to f, U, and R then Φ_R is subdifferentially regular at 0, and lower-$\mathcal{C}^2$ around this point.*

Proof:

By Theorem 3.8 we have $\partial\Phi_R(0) = \{\bar{w}_u\}$. Since $\bar{w}_u \in \hat{\partial}\Phi_R(0)$ (by Corollary 2.13) and $\hat{\partial}\Phi_R(0) \subseteq \partial\Phi_R(0)$ we have equality of the regular and general subgradient maps. Corollary 2.11 shows Φ_R is lower semi-continuous, and therefore Φ_R is subdifferentially regular at 0.

Next Theorem 3.2 shows Φ_R is prox-regular at 0 for $\bar{w}_u$. Since $\partial\Phi_R(0) = \{\bar{w}_u\}$ we can restate this as Φ_R is prox-regular at 0. Fact 3.9 and Theorem 3.8 therefore complete the proof.

♠

We have now shown that $N_{\partial f(\bar{x})}(\bar{w})$ is an excellent choice for the subspace U, so have an understanding as to why Lemaréchal, Oustry, and Sagastizábal selected it. But the question of whether it is the "best" subspace still arises. In the next theorem we see that $N_{\partial f(\bar{x})}(\bar{w})$ is the largest subspace for which the Quadratic Sub-Lagrangian is differentiable at 0. Besides answering the question of the "best" subspace for U, Theorem 3.12 provides us with two interesting results. First whenever Φ_R is differentiable at 0 it is strictly differentiable there, and lower-$\mathcal{C}^2$ in a neighbourhood of 0. Secondly Theorem 3.12 generalizes perfectly to the convex case, since for convex functions one can take $R = \rho = 0$; therefore Lemaréchal, Oustry, and Sagastizábal used, for their choice of U, the largest subspace for which their results would hold. Before proving this final result for Chapter 3, we require one simple fact.

Fact 3.11 *([8], Exercise 8.8)*
If f is differentiable at $\bar{x}$, then $\hat{\partial} f(\bar{x}) = \{\nabla f(\bar{x})\}$

Theorem 3.12 *If f: $\Re^n \rightarrow \bar{\Re}$ is a proper lsc function that is subdifferentially regular at $\bar{x}$, prox-regular there for $\bar{w} \in \text{rint}\partial f(\bar{x})$ with respect to ρ and ε, and bounded below by a Quadratic. Then for any $R \geq \rho$, $U = N_{\partial f(\bar{x})}(\bar{w})$ is maximal in the set of all subspaces for which the Quadratic Sub-Lagrangian of f with respect to R and U is differentiable at 0.*

Proof:

Suppose U is not contained in $N_{\partial f(\bar{x})}(\bar{w})$, then $U \bigcap T_{\partial f(\bar{x})}(\bar{w}) \neq \emptyset$. Since $\bar{w} \in \text{rint}\partial f(\bar{x}) = \text{rint}\hat{\partial} f(\bar{x})$ there exists some $\tilde{w} \in U$ such that $\tilde{w} \neq \bar{w}$, $(\tilde{w} + \bar{w}_v) \in \partial f(\bar{x})$, and $|(\tilde{w} + \bar{w}_v) - \bar{w}| < \varepsilon$. By Lemma 2.12 we notice $P_U(\tilde{w} + \bar{w}_v) = \tilde{w} \in \hat{\partial}\Phi_R(0)$. Since $\tilde{w} \neq \bar{w}_u = \nabla\Phi_R(0)$, we have $\hat{\partial}\Phi_R(0) \neq \{\nabla\Phi_R(0)\}$, and the contrapositive of Fact 3.11 shows that Φ_R cannot be differentiable at 0.

◇

Chapter 4

Second Order Properties of the Quadratic Sub-Lagrangian

And out of the confusion,
where the river meets the sea,
something new will arrive,
something better will arrive.

Sting

So far we have focused on studying the first order behaviour of the Quadratic Sub-Lagrangian. We began by showing that the Quadratic Sub-Lagrangian was a well defined envelope function. From there we demonstrated that by the correct UV decomposition we could ensure that this envelope was well behaved near the origin. We even went to the point where we could ensure strict differentiability there and showing that Φ_R is lower-C^2 near 0. We now turn our attention to the second order behavior of the Quadratic Sub-Lagrangian.

We say a function has a quadratic expansion at $\bar{x}$ if it is differentiable at $\bar{x}$ and there is an operator $A : \Re^n \to \Re^n$ such that

$$f(x) = f(\bar{x}) + \langle \nabla f(\bar{x}), x - \bar{x}\rangle + \frac{1}{2}\langle x - \bar{x}, A(x - \bar{x})\rangle + o(|x - \bar{x}|^2),$$

where $o(|x - \bar{x}|^2)$ refers to a term with the property that $\frac{o(|x-\bar{x}|^2)}{|x-\bar{x}|^2} \to 0$ as $x \to \bar{x}$. If A has the additional property that

$$\nabla f(x) = \nabla f(\bar{x}) + A(x - \bar{x}) + o(|x - \bar{x}|)$$

for x in the domain of ∇f then we say f is twice differentiable at $\bar{x}$. In such a case A is necessarily unique so we label it by $\nabla^2 f(\bar{x})$ and refer to it as the Hessian of f. It would be too much to expect the Quadratic Sub-Lagrangian to always provide us with a quadratic expansion at 0, so next we seek situations for which this occurs. Before developing some of these properties we shall provide some inspiration for the search. More precisely we shall show that a Quadratic expansion of Φ_R can be used (to a degree) to create a quadratic expansion of the original function.

4.1 Definition of H_R

Once again we begin with a proper lower semi-continuous function, $f : \Re^n \to \bar{\Re}$, that is subdifferentially regular at $\bar{x}$, prox-regular there for $\bar{w} \in \operatorname{rint}\partial f(\bar{x})$with respect to ρ and ε, and bounded below by a Quadratic also with respect to ρ. Let $U = V^{\perp}$ be subspaces of $\Re^n$ and $R > \rho$. Suppose Φ_R has a quadratic expansion at 0; i.e.,

$$\Phi_R(u) = \Phi_R(0) + \langle \nabla\Phi_R(0), u\rangle + \frac{1}{2}\langle u, Au\rangle + o(|u|^2),$$

where $A : U \to U$. Immediately we note that Theorem 3.12 forces $U \subseteq N_{\partial f(\bar{x})}(\bar{w})$. We shall define H_R by

$$\begin{array}{rccc} H_R := & U \oplus V & \to & U \oplus V \\ & (u+v) & \mapsto & (Au - Rv). \end{array} \tag{4.1}$$

Since $W_R(u)$ is set valued, we shall encounter some difficulty in notation; to avoid this we create a selection function for W_R. That is we define $\omega : U \mapsto V \bigcup\{\infty\}$ to be any $v \in W_R(u)$ when $W_R(u)$ is nonempty and ∞ whenever $W_R(u) = \emptyset$. This actually creates a family of functions from U into $\bar{V}$, but for our purposes we consider ω to be any fixed one of them. Now we note that by Theorem 3.8 we have $\nabla\Phi_R(0) = \bar{w}_u$, and by Proposition 2.7, $W_R(u) \neq \emptyset$ for any u in the domain of Φ_R. Since a quadratic expansion of Φ_R implies differentiability of Φ_R, which in turn implies continuity of Φ_R, we have $\omega(u) \neq \infty$ when u is sufficently close to 0. Therefore we find:

$$f(\bar{x}+u+\omega(u)) - \langle \bar{w}_v, \omega(u)\rangle + \frac{R}{2}|\omega(u)|^2 = f(\bar{x}) + \langle \bar{w}_u, u\rangle + \frac{1}{2}\langle u, Au\rangle + o(|u|^2),$$

$$f(\bar{x}+u+\omega(u)) = f(\bar{x}) + \langle \bar{w}, (u+\omega(u))\rangle + \frac{1}{2}\langle (u+\omega(u)), (Au - R\omega(u))\rangle + o(|u|^2),$$

$$f(\bar{x}+u+\omega(u)) = f(\bar{x})+\langle\bar{w},(u+\omega(u))\rangle+\frac{1}{2}\langle(u+\omega(u)),H_R(u+\omega(u))\rangle+o(|u|^2).$$

Lastly we note that $\frac{o(|u|^2)}{|u|^2} \to 0$ as $u \to 0$ implies $\frac{o(|u|^2)}{|u|^2+|\omega(u)|^2} \to 0$ as $u \to 0$. Therefore we may replace the $o(|u|^2)$ with $o(|u|^2 + |\omega(u)|^2) = o(|u+\omega(u)|^2)$, yielding

$$f(x) = f(\bar{x}) + \langle\bar{w}, x-\bar{x}\rangle + \frac{1}{2}\langle x-\bar{x}, H_R(x-\bar{x})\rangle + o(|x-\bar{x}|^2),$$

where x is of the form $x = \bar{x} + (u+\omega(u))$. Although this is not a quadratic expansion of f itself, it does provide a quadratic expansion of f along the manifold defined by $\{\bar{x} + (u+v) : u \in U, v \in W_R(u)\}$.

4.2 Existence of Quadratic Expansions of Φ_R

Having now established a purpose for our search we shall seek some properties that ensure a quadratic expansion for Φ_R. To begin we provide an alternate approach to the problem.

Fact 4.1 *([5], Theorem 3.1) Suppose that f is prox-regular at $\bar{x}$ for $\bar{w}$. Then f is differentiable at $\bar{x}$ and has a second order expansion at $\bar{x}$ if and only if f is twice epi-differentiable at $\bar{x}$ and $f''_{\bar{x},\bar{w}}$ is finite everywhere.*

This fact inspires the next major result (Theorem 4.4) in which we develop an if and only if statement of when the Quadratic Sub-Lagrangian is twice epi-differentiable. Although this is weaker than a quadratic expansion it shall provide a base for our final results. In order to achieve this goal we shall apply the following fact.

Fact 4.2 *([6], Theorem 6.5) Let f be prox-regular at $\bar{x} = 0$ for $\bar{w} = 0$ with respect to ε and ρ, and let $\lambda \in (0, \frac{1}{\rho})$. Then f is twice epi-differentiable at 0 for 0 if and only if e_λ is twice epi-differentiable at 0 for 0.*

With this fact in mind we proceed to examine the Moreau envelope of the Quadratic Sub-Lagrangian. In effect we will take an envelope of an envelope in order to see when the later is twice epi-differentiable.

Lemma 4.3 *Let f: $\Re^n \to \bar{\Re}$ be a proper lsc function that is subdifferentially regular at $\bar{x}$, prox-regular there for $\bar{w} \in \partial f(\bar{x})$ with respect to ρ and ε, and bounded below by a Quadratic. Let U and V be any perpendicular subspaces of $\Re^n$, then*

$$e_{\frac{1}{R}} f(\bar{x} + (u + \frac{\bar{w}_v}{R})) = e_{\frac{1}{R}} \Phi_R(u) + \frac{1}{2R}|\bar{w}_v|^2.$$

Proof:

$$\begin{aligned}
& e_{\frac{1}{R}} f(\bar{x} + \tfrac{\bar{w}_v}{R} + (u+0)) \\
= \;& \inf_{x \in \Re^n} \{ f(x) + \tfrac{R}{2}|(\bar{x} + \tfrac{\bar{w}_v}{R} + (u+0)) - x|^2 \} \\
= \;& \inf_{\tilde{u} \in U} \{ \inf_{\tilde{v} \in V} \{ f(\bar{x} + (\tilde{u} + \tilde{v})) + \tfrac{R}{2}|\tfrac{\bar{w}_v}{R} + (u+0) - (\tilde{u} + \tilde{v})|^2 \} \} \\
= \;& \inf_{\tilde{u} \in U} \{ \inf_{\tilde{v} \in V} \{ f(\bar{x} + (\tilde{u} + \tilde{v})) + \tfrac{R}{2}|\tfrac{\bar{w}_v}{R} - \tilde{v}|^2 \} + \tfrac{R}{2}|u - \tilde{u}|^2 \} \\
= \;& \inf_{\tilde{u} \in U} \{ \inf_{\tilde{v} \in V} \{ f(\bar{x} + (\tilde{u} + \tilde{v})) + \tfrac{1}{2R}|\bar{w}_v|^2 - \langle \bar{w}_v, \tilde{v} \rangle + \tfrac{R}{2}|\tilde{v}|^2 \} + \tfrac{R}{2}|u - \tilde{u}|^2 \} \\
= \;& \inf_{\tilde{u} \in U} \{ \Phi_R(\tilde{u}) + \tfrac{1}{2R}|\bar{w}_v|^2 + \tfrac{R}{2}|u - \tilde{u}|^2 \}
\end{aligned}$$

Therefore:

$$e_{\frac{1}{R}} f(\bar{x} + (u + \frac{\bar{w}_v}{R})) = e_{\frac{1}{R}} \Phi_R(u) + \frac{1}{2R}|\bar{w}_v|^2. \tag{4.2}$$

♣

Theorem 4.4 *Let f: $\Re^n \to \bar{\Re}$ be a proper lsc function that is subdifferentially regular at $\bar{x}$, prox-regular there for $\bar{w} \in \partial f(\bar{x})$ with respect to ρ and ε, and bounded below by a Quadratic. Let U be any subspace along which $f_U := f(\bar{x} + (u+0))$ is continuous at 0, and $V = U^\perp$. Then for sufficently large R, Φ_R is twice epi-differentiable at 0 for $\bar{w}_u$ if and only if $e_{\frac{1}{R}} f(\bar{x} + (\cdot_u + \frac{\bar{w}_v}{R}))$ has this property.*

Proof: Case I: if $\bar{x} = 0$, and $\bar{w} = 0$.

By Theorem 3.2 we have Φ_R is prox-regular at $\bar{u} = 0$ for $\bar{w}_u = 0$ with respect to $R' = (R-\rho)/2$. Then $R > R'$, (which is true for any $R > 2\rho$) implies $\frac{1}{R} \in (0, \frac{1}{R'})$. Therefore we may apply Fact 4.2 to note Φ_R is twice epi-differentiable at 0 for $\bar{w}_u$ if and only if $e_{\frac{1}{R}}\Phi_R$ has this property. From equation (4.2) we have Φ_R is twice epi-differentiable at 0 for 0 if and only if $e_{\frac{1}{R}} f(\bar{x} + (\cdot_u + \frac{\bar{w}_v}{R}))$ is twice epi-differentiable at 0 for 0, which proofs case I.
Case II: general f

Let $\tilde{f}(x) = f(\bar{x} + x) - \langle \bar{w}, x \rangle$, then by similar calculations used in 2.7 (see equation (2.9)), we see

$$\Phi_R(u) = \tilde{\Phi}_R(u) + \langle \bar{w}_u, u \rangle,$$

which is twice epi-differentiable if and only if $\tilde{\Phi}_R$ maintains this property. Case I then completes the proof.

$\diamond$

Having now established an if and only if statement of twice epi-differentiability for the Quadratic Sub-Lagrangian we have reduced the problem to showing when the second epi-derivative is finite valued. Our next concern will be to show that if $f_U = f(\bar{x} + (\cdot_u + 0))$ has a second order expansion, then the Quadratic Sub-Lagrangian maintains this property. While proving this result we shall also show several other cases which guarantee a Quadratic expansion for Φ_R.

Theorem 4.5 *Let $f\colon \Re^n \to \bar{\Re}$ be a proper lsc function that is subdifferentially regular at $\bar{x}$, prox-regular there for $\bar{w} \in \mathrm{rint}\partial f(\bar{x})$ with respect to ρ and ε, and bounded below by a Quadratic. Let $U \subseteq N_{\partial f(\bar{x})}(\bar{w})$ be a subspace, and V be it's perpendicular subspace. Let $f_U = f(\bar{x} + (\cdot_u + 0))$. If $e_{\frac{1}{R}} f(\bar{x} + (\cdot_u + \frac{\bar{w}_v}{R}))$ is twice epi-differentiable at 0 for $\bar{w}_u$ then any of the following conditions imply Φ_R has a second order expansion at 0:*

i) f, or f_U has a quadratic expansion at $\bar{x}$,

ii) $f''_{\bar{x},\bar{w}}$, or $(f_U)''_{\bar{x},\bar{w}}$ is finite everywhere,

iii) $f \in C^{1+}$, or $f_U \in C^{1+}$ on some neighbourhood of $\bar{x}$,

iv) f_U is bounded above by a quadratic on some neighbourhood of $\bar{x}$,

v) There exists some $r > 0$, and $m \geq 2$ fixed such that for any $u \in U$ one can find $v \in V$ such that $h_R(u + v) \leq f(\bar{x}) + r|u|^m$,

vi) Condition (v) is true on some neighbourhood of 0.

Proof:

Since $e_{\frac{1}{R}} f(\bar{x} + (\cdot_u + \frac{\bar{w}_v}{R}))$ is twice epi-differentiable at 0 for $\bar{w}_u$, we have Φ_R is twice epi-differentiable there. Thus by Fact 4.1 Φ_R has a second order expansion if and only if $(\Phi_R)''_{0,\bar{w}_u}$ is finite everywhere. Since $(\Phi_R)''_{0,\bar{w}_u}$ is proper we need only show an upper bound for $(\Phi_R)''_{0,\bar{w}_u}$ in each case. As usual we begin with the case $\bar{w} = 0$,

Case I: $\bar{w} = 0$

Notice,

$$\begin{aligned} (\Phi_R)''_{0,\bar{w}_u}(d) &= e - \lim_{\tau \searrow 0} \frac{\Phi_R(\tau d) - \Phi_R(0)}{\frac{1}{2}\tau^2} \\ &\leq e - \lim_{\tau \searrow 0} \frac{f(\bar{x} + (\tau d + 0)) - f(\bar{x})}{\frac{1}{2}\tau^2}. \end{aligned} \tag{4.3}$$

i) If f has a quadratic expansion then we may apply the approximation $f(\bar{x} + (\tau d + 0)) = f(\bar{x}) + \langle(\nabla f(\bar{x}))_u, \tau d\rangle + \langle \tau d, A\tau d\rangle + o(|\tau d|^2)$, where $(\nabla f(\bar{x}))_u = \bar{w}_u = 0$. Therefore equation (4.3) yields,

$$(\Phi_R)''_{0,\bar{w}_u}(d) \leq e - \lim_{\tau \searrow 0} \frac{\langle \tau d, A\tau d\rangle + o(|\tau d|^2)}{\frac{1}{2}\tau^2} = 2\langle d, Ad\rangle,$$

thus $(\Phi_R)''_{0,\bar{w}_u}(d)$ is bounded above by $2\langle d, Ad\rangle < \infty$. The same proof works for the case of f_U having a quadratic expansion at $\bar{x}$.

ii) If $f''_{\bar{x},\bar{w}}$, or $(f_U)''_{\bar{x},\bar{w}}$ is finite everywhere, we need only notice that equation (4.3) shows $(\Phi_R)''_{0,\bar{w}_u}(d) \leq (f)''_{\bar{x},\bar{w}}(d)$ or $(\Phi_R)''_{0,\bar{w}_u}(d) \leq (f_U)''_{\bar{x},\bar{w}}(d)$ as appropriate.

iii) If $f \in \mathcal{C}^{1+}$ on some neighbourhood of $\bar{x}$, then by the mean value theorem we note for all $\tau d \in U$ there exists some $x_\tau = \bar{x} + \eta d, \eta \in (0, \tau)$ and $w_\tau \in \partial f(x_\tau) = \nabla f(x_\tau)$ such that

$$f(\bar{x} + \tau d) - f(\bar{x}) = \langle w_\tau, \tau d\rangle.$$

Since $f \in \mathcal{C}^{1+}$ we have the existence of some $K > 0$ such that $|w_\tau| \leq K|\tau d|$, when τ is sufficiently small. Thus (4.3) becomes,

$$(\Phi_R)''_{0,\bar{w}_u}(d) \leq e - \lim_{\tau \searrow 0} \frac{\langle w_\tau, \tau d\rangle}{\frac{1}{2}\tau^2} \leq e - \lim_{\tau \searrow 0} \frac{K|\tau d|^2}{\frac{1}{2}\tau^2} = 2K|d| < \infty.$$

The proof for $f_U \in \mathcal{C}^{1+}$ is identical.

iv), v), and vi) It will suffice in this case to prove vi), as iv) and v) can be thought of as less generalized cases of vi). Note without loss of generality the $v \in V$ such that such that $h_R(u + v) \leq r|u|^m$, can be considered to converge to 0 as $u \to 0$ (see equation (2.8)). Therefore applying the assumption of vi) to $(\Phi_R)''_{0,\bar{w}_u}(d)$ we see,

$$\begin{aligned}
(\Phi_R)''_{0,\bar{w}_u}(d) &= e - \lim_{\tau \searrow 0} \frac{\Phi_R(\tau d) - \Phi_R(0)}{\frac{1}{2}\tau^2} \\
&= e - \lim_{\tau \searrow 0} \frac{h_R(\tau d + v) - f(\bar{x})}{\frac{1}{2}\tau^2}, \ v \in W_R(\tau d) \\
&\leq e - \lim_{\tau \searrow 0} \frac{r|\tau d|^m}{\frac{1}{2}\tau^2} < \infty.
\end{aligned}$$

Case II, general $\bar{w}$.

Let $\tilde{f}(x) := f(x) - \langle \bar{w}, x\rangle$, and $\tilde{\Phi}$ be its Quadratic sub-Lagrangian. Then by equation (2.9) we note that Φ_R has a second order expansion if and only if $\tilde{\Phi}_R$ has one. Since $\tilde{\Phi}_R$ satisfies case I the proof is complete.

♡

Chapter 5

Conclusion, and Further Areas of Exploration

Still, round the corner there may wait
A new road or a secret gate;

J. R. R. Tolkien

5.1 Summary of Results

The purpose of this thesis was to develop a new envelope function that could be applied to the broad range of functions known as prox-regular. To do this we combined two previously developed envelopes, the U-Lagrangian and the Moreau envelope. The first of these, the U-Lagrangian, was developed recently by Lemaréchal, Oustry, and Sagastizábal, in their paper entitled "The U-Lagrangian of a Convex Function" [2]. This paper provided a base for this thesis and much of the inspiration on the results regarding "good" UV decompositions. The second envelope used, the Moreau envelope, was developed in 1963 by J. J. Moreau [3] [4], and later extended to prox-regular equations by R. A. Poliquin and R. T. Rockafellar in 1996 [5]. These works were used as tools in the study of Quadratic Sub-Lagrangian.

The work of Lemaréchal, Oustry, and Sagastizábal provided three major results on the U-Lagrangian. The first of these was given a finite valued convex function the U-Lagrangian is also a finite valued convex function. Next they showed that the U-Lagrangian was differentiable at 0, with derivative related to the subgradient used in it's definition. Lastly they developed some second order theory for the U-Lagrangian. In this they showed how a 2nd derivative of the U-Lagrangian could be used as a second order expansion of the original function, and provide an if and only if proof regarding when the U-Lagrangian was twice differentiable.

In extending the first major result we sought to show first that the Quadratic Sub-Lagrangian of a proper prox-regular function was also proper and prox-regular. Unlike Lemaréchal, Oustry, and Sagastizábal, who assumed a specific UV decomposition, we showed this was indeed true whenever the original function was continuous along the U portion of the decomposition. This not only extended their result, as the UV decomposition they selected forced this criterion to be true; but, formed the stronger result that when using this broad range of decompositions the Quadratic Sub-Lagrangian was continuous at 0. Although they demonstrated the for their particular choice of UV decomposition the U-Lagrangian we differentiable at 0, they never explored the question of what other subspaces would provide continuity there.

To extend their second result we sought to discover for what UV decompositions the Quadratic Sub-Lagrangian was differentiable at 0. In doing this, we turned to the work of Lemaréchal, Oustry, and Sagastizábal for inspiration. In examining their choice of UV decomposition we showed that it provided a satisfactory decomposition to ensure the Quadratic Sub-Lagrangian was differentiable at 0. More than that we showed it maximized the subspace U in the UV decomposition. By this we mean if the Quadratic Sub-Lagrangian was differentiable at 0, then U had to be a subspace of the one used in Lemaréchal, Oustry, and Sagastizábal's decomposition. One of the interesting points about this result was that it generalized perfectly to the U-Lagrangian as applied to convex functions since the quadratic factor could then be taken to be 0, yielding the U-Lagrangian. Whether or not Lemaréchal, Oustry, and Sagastizábal were aware of this fact is unknown. Another interesting result that came out of this section was that every subdifferentially regular prox-regular function has a subspace along which it is strictly differentiable and this subspace is related to the relative interior of the subgradient mapping.

To generalize the last result of Lemaréchal, Oustry, and Sagastizábal we examined the quadratic expansions of the Quadratic Sub-Lagrangian at 0. We first showed that, like the U-Lagrangian, these expansions could be used to provide quadratic expansions of the original function. This was slightly more general than the U-Lagrangian case, as they focused on the existence of a Hessian matrix. However, Lemaréchal, Oustry, and Sagastizábal's results were stronger in the end as they were able to find an if and only if proof for the existence of such Hessians with regards to f, whereas we were only able to provide an if and only if statements regarding when the Quadratic Sub-Lagrangian was twice epi-differentiable at 0.

In extending these major results we also showed many smaller, but nonetheless important, results for the Quadratic Sub-Lagrangian. One of the more useful ones was that the Quadratic Sub-Lagrangian's related proximal map was nonempty whenever the Quadratic Sub-Lagrangian was finite. In showing this result we also proved two very useful equations. The first showed a norm bound for the proximal mapping of the Quadratic Sub-Lagrangian. This was very useful as it showed that if u converged to 0 and v was in the proximal map of the Quadratic Sub-Lagrangian evaluated at u, then v also converged to 0. This fact was used in various proofs including the proof of prox-regularity of the Quadratic Sub-Lagrangian at 0.

The second equation showed a relation between the Quadratic Sub-Lagrangian of a function and the Quadratic Sub-Lagrangian of the function after linear adjustments. By showing this we provided the ability to break down the proofs into simpler sections. This allowed us to begin most proofs assuming that the function was prox-regular at $\bar{x}$ for 0, then extend the result to prox-regularity for any $\bar{w}$ in the subgradient map of the function. This two step process made proofs easier to follow by removing much of the excessive notation that would be required for direct proofs.

As mentioned before, the other key envelope in the development of the Quadratic Sub-Lagrangian was the Moreau envelope. Unlike the U-Lagrangian, the Moreau envelope was used more as a tool than an inspiration. Other than showing the Quadratic Sub-Lagrangian formed a closer envelope than the Moreau envelope we did not expand on results regarding Moreau envelopes and prox-regular functions. Instead we made use of these results to gain a better understanding of the Quadratic Sub-Lagrangian.

The first fact discussed about Moreau envelopes was a very basic one, namely that Moreau envelopes converge to the original function as the envelope factor gets larger. Since the Quadratic Sub-Lagrangian was squeezed between the two functions it was a simple corollary that it too converged towards the original function.

We did not involve Moreau envelopes again until Chapter four where they were of some use in examining the second order behaviour of the Quadratic Sub-Lagrangian. More specifically we showed that the Quadratic Sub-Lagrangian is twice epi-differentiable if and only if the Moreau envelope of the original function, examined along the subspace U, has this property. As a simple result from this, we learn that if f is twice epi-differentiable then it's Quadratic Sub-Lagrangian is also twice epi-differentiable. In the case of Moreau envelopes this is an if and only if condition so our result is weaker in that respect. However, one could view this as a stronger property since the Quadratic Sub-Lagrangian doesn't require f to be twice epi-differentiable in order for it to have this property. The question of what happens when $f_U = f(\bar{x} + (\cdot_u + 0))$ is twice epi-differentiable has yet to be determined. In fact there are several areas for future research in the field of Quadratic Sub-Lagrangians, as the next section will relate.

5.2 Future Areas of Study

There are two clear directions for future exploration in the area of Quadratic Sub-Lagrangians. The first is to extend the space for which they are defined, the second is to examine the class of functions for which the theorems of this thesis hold.

Since Quadratic Sub-Lagrangians are defined in terms of subgradients, inner products, and perpendicular subspaces the definition can be extended to any space for which these things are well defined. Specifically Quadratic Sub-Lagrangians are well defined in any Hilbert space.

Like Quadratic Sub-Lagrangians the concept of prox-regularity is defined in terms of subgradients and inner products, so it too can be examined in general Hilbert spaces. In S. K. Boralugoda's Ph.D. thesis, "Prox-Regular Functions in Hilbert Spaces", the idea of a prox-regular function was developed for Hilbert spaces [1]. Using the definitions and theorems of this work it may be possible to extend this thesis to Hilbert spaces as well.

Although on the surface most of the theorems in this thesis appear to be adaptable to Hilbert spaces, they are not. The problem occurs in Proposition 2.7 where we showed that for any u in the domain of Φ_R that W_R is nonempty. In proving this we showed that h_R achieved it's infimum inside of a closed ball. Since, in $\Re^n$, closed balls are compact, and h_R was lower semicontinuous we were able to change this infimum to a minimum, and conclude that W_R was nonempty. Unfortunately closed balls are not compact in any infinite dimensional Hilbert space.

This is not the end of the difficulties in extending Quadratic Sub-Lagrangians to Hilbert spaces. Although not cited directly many other theorems applied Proposition 2.7 in their proofs. Any theorem that used equation (2.8) to state that v converged to 0 as u converged to 0, (Theorem 3.2, and its Corollary for example), also used the fact that v existed for any u in the domain of Φ_R. Therefore, if Proposition 2.7 cannot be extended to Hilbert spaces then these theorems would have to be redone in another way. Since Theorem 3.2 states when the Quadratic Sub-Lagrangian is continuous and prox-regular, one of the key results of this thesis, it could not be ignored in extending this envelope to infinite dimensional Hilbert spaces.

Further challenge occurs in that some of the facts used in this thesis are not yet extended to Hilbert spaces. Fact 4.2 for example is not yet extended to Hilbert spaces [1]. Since this fact is used in proofs of Corollary 4.4 and Theorem 4.5 one would either have to show this fact holds true in general Hilbert spaces, or reexamine the proofs in which they were used.

In extending the thesis in the other direction, one could examine how the results of this thesis extend to broader classes of functions. As mentioned before the question of what occurs when $f(\bar{x}+(\cdot_u+0))$ is twice epi-differentiable has yet to be resolved. As a second example, given a function on $f : \Re^n \longrightarrow \bar{\Re}$, a point $\bar{x}$ where f is finite, $\bar{w} \in \partial f(\bar{x})$, and a UV decomposition along which $f_U = f(\bar{x} + (\cdot_u + 0))$ is continuous, is it true that the Quadratic Sub-Lagrangian is continuous? (The answer to this appears to be no, but further exploration would be required to provide a counter example). Another question is, what do we have to add to the above conditions to force Φ_R to be continuous, or prox-regular? Would it suffice to have the function bounded below by a quadratic, or perhaps bounded below by any continuous function?

In answering these questions one should keep in mind that prox-regular functions are dense in the space of measurable functions on $\Re^n$. Thus any function can be approximated within ε by a prox-regular one. So another question is how can this be used? That is, if one takes a sequence of prox-regular functions converging to a given function, then what can be said about the corresponding Quadratic Sub-Lagrangians? And, what additional properties does the function have to have in order for this convergence to be of use? Unlike the problem of extending Quadratic Sub-Lagrangians to Hilbert spaces, the problems that will be encountered in enlarging the function class used are not obvious, and only future study will reveal them.

5.3 Conclusion

The study of Optimization has come a long way since Fermat entered into the field in the 1620s ([9], p. 430). The development of calculus alone provided vast insight into the study of local minimums and maximums. Calculus yielded many powerful results such as the mean value theorem and the derivative test for convexity of functions. Unfortunately calculus was limited to differentiable functions, so the study of Optimization was far from complete. Although most functions do not fall into the category of differentiable, it is well know that every function can be approximated to any degree by a differentiable function. So the major question in Optimization today is how to find these approximations. This thesis provided one method through the development of a new envelope function we called the Quadratic Sub-Lagrangian. It also showed how this envelope forms a more accurate estimate of a given function than the Moreau envelope and behaves well when applied to the broad range of functions known as prox-regular. However development of the Quadratic Sub-Lagrangian is far from complete. The questions of generalizing these results to Hilbert spaces and to broader ranges of functions are still unanswered. Undoubtedly answering these questions will only lead to further questions, and perhaps other envelopes. And so, like all fields of mathematics, Optimization was a short time in creation, but will be a long time in completion.

Bibliography

[1] S. K. Boralugoda, *Prox-Regular Functions in Hilbert Spaces*, Ph.D. Thesis (1998), Univerty of Alberta.

[2] Claude Lemaréchal, François Oustry, and Claudia Sagastizábal, *The U-Lagrangian of a Convex Function*, Tans. Amer. Math. Soc., 352 (2000), pp. 711-729.

[3] J. J. Moreau, *Propriétés des applications "prox"*, C. R. Acad. Sci. Paris, 256 (1963), pp. 1069-1071.

[4] J. J. Moreau, *Proximité et dualité dans un espace hilbertien*, Bull. Soc. Math. France, 93 (1965), pp. 273-299.

[5] R. A. Poliquin and R. T. Rockafellar, *Generalized Hessian Properties of Regularized Nonsmooth Functions*, SIAM J. Optimization, Vol. 6 No. 4 (1996), pp. 1121-1137.

[6] R. A. Poliquin and R. T. Rockafellar, *Prox-Regular Functions in Variational Analysis*, Trans. Amer. Math. Soc., 348 (1996), pp. 1805-1838.

[7] R. T. Rockafellar, *The Multiplier Method of Hestenes and Powell Applied to Convex Programming*, Journal of Optimization Theory and Applications, 12 (1973), pp. 555-562.

[8] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Springer-Verlag, New York, 1998.

[9] Victor J. Katz, *A History of Mathematics, an Introduction*, Harper Collins College Publishers, New York, 1993.