Fine and Pathological Properties of Subdifferentials

by

Shawn Xianfu Wang
B.Sc., Lanzhou University, 1986
M.Sc., Central South University of Technology, 1989
M.Sc., Simon Fraser University, 1995

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
Doctor of Philosophy
in the Department
of
Mathematics & Statistics

© Shawn Xianfu Wang 1999
SIMON FRASER UNIVERSITY
August 1999

All rights reserved. This work may not be
reproduced in whole or in part, by photocopy
or other means, without the permission of the author.
APPROVAL

Name: Shawn Xianfu Wang
Degree: Doctor of Philosophy
Title of thesis: Fine and Pathological Properties of Subdifferentials

Examining Committee: Dr. R. Lockhart
Chair

Dr. J.M. Borwein, Senior Supervisor

Dr. B.S. Thomson

Dr. R. Choksi

Dr. P. Borwein

Dr. A.S. Lewis, External Examiner
Dept. of Combinatorics & Optimization
The University of Waterloo, Waterloo, Ontario

Date Approved: ________________________________
Abstract

In this thesis, we explore a Baire categorical approach for studying the subdifferentiability of Lipschitz functions and continuous functions.

On the line, we show that nowhere monotone functions are the key ingredients to construction of continuous functions, absolutely continuous functions, and Lipschitz functions with large subdifferentials. We also explicitly construct Lipschitz functions which are simultaneously Hölder subdifferentiable and superdifferentiable only on a countable dense subset.

On separable Banach spaces, we prove that the class of functions with maximal Clarke and approximate subdifferentials is residual in an appropriate complete metric space. In particular, in the space of nonexpansive functions endowed with the supremum metric, the set of functions whose Clarke subdifferential and approximate subdifferential are identically equal to the dual unit ball is residual; in the space of bounded continuous functions endowed with the supremum metric, the set of functions whose Clarke subdifferential and approximate subdifferential are identically equal to the whole dual space is residual. On general Banach spaces, we show that for a locally Lipschitz function $f$ there exists a residual subset $G$, in an appropriate complete metric space, such that $(-\partial_x(-f)) \cap \partial_x g \neq \emptyset$ everywhere for each $g \in G$ where $\partial_x$ stands for the approximate or Clarke subdifferential. Diverse implications are given. In order to recover a locally Lipschitz function from its Clarke subdifferential, we also study line integrals in general Banach spaces.

Such a categorical approach allows us to generalize and simplify many known results on subdifferential integrability, and to illustrate the stark difference between the subdifferentiability of convex functions and the subdifferentiability of nonconvex functions.
Dedication

Dedicated to my wife Shirely, son Kevin, and daughter Amy.
Acknowledgements

It is my great honor to study under the first class mathematician: Dr. J.M. Borwein, whose profound mathematical expertise has taught me a great deal. I sincerely thank him for giving me the opportunity to study under him in CECM. Without him this thesis could not have been written.

I am grateful to the members of my Examining Committee and the External Examiner for taking the time to read this thesis and suggestions on its improvement. During the writing of the thesis, I also benefited from communication and discussions with: Dr. W.B. Moors, Dr. J. Vanderwerff.

My thanks also go to the SFU Mathematics Department for teaching assistantships and fellowships, and CECM for making such a stimulating study environment.
# Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approval</td>
<td>ii</td>
</tr>
<tr>
<td>Abstract</td>
<td>iii</td>
</tr>
<tr>
<td>Dedication</td>
<td>iv</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>v</td>
</tr>
<tr>
<td>List of Figures</td>
<td>ix</td>
</tr>
<tr>
<td>Overview</td>
<td>1</td>
</tr>
<tr>
<td>1 Preliminaries</td>
<td>4</td>
</tr>
<tr>
<td>1.1 Classical differentiability</td>
<td>5</td>
</tr>
<tr>
<td>1.2 Subderivatives and subdifferentials</td>
<td>8</td>
</tr>
<tr>
<td>1.3 Subdifferential integrability</td>
<td>14</td>
</tr>
<tr>
<td>1.4 Minimal subdifferentials and partially strict differentiability</td>
<td>15</td>
</tr>
<tr>
<td>2 The tool-box</td>
<td>17</td>
</tr>
<tr>
<td>2.1 Baire category and Zorn’s lemma</td>
<td>17</td>
</tr>
<tr>
<td>2.2 The strong separation theorem</td>
<td>18</td>
</tr>
<tr>
<td>2.3 Mean-value theorems</td>
<td>19</td>
</tr>
<tr>
<td>2.4 Generating uscos and cuscos</td>
<td>20</td>
</tr>
<tr>
<td>2.5 Haar null sets</td>
<td>22</td>
</tr>
<tr>
<td>2.6 Lipschitz functions with minimal Clarke subdifferentials</td>
<td>23</td>
</tr>
</tbody>
</table>
2.7 Variational principles ........................................... 27
2.8 Some properties of dual balls of Banach spaces ............... 29
3 Subdifferentials in $\mathcal{H}$ ........................................ 31
  3.1 Nowhere monotone functions .................................. 31
  3.2 Properties of nowhere monotone functions .................... 33
  3.3 Rockafellar type functions ................................... 34
  3.4 The Michel-Penot subdifferential on null sets ............... 38
  3.5 The space of nondecreasing continuous functions .......... 40
  3.6 The space of automorphisms .................................. 42
  3.7 The space of continuous functions $C[0,1]$ .................... 44
  3.8 Such a pathological behavior is prevalent too! ............. 51
  3.9 Typical Lipschitz functions have constant subdifferentials .. 53
  3.10 Can the pseudo-regular points generate the subdifferential? 58
  3.11 A comparison to convex analysis .............................. 59
4 Prescribed Hölder subdifferentials ................................ 61
  4.1 Domains of Hölder super-sub-differentials may be countable .. 61
  4.2 Basic construction ............................................. 62
  4.3 Computing the Hölder super-sub-differentials of $f$ .......... 72
5 Subdifferentials in separable Banach spaces ....................... 86
  5.1 Properties of $T$-Lipschitz functions ......................... 87
  5.2 When is a set-valued map a Clarke subdifferential? .......... 89
  5.3 A characterization of the Clarke subdifferential ............. 97
  5.4 The approximate subdifferential ............................... 100
  5.5 The Gâteaux differentiability of $C$-Lipschitz functions .... 102
  5.6 Further properties of approximate subdifferentials .......... 105
5.7 The subdifferentiability of permutation-invariant functions . . . . . . 108
5.8 The subdifferentiability properties of $BC(A)$ . . . . . . . . . . . . . 110
5.9 The subdifferentiability properties of $C(A)$ . . . . . . . . . . . . . . 115
5.10 The viscosity subdifferentiability of $BC(A)$ . . . . . . . . . . . . . . 118
5.11 The subdifferentiability of $C$-Lipschitz extensions . . . . . . . . . 121
5.12 Sum rules and Jacobians . . . . . . . . . . . . . . . . . . . . . . . . . . 128

6 Subdifferentials in arbitrary Banach spaces . . . . . . . . . . . . . . . 135
6.1 A dense approximation lemma . . . . . . . . . . . . . . . . . . . . . . 135
6.2 Approximate subdifferentials in smooth Banach spaces . . . . . . . . 136
6.3 Clarke subdifferentials in arbitrary Banach spaces . . . . . . . . . . . 139
6.4 $\sigma$-minimal Clarke subdifferentials plus the dual ball . . . . . . . 144
6.5 When is $\mathcal{A}_T$ non-empty? . . . . . . . . . . . . . . . . . . . . . . 145
6.6 The existence of $s$-minimal subdifferentials . . . . . . . . . . . . . . 148
6.7 The Mordukhovich-Shao sequential subdifferential . . . . . . . . . . 150

7 Green’s theorem in Banach spaces . . . . . . . . . . . . . . . . . . . . . 154
7.1 Line integrals . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 154
7.2 Integration of subdifferentials of essentially smooth functions . . . . 157
7.3 Green’s theorem and almost everywhere differentiability . . . . . . . 157
7.4 Another characterization of the Clarke subdifferential . . . . . . . . . 162
7.5 Examples . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 163

Bibliography . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 166
Index . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 174
List of Figures

4.1 The function of $g$ on $[a_{n,3k}, a_{n,3k+3}]$ 65
4.2 The upper and lower bounds of function $g$ on $[a_{n,3k}, a_{n,3k+3}]$ 67
Overview

Nonsmooth analysis refers to differential analysis in the absence of differentiability. In order to generalize the notion of classical differentiability, nonsmooth analysts have introduced “derivative objects”: subdifferentials. The main competitors in recent years are: the Clarke subdifferential, Ioffe-Mordukhovich-Kruger approximate subdifferential, Michel-Penot subdifferential, and Dini-Hadamard subdifferential. Just as derivatives can be used to characterize many properties of functions: such as convexity, monotonicity, necessary optimization conditions, implicit function theorems, etc, the subdifferentials have played the same role, or an even more powerful role than the derivative both in theory and in practice. Surfing through classical analysis and comparing to nonsmooth analysis, we observe a gap. In classical analysis, typical functions have no derivative at all, whereas in nonsmooth analysis it is not clear at all what are the typical subdifferentiability properties of Lipschitz functions or continuous functions.

In this thesis, we study the possible pathological behavior of subdifferentials of continuous functions and Lipschitz functions. We shall see that most continuous functions and Lipschitz functions exhibit very pathological behavior. Once we understand those properties we are in a position to view the pathological properties as inevitable or natural consequences of nonsmooth analysis.

In chapter 1, we gather basic concepts in nonsmooth analysis. The Clarke subdifferential, Ioffe-Mordukhovich-Kruger approximate subdifferential, Dini subdifferential, and Michel-Penot subdifferential are introduced. The relationship between subdifferentiability and classical differentiability is discussed.

In chapter 2, we list the main tools used throughout the thesis: Baire category, Haar measure, mean-value theorems, and variational principles. Generating uscos and cuscos
from densely defined multifunctions will allows us to produce many useful generalized sub-differential maps in later chapters. Lipschitz functions with minimal Clarke subdifferentials are included for comparison purpose.

In chapter 3, we show that nowhere monotone functions are the key ingredients to construction of continuous functions, absolutely continuous functions, and Lipschitz functions with large subdifferentials on the real line. Let \( \partial_c f, \partial_a f \) denote the Clarke subdifferential and approximate subdifferential respectively. In the Banach space of continuous functions defined on \([0, 1]\), denoted by \( C[0, 1] \), with the uniform norm, we show that there exists a residual and prevalent set \( D \subset C[0, 1] \) such that \( \partial_a f = \partial_c f = \mathbb{R} \) on \([0, 1]\) for every \( f \in D \). In the space of automorphisms we prove that most functions \( f \) satisfy \( \partial_a f = \partial_c f = [0, +\infty) \) on \([0, 1]\). The subdifferentiability of the Weierstrass function and the Cantor function are completely analyzed. We also construct absolutely continuous functions on \( \mathbb{R} \) such that \( \partial_a f = \partial_c f \equiv \mathbb{R} \).

In chapter 4, we construct Lipschitz functions on the line with prescribed Hölder subdifferentials. More precisely, suppose that \( S_1 \) and \( S_2 \) are countable and dense in \( \mathbb{R} \) with \( S_1 \cap S_2 = \emptyset \). Then there exist two countable sets \( D_1 \subset S_2 \) and \( D_2 \subset S_1 \) with both \( D_1 \) and \( D_2 \) dense in \( \mathbb{R} \) such that there exists a Lipschitz function \( f : \mathbb{R} \to \mathbb{R} \) whose Hölder subdifferential map \( \partial_{hs} \) satisfies:

\[
\partial_{hs} f(x) = (-1, 1) \quad \text{if } x \in D_2 \quad \text{and} \quad \partial_{hs} f(x) = \emptyset \quad \text{if } x \notin D_2, \\
\partial_{hs} (-f)(x) = (-1, 1) \quad \text{if } x \in D_1 \quad \text{and} \quad \partial_{hs} (-f)(x) = \emptyset \quad \text{if } x \notin D_1,
\]

for all \( s > 0 \). This shows that the “Proximal density theorem” of Clarke, Ledyaev and Wolenski’s can not be significantly improved.

In chapter 5, we prove the existence of continuous functions and Lipschitz functions with maximal Clarke subdifferentials on separable Banach spaces. Examples of such functions have been given by Rockafellar, Borwein and Fitzpatrick, Jouini etc. The nature of these functions as contrasted with familiar nonsmooth functions tends to suggest that such functions are the exception rather than the rule. But we show that the actual situation is quite the reverse. We prove that the class of functions with maximal subdifferential is residual in an appropriate complete metric space, \( \mathcal{X}_T \), so that we nonsmooth analysts should be surprised in some sense whenever we encounter a function such that its Clarke subdifferential is
singleton at some point. The same hold for approximate subdifferentials. Jouini’s, Borwein and Fitzpatrick’s, Borwein, Moors and Wang’s results are generalized simultaneously.

In chapter 6, we extend the results in chapter 5 to general Banach spaces. The main result is that for a locally Lipschitz function $f \in \mathcal{X}_T$, one may find a residual subset $G \subseteq \mathcal{X}_T$ such that $(-\partial_h(-f)) \cap \partial_h g \neq \emptyset$ everywhere for each $g \in G$ where $\|\|$ stands for ‘a’ or ‘c’. This implies each weak* cusco generated by a countable number of minimal Clarke subdifferentials is a Clarke subdifferential. We examines the existence of Lipschitz functions with ‘s-minimal’ subdifferential mappings which provides a stark contrast to the results concerning the existence of ‘maximal’ subdifferentials. Finally, we show that on weakly compactly generated Asplund spaces with Fréchet differentiable norms most nonexpansive Lipschitz functions identically have the dual unit ball as their Mordukhovich-Shao sequential subdifferentials.

In chapter 7, by introducing line integrals for Borel measurable maps, we obtain an infinite dimensional analogue of Green’s Theorem. In separable Banach space, we show that the integral is Gâteaux differentiable everywhere except for a Haar null set, and its Gâteaux derivative equals that Borel measurable map except for a Haar null set. Examples are given to show that this fails for Fréchet differentiability. Borwein, Moors and Shao’s recent related work is extended.

An index appears at the end of the thesis.
Chapter 1

Preliminaries

In this chapter we gather basic concepts that will be used repeatedly throughout the rest of the thesis.

For a Banach space \((X, \|\cdot\|)\), we denote by \((X^*, w^*)\) its norm dual space endowed with the weak* topology, which is a locally convex and Hausdorff topological vector space; \(\langle x^*, x \rangle\) is the canonical pairing on \(X^* \times X\);

\[
[a, b] := \{tb + (1 - t)a : 0 \leq t \leq 1\}; \quad S_X := \{x \in X : \|x\| = 1\};
\]

\[
B_X := \{x \in X : \|x\| \leq 1\}; \quad B_{X^*} := \{x^* \in X^* : \|x^*\| \leq 1\};
\]

\[
B_\delta(x) := \{y \in X : \|x - y\| < \delta\}; \quad B_\delta[x] := \{y \in X : \|x - y\| \leq \delta\}.
\]

For a non-empty subset \(E\) of \(X^*\) we denote by \(w^*\text{cl}E\) the \(w^*\)-closure of \(E\); \(\overline{\text{co}}^{w^*} E\) the \(w^*\)-closed convex hull of \(E\). For \(x \in X\) we define

\[
\sigma_E(x) := \sup\{\langle x^*, x \rangle : x^* \in E\}.
\]

If \(\{Q_\alpha\}\) is a set of sets from \(X^*\), we use \(\limsup Q_\alpha\) to denote the collection of limits of \(w^*\)-converging subnets \(\{x_\alpha\}\) with \(x_\alpha \in Q_\alpha\) for every \(\alpha\). On \(\mathbb{R}^n\), we use ‘\(\mu\)’ to denote the Lebesgue measure.

For a function \(f : X \to \mathbb{R}\) and \(S \subset X\) we set

\[
\text{epi} f := \{(\alpha, x) \in \mathbb{R} \times X : \alpha \geq f(x)\}; \quad \text{dom} f := \{x \in X : |f(x)| < \infty\};
\]
1.1 Classical differentiability

\[ I_S(x) := \begin{cases} 
0 & \text{if } x \in S, \\
+\infty & \text{if } x \not\in S.
\end{cases} \]

\[ \chi_S(x) := \begin{cases} 
1 & \text{if } x \in S, \\
0 & \text{if } x \not\in S.
\end{cases} \]

\[ d_S(x) := \inf\{\|x - y\| : y \in S\}; \quad f_S(x) := \begin{cases} 
f(x) & \text{if } x \in S, \\
+\infty & \text{if } x \not\in S.
\end{cases} \]

By \( u \to f x \), we mean \( u \to x \) and \( f(u) \to f(x) \). We say that \( f \) is proper lower semicontinuous on \( X \) if \( f \) is lower semicontinuous, not everywhere equal to \(+\infty\), and nowhere \(-\infty\); and \( f \) is Lipschitz on \( A \) if there exists an \( L > 0 \) such that \(|f(z) - f(y)| \leq L\|z - y\|\) for all \( z, y \in A \). If for some \( \delta > 0 \), \( f \) is Lipschitz on \( B_{\delta}(x) \), we say that \( f \) is locally Lipschitz around \( x \).

1.1 Classical differentiability

A \( \beta \)-bornology of \( X \) is a family of closed bounded and centrally symmetric subsets of \( X \) whose union is \( X \), which is closed under multiplication by scalars and is directed upwards (that is, the union of any two members of \( \beta \) is contained in some member of \( \beta \)). We will denote by \( X_{\beta}^* \) the dual space of \( X \) endowed with the topology of uniform convergence on \( \beta \)-sets. The following choices for \( \beta \) are of main interest: (i) \( G \) the Gâteaux bornology consists of all finite symmetric sets. (ii) \( H \) the Hadamard bornology, consisting of all compact symmetric sets. (iii) \( F \) the Fréchet bornology, consisting of all bounded symmetric sets.

**Definition 1.1.1** A real-valued function \( f \) is said: (i) to have directional derivative at \( x \) in the direction \( y \) if

\[ f'(x; y) := \lim_{t \to 0^+} \frac{f(x + ty) - f(x)}{t} \]

exists;

(ii) to be \( \beta \)-differentiable at \( x \) and \( x^* \in X_{\beta}^* \) is called its \( \beta \)-derivative at \( x \), if for each \( S \in \beta \),

\[ f(x + ty) - f(x) - \langle x^*, y \rangle \to 0, \]

as \( t \to 0^+ \) uniformly for \( y \in S \). We denote the \( \beta \)-derivative of \( f \) at \( x \) by \( \nabla_\beta f(x) \). We say that \( f \) is \( \beta \)-smooth at \( x \) if \( \nabla_\beta f : X \to X_{\beta}^* \) is continuous in a neighbourhood of \( x \).

We obtain the Fréchet derivative by applying this definition to the bounded bornology, the Hadamard derivative in the case of the norm-compact bornology and the Gâteaux derivative for the finite bornology. When \( f \) is Lipschitz near \( x \) then Hadamard and Gâteaux differentiability coincide. Hadamard differentiability has the composition property but this fails.
for Gâteaux differentiability [94]. In finite dimensions Hadamard and Fréchet differentiability are equivalent, and for Lipschitz functions Gâteaux and Fréchet differentiability are the same. The following example [11] illustrates the distinctions.

Example 1.1.2 Consider the function

\[ f(x, y) := \begin{cases} 
\frac{x^ay^b}{(x^p + y^q)} & \text{if } (x, y) \neq 0, \\
0 & \text{if } (x, y) = (0, 0).
\end{cases} \]

(1). If \( a = 2, b = 3, p = 2, q = 4 \) then \( f \) is Fréchet differentiable at \((0, 0)\).

(2). If \( a = 1, b = 3, p = 2, q = 4 \) then \( f \) is Gâteaux differentiable at \((0, 0)\) but not Fréchet differentiable.

(3). If \( a = 2, b = 4, p = 4, q = 8 \) then \( f \) is Gâteaux differentiable but not continuous at \((0, 0)\).

(4). If \( a = 1, b = 2, p = 2, q = 2 \) then \( f \) is not Gâteaux differentiable but has directional derivative in every direction and is locally Lipschitz around \((0, 0)\).

(5). If \( a = 1, b = 4, p = 2, q = 4 \) then \( f \) is Fréchet differentiable at \((0, 0)\) but not continuously differentiable at \((0, 0)\) since \( \nabla f(0, 0) = (0, 0) \) and \( \nabla f(0, y) = (1, 0) \) for \( y \neq 0 \).

(6). If \( a = 1, b = 1, p = 2, q = 2 \) then \( f \) does not have a derivative in directions other than those of the coordinate axes. Moreover \( f \) is not continuous at \((0, 0)\).

The most important differentiability theorem in \( \mathbb{R}^n \) is:

Theorem 1.1.3 (Rademacher) Let \( U \) be a nonempty open subset of \( \mathbb{R}^n \) and suppose that \( f : U \to \mathbb{R} \) is locally Lipschitz. Then \( f \) is differentiable almost everywhere.

If we assume one of the partial \( \partial f(x)/\partial x_1, \ldots, \partial f(x)/\partial x_n \) exists at \( x \) and the remaining \( n - 1 \) partials exists in some neighborhood and are continuous at \( x \), then \( f \) is differentiable at \( x \) [1]. A nice property of convex functions [52, 49] comes as follows:

Theorem 1.1.4 Let \( f \) be a convex function on \( \mathbb{R}^n \), and \( |f(x)| < \infty \). Then \( f \) is differentiable at \( x \) if and only if for some basis of \( \mathbb{R}^n \) in which \( x = (x_1, \ldots, x_n) \), the partial derivatives \( \partial f(x)/\partial x_i \) exist at \( x \) and are finite for \( i = 1, \ldots, n \).
1.2 Classical differentiability

The following example [72] shows that in general Banach space, some hypotheses are required for the validity of a differentiability theorem for Lipschitz functions.

**Example 1.1.5** (1) For \( x = (x_n) \) in \( l^\infty \) define a seminorm \( p \) by \( p(x) := \limsup |x_n| \), then \( p \) is continuous and convex, but nowhere Gâteaux differentiable.

(2) The norm \( \|x\|_1 := \sum |x_n| \) in \( l^1 \), the space of all real sequences for which \( \sum |x_n| \) is convergent, is Gâteaux differentiable precisely at those point \( x = (x_n) \) for which \( x_n \neq 0 \) for all \( n \). In this case, the Gâteaux differential is the bounded sequence \( \text{sgn} x_n \in l^\infty \). On the other hand \( \| \cdot \|_1 \) is not Fréchet differentiable at any point.

A norm in a Banach space \( X \) is said to be \( \beta \)-**smooth** if it is \( \beta \)-differentiable at each \( x \neq 0 \). We call a Banach space \( \beta \)-**smoothable** if it has an equivalent \( \beta \)-smooth norm. A Banach space \( X \) is said to be **smoothable** if it admits an equivalent norm that is Gâteaux differentiable everywhere except at 0. A celebrated differentiability result [74] is the following:

**Theorem 1.1.6 (Preiss)** Let \( X \) be a \( \beta \)-smoothable Banach space. Then every locally Lipschitz function defined on an open subset \( A \) of \( X \) is \( \beta \)-differentiable on a dense subset of \( A \). Moreover, the mean-valued theorem holds for the \( \beta \)-derivative of locally Lipschitz functions.

**Strict differentiability**

Let \( f \) be Lipschitz near \( x \). We say that \( f \) is **strictly Gâteaux differentiable** at \( x \) if for each \( v \in X \) one has

\[
\lim_{t \to 0} \frac{f(x' + tv) - f(x')}{t} = \langle \nabla f(x), v \rangle.
\]

If the limit is uniform for \( v \in B_X \), we say \( f \) is **strictly Fréchet differentiable** at \( x \).

On separable Banach spaces for a locally Lipschitz function defined on an open set, the set of points where it is Gâteaux differentiable but not strictly Gâteaux differentiable is of first category [50]. For every real-valued function \( f \) defined on a Banach space the set of points at which \( f \) is Fréchet differentiable but is not strictly Fréchet differentiable is of the first category [97]. If \( f \) is continuous and convex on nonempty open convex subset \( A \) of \( X \), then \( f \) is Fréchet (Gâteaux) differentiable at a point \( x \in A \) if and only if \( f \) is strictly Fréchet (Gâteaux) differentiable at \( x \) [72].
1.2 Subderivatives and subdifferentials

A set-valued mapping $T$ between Hausdorff topological spaces $X$ and $Y$ is an *usco* if $T(x)$ is a non-empty compact subset of $Y$ for each $x \in X$, and if $T$ is *upper semicontinuous*: that is whenever $U$ is open in $Y$ then $\{x \in X : T(x) \subset U\}$ is open in $X$. If, in addition, $Y$ is a topological vector space and $T(x)$ is convex, then $T$ is a *cusco*. Moreover, $T$ is called a *minimal cusco (usco)* on $X$ if its graph does not properly contain the graph of any other cusco (usco) on $X$. We say that $T$ is *continuous* at $x_0 \in X$ if $T(x_0)$ is a singleton and, for every open set $U$ containing $T(x_0)$, the set $\{x \in X : T(x) \subset U\}$ is a neighbourhood of $x_0$. We denote the graph of $T$ by

$$\text{Gr}(T) := \{(x, x^*) \in X \times Y : x^* \in T(x)\}.$$ 

Every usco $T : X \to 2^Y$ has a closed graph in $X \times Y$ [72]. If $Y = Z^*$ is a dual Banach space, saying that a usco (cusco) map $T$ from $X$ into $2^Y$ is *w*-usco (cusco) means that we are taking $Y$ to be $Z^*$ in its w*-topology.

For a lower semicontinuous function $f$ on a Banach space $X$ with $|f(x)| < \infty$, we define the following subdifferentials:

**The $\beta$-subdifferential**

The $\beta$-subdifferential of $f$ at $x$ is:

$$\partial_\beta f(x) := \{x^* \in X^* : \liminf_{t \to 0} \inf_{h \in V} t^{-1}(f(x + th) - f(x) - t\langle x^*, h \rangle) \geq 0 \text{ for all } V \in \beta\}.$$ 

If $f$ is $\beta$-differentiable at $x$, then $\partial_\beta f(x)$ is singleton. In particular, the Frechet subdifferential of $f$ at $x$ is:

$$\partial^F f(x) = \{x^* \in X^* : \liminf_{y \to x} \frac{f(y) - f(x) - \langle x^*, y - x \rangle}{\|y - x\|} \geq 0\};$$

while the Dini-Hadamard subdifferential and Dini-Hadamard superdifferential corresponding to the compact bornology transpires to be equal to

$$\partial_- f(x) = \{x^* : \langle x^*, v \rangle \leq f^-(x; v) \text{ for all } v \in X\},$$

$$\partial^+ f(x) = \{x^* : \langle x^*, v \rangle \leq f^+(x; v) \text{ for all } v \in X\},$$
1.2 Subderivatives and subdifferentials

where the Dini-Hadamard type lower and upper derivatives of $f$ at $x$ in the direction $v$ are given respectively by

$$f^+(x;v) := \limsup_{t \to 0^+} \frac{f(x + th) - f(x)}{t}, \quad f^-(x;v) := \liminf_{t \to 0^-} \frac{f(x + th) - f(x)}{t}.$$  

When $f$ is locally Lipschitz at $x$, one may set $v \equiv h$ in the limits.

On $\beta$-smoothable Banach spaces, each lower semicontinuous function $f$ is densely $\beta$-subdifferentiable on its domain [5]. If $f$ is defined on the real line, the set $A := \{x : \partial_- f(x) \text{ is not singleton}\}$ is at most countable. For $f : X \to \mathbb{R}$ with $X$ being a separable normed vector space, the set $A$ of points where $f$ has a non-unique Fréchet subgradient is first category in $X$. Moreover, if $X = \mathbb{R}^n$, then $A$ is further of measure zero [48].

The $\beta$-viscosity subdifferential

We say $f$ is $\beta$-viscosity subdifferentiable and $x^*$ is a $\beta$-viscosity subderivative of $f$ at $x$ if there exists a locally Lipschitz function $g$ such that $g$ is $\beta$-smooth at $x$, $\nabla \beta g(x) = x^*$, and $f - g$ attains a local minimum at $x$. We denote the set of all $\beta$-viscosity subderivatives of $f$ at $x$ by $D_\beta f(x)$.

A special type of viscosity subdifferential is the $s$-Hölder subdifferential: $x^* \in X^*$ is called an $s$-Hölder subgradient of $f$ at $x$ if $f(x)$ is finite and for some $\sigma > 0$ and $\delta > 0$ one has

$$f(y) \geq f(x) + \langle x^*, y - x \rangle - \sigma \| y - x \|^{1+s},$$

when $\|y - x\| < \delta$. We write $x^* \in \partial_{hs} f(x)$. The $s$-Hölder superdifferential of $f$ is defined as $-\partial_{hs}(-f)$. When $s = 1$ such a subdifferential is called proximal subdifferential, denoted by $\partial_p f$.

In a finite dimensional space, if $f$ is lower semicontinuous at $x$ then $\partial_- f = \partial^F f = D_F f$ [2]. In Banach space which admits a Lipschitz and Fréchet differentiable bump, $\partial^F f = D_F f$ [37]. If $f$ is Gâteaux differentiable at $x$, we have $\partial_- f(x) \subset \{\nabla f(x)\}$. In $\mathbb{R}^2$ it may happen that $f$ is Gâteaux differentiable at $x$ but $\partial_- f(x) = \emptyset$ as the following example shows,

**Example 1.2.1** (1). Let $f : \mathbb{R}^2 \to \mathbb{R}$ be Gâteaux but not Fréchet differentiable at $x$ (see Example 1.1.2). Then either $D_F f(x) = \emptyset$ or $D_F(-f)(x) = \emptyset$. In $\mathbb{R}^n$, this happens only when $f$ is not locally Lipschitz.
(2). In $l^1$ there exists an equivalent norm $\| \cdot \|$ which is Gâteaux differentiable but nowhere Fréchet differentiable [72, page 86]. Therefore $\nabla(\|x\|)$ exists at nonzero point $x$ but $\partial F(\|x\|) = \emptyset$ and $\partial F(\|x\|) = \nabla\|x\|$.

**The approximate subdifferential**

For $x \in \text{dom } f$ we set $\partial^- f(x) := \{ x^* \in X^* : \langle x^*, h \rangle \leq f^-(x; h) + \epsilon \|h\|, \text{ for all } h \in X \}$. Let $\mathcal{F}$ denote the collection of all finite dimensional subspaces of $X$. The set

$$\partial_a f(x) := \bigcap_{L \in \mathcal{F}} \limsup_{u \to x} \partial^- f_{u+L}(u),$$

(1.1)

is called the Ioffe-approximate subdifferential of $f$ at $x$. It follows from the definition that $\partial_a f(x)$ is $w^*$-closed and $\partial_a f(x) = \limsup_{u \to x} \partial_a f(u)$. When $X$ is finite dimensional, it reduces to the Ioffe-Kruger-Mordukhovich approximate subdifferential defined by:

$$\partial_a f(x) := \bigcap_{\delta > 0} \text{cl}\{ z^* : z^* \in \partial_- f(z), \|z - x\| < \delta, |f(z) - f(x)| < \delta \}. \quad (1.2)$$

In $\mathbb{R}^n$, $\partial_a f$ is countably generated [34], that is, there exists a countable set $\Omega \subset \text{Gr}[\partial_- f]$ such that

$$\partial_a f(x) = \{ \lim_{i \to \infty} \xi_i : (x_i, \xi_i) \in \Omega \text{ for all } i, \lim_{i \to \infty} x_i = x \}.$$  

This follows from the fact that $\text{Gr}[\partial a f] = \text{cl}(\text{Gr}[\partial_- f])$ and $\mathbb{R}^n \times \mathbb{R}^n$ is separable. Note that $\partial_a f$ is sensitive to null sets, even a finite number of points.

If in addition $X$ is smoothable, then $\partial_a f(x) = \limsup_{u \to x} \partial^- f_{u}(u)$ for any lower semi-continuous function $f$ on $X$ [57]. When $f$ is locally Lipschitz at $x$, by [4] one may write

$$\partial_a f(x) = w^*\text{cl}\{ w^* - \lim_{n \to \infty} x^*_n : x^*_n \in \partial_- f(x_n), x_n \to x \}.$$  

When $X$ is separable, $\partial_a f(x) = \{ w^* - \lim_{n \to \infty} x^*_n : x^*_n \in \partial_- f(x_n), x_n \to x \}$.

**The Clarke subdifferential**

The Rockafellar directional derivative at $x$ in the direction $v$ and Clarke-Rockafellar subdifferential at $x$ are given respectively by:

$$f^+(x; v) := \lim_{\epsilon \downarrow 0} \inf_{y \in y + B_x} \limsup_{\epsilon \downarrow 0} \frac{f(y + tw) - f(y)}{t}.$$  

(1.3)

$$\partial_c f(x) := \{ x^* \in X^* : \langle x^*, v \rangle \leq f^+(x, v) \text{ for all } v \in X \}.$$
1.2 Subderivatives and subdifferentials

Note that $\partial_c f(x) \subset X^*$ is a $w^*$-closed convex set (possibly empty).

We say that $f$ is directionally Lipschitz at $x$ if $f$ is finite at $x$ and for some $v \in X$:

$$\limsup_{y \to x, \frac{t}{t_0} \to v} \frac{f(y + tw) - f(y)}{t} < +\infty.$$ 

If $f$ is directionally Lipschitz at $x$, then $\partial_a f(x) \subset \partial_c f(x)$ [57]. If $f$ is not directionally Lipschitz at $x$, there is no definite relationship between the two. If $f$ is directionally Lipschitz at $x$, then $\partial_c f(x) = -\partial_c (-f)(x)$ and

$$f^\dagger(x; v) = \limsup_{y \to x, \frac{t}{t_0} \to v} \frac{f(y + tw) - f(y)}{t} \text{ for each } v \in \text{int}\{v : f^\dagger(x; v) < \infty\}.$$ 

When $f$ is locally Lipschitz around $x$, $f^\dagger(x; v)$ and $\partial_c f$ reduce to the Clarke directional derivative and Clarke subdifferential $\partial_c f$ given respectively by

$$f^0(x; v) := \limsup_{y \to x, \frac{t}{t_0} \to v} \frac{f(y + tw) - f(y)}{t}, \quad \partial_c f(x) := \{x^* : \langle x^*, v \rangle \leq f^0(x; v) \text{ for all } v \in X\}.$$ 

Then $\partial_c f(x)$ is a singleton if and only if $f$ is strictly Gâteaux differentiable [33]. In this case $\partial_c f$ is norm-to-weak* continuous. If $f$ is strictly Fréchet differentiable at $x$, then $\partial_c f$ is norm-to-norm continuous at $x$. On $\mathbb{R}^n$ when $f$ is locally Lipschitz,

$$\partial_c f(x) = \text{co}\{\lim \nabla f(x_i) : x_i \to x, x_i \notin \Omega\},$$

where $\Omega$ is any null set containing the set of points at which $f$ fails to be differentiable. Thus $\partial_c f$ is minimal among $w^*$-cuscors containing the derivative of $f$ and the definition of $\partial_a f$ is a natural generalization of $\partial_c f$.

On $\mathbb{R}^n$, for any l.s.c. function $f$ and any $z$ such that $|f(z)| < +\infty$ we have $\partial_a f(z) \subset \partial_c f(z)$, and $\partial_c f(z) \neq \emptyset \Rightarrow \partial_a f(z) \neq \emptyset$. On a Banach space, when $f$ is locally Lipschitz at $x$, $\partial_c f(x) = \overline{co}^w \partial_a f(x)$ [56].

The Michel-Penot subdifferential

The Michel-Penot directional derivative at $x$ in the direction $v$ and subdifferential at $x$ are given respectively by:

$$f^\circ(x; v) := \sup_{w} \limsup_{t \to 0} \frac{f(x + tw + tv) - f(x + tw)}{t}, \quad (1.4)$$
\[ \partial_{mp} f(x) := \{ x^* : \langle x^*, v \rangle \leq f^*(x; v) \text{ for all } v \in X \}. \]

In general, the Michel-Penot subdifferential is smaller than the Clarke subdifferential, but it has no general continuity property corresponding to the \(w^*\)-upper semicontinuity of Clarke’s subdifferential. Unlike \( \partial_c f \), the Michel-Penot subdifferential \( \partial_{mp} f(x) \) is singleton if and only if \( f \) is Gâteaux differentiable at \( x \). For each locally Lipschitz function \( f \) defined on an open set \( A \) of a separable Banach space \( X \), the set [50]

\[ \{ x \in A : \partial^+ f(x) = \partial_{mp} f(x) = \partial_c f(x) \} \]

is residual in \( A \).

### Convex functions

If \( f \) is convex and \( |f(x)| < \infty \), the subdifferential of \( f \) at \( x \) is:

\[ \partial f(x) := \{ x^* : \langle x^*, y - x \rangle \leq f(y) - f(x) \text{ for all } y \in \text{dom } f \}. \tag{1.5} \]

A convex function \( f \) on an open convex subset \( A \) of a Banach space \( X \) is Gâteaux differentiable at \( x \in A \) if and only if \( f \) has a unique subgradient [49, 72]. For a continuous and convex function \( f \) on a nonempty open subset \( A \), the mapping \( \partial f \) is \( \beta \)-continuous at \( x \in A \) if and only if \( f \) is \( \beta \)-differentiable at \( x \) [76]. If \( f \) is continuous convex function on an open set \( A \), then \( f \) is locally Lipschitz [33], and all generalized subdifferentials become \( \partial f \) [72, 33, 58]. Furthermore,

**Theorem 1.2.2** [72] If \( f \) is a continuous convex function on an open subset \( A \) of a Banach space \( X \), then \( \partial f \) is a minimal \( w^* \)-cusc on \( A \).

**Example 1.2.3** A lower semicontinuous convex function \( f \) such that \( \partial f(x) \) is unbounded or empty for some \( x \in \text{dom } f \). Define

\[ f(x_1, x_2) = \begin{cases} \max\{1 - \sqrt{x_1}, |x_2|\} & \text{if } x_1 \geq 0, \\ +\infty & \text{otherwise}. \end{cases} \]

Directly using equation (1.5), we obtain: \((-\infty, 0) \times \{1\} \subset \partial f(0, 1); \partial f(0, x_2) = \emptyset \text{ if } |x_2| < 1; \) and \((-\infty, 0) \times \{-1\} \subset \partial f(0, -1). \)

When \( f \) is locally Lipschitz at \( x \), \( f^-(x; \cdot) \leq f^c(x; \cdot) \leq f^*(x; \cdot) \leq f^0(x; \cdot) \) always hold. We say that \( f \) is regular at \( x \in A \) if \( f^-(x; v) = f^0(x; v) \) for each \( v \in X \), and \( f \) is pseudo-regular
1.3 Subderivatives and subdifferentials

at $x \in A$ if $f^+(x; v) = f^0(x; v)$ for each $v \in X$. If $f$ is convex and Lipschitz near $x$, then $f$ is regular at $x$ [33]. For a lower semicontinuous function $f$, if $\partial f(x) \neq \emptyset$ then $f$ is Clarke subdifferentially regular at $x$ if and only if $\partial_-(f(x)) = \partial_+ f(x)$ [80]. Mordukhovich and Ioffe define $f$ to be approximate subdifferentially regular at $x$ if $\partial_a f(x) = \partial_-(f(x))$ [56, 69].

The following example largely from [11] illustrates the distinctions among $\partial_-$, $\partial_a$, $\partial_c$.

**Example 1.2.4** The $k$'th order statistic $\phi_k : \mathbb{R}^n \to \mathbb{R}$ is the $k$'th largest element of $\{x_1, x_2, \ldots, x_n\}$, and $\phi_k$ is Lipschitz with respect to $x$. We denote the canonical basis in $\mathbb{R}^n$ by $\{e^1, e^2, \ldots, e^n\}$. When $k = 1$, $\phi_1$ is convex, all subdifferentials are convex subdifferentials. When $k > 1$:

1. At any point $x \in \mathbb{R}^n$, directly using the definition we have

$$
\partial_\phi \phi_k(x) = \begin{cases}
\co \{ e^i : x_i = \phi_k(x) \}, & \text{if } \phi_{k-1}(x) > \phi_k(x), \\
\emptyset, & \text{otherwise.}
\end{cases}
$$

2. At $0$ we have $\partial_c \phi_k(0) = \Delta$ whereas $\partial_a \phi_k(0) = \{ y \in \Delta : | \supp y | \leq n - k + 1 \}$, where the number of elements in $\supp y := \{ i : y_i \neq 0 \}$ is denoted by $| \supp y |$ and $\Delta := \{ y \in \mathbb{R}^n : y \geq 0, \sum y_i = 1 \}$.

Indeed, $\partial_a \phi_k(0) \subseteq \{ y \in \Delta : | \supp y | \leq n - k + 1 \}$ by (1). Conversely, given any vector $y$ belonging to the right-hand-side, choose a subset $J$ of exactly $(k - 1)$ indices $i$ for which $y_i$ is zero. Then, for any small $\delta > 0$, $y \in \co \{ e^i : i \notin J \} = \partial_\phi \phi_k(\delta \sum_{i \in J} e_i)$, whence, by taking limits, $y \in \partial_a \phi_k(0)$. $\partial_c \phi_k(0)$ follows by taking convex hulls.

3. However, for every $x \in \mathbb{R}^n$ we have $\partial_c \phi_k(x) = \partial_{mp} \phi_k(x) = \co \{ e^i : x_i = \phi_k(x) \}$.

When $\phi_{k-1}(x) > \phi_k(x)$, this follows from (1). Assume $\phi_{k-1}(x) = \phi_k(x)$. Suppose

$$
x_{i_1} \geq \ldots \geq x_{i_{k-1}} = x_{i_k} = \ldots = x_{i_s} > x_{i_{s+1}} \geq \ldots \geq x_{i_n}.
$$

Let $u = e_{i_1} + e_{i_2} + \ldots + e_{i_{k-1}}$. Then $\phi_k(x + tu) = \phi_k(x)$ for $t > 0$. For sufficiently small $\|h\|, t$ we have $\min \{ x_{i_1} + t + h_{i_1}, \ldots, x_{i_{k-1}} + t + h_{i_{k-1}} \} > x_{i_k} + th_{i_k}$, and

$$
x_{i_s} + th_{i_s} > \max \{ x_{i_{s+1}} + th_{i_{s+1}}, \ldots, x_n + th_n \},
$$

which implies $\phi_k(x + th + tu) = \phi_k(x) + t \max_{m \in I \setminus \{i_{k-1}\}} \{ h_m \}$, so $\phi_k'(x, h) \geq \max_{m \in I \setminus \{i_{k-1}\}} \{ h_m \}$. By positive homogeneity of $\phi_k'(x, \cdot)$, $\phi_k'(x, h) \geq \max_{m \in I \setminus \{i_{k-1}\}} \{ h_m \}$ for all $h \in \mathbb{R}^n$. As $i_{k-1}$ is arbitrary from $I$, $\phi_k'(x, h) \geq \max_{m \in I \setminus \{i_{k-1}\}} \{ h_m \}$ for all $h \in \mathbb{R}^n$, so $\partial_{mp} \phi_k(x) \supset \co \{ e_m : m \in I \}$. Since $\partial_{mp} \phi_k(x) \subset \partial_c \phi_k(x) = \co \{ e_m : m \in I \}$, we have $\partial_{mp} \phi_k(x) = \co \{ e_m : m \in I \}$. 


1.3 Subdifferential integrability

Whitney constructed a $C^1$ function $f$ on $[0,1] \times [0,1]$ such that $\nabla f(x) \equiv 0$ for each $x \in A$, where $A$ is a connected set, but $f$ is not constant on $A$ [92]. Of course the interior of $A$ is empty, and $A$ is not a union of rectifiable arcs. This shows that it only makes sense to discuss integrability on open connected sets.

**Definition 1.3.1** A locally Lipschitz function $f$ defined on an open convex set $U$ is $\mathcal{H}$-integrable if for any $g$ satisfying $\partial_2 g(x) \subset \partial_2 f(x)$ for all $x \in U$ we may deduce that $f$ and $g$ differ only by a constant, where $\mathcal{H}$ stands for one of ‘c’, ‘a’, ‘mp’, ‘hs’.

We will see that a Lipschitz function can be uniquely determined by its approximate or Clarke subdifferential largely because its derivative exists on many points. Moreover, the derivative map has some continuity properties. Thus its Clarke subdifferential or approximate subdifferential is minimal in an appropriate sense. The derivative of a Lipschitz function with a large subdifferential map exists only on first category sets (see pages 57, 104 and 143).

**Definition 1.3.2** A set-valued map $T : X \to 2^{X^*}$ is said to be $n$-cyclically monotone provided $\sum_{1 \leq k \leq n}(x_k^*, x_k - x_{k-1}) \geq 0$ whenever $n \geq 2$ and $x_0, x_1, \ldots, x_n = x_0 \in X$, and $x_k^* \in T(x_k)$, $k = 1, \ldots, n$. We say $T$ is cyclically monotone if it is $n$-cyclically monotone for every $n$. A monotone operator $T$ is said to be maximal cyclically monotone provided $T = S$ whenever $S$ is cyclically monotone and $\text{Gr}(T) \subset \text{Gr}(S)$.

The integrability of convex functions has been characterized [78, 79]:

**Theorem 1.3.3 (Rockafellar)** If $f$ is a proper lower semicontinuous convex function on a Banach space $X$, then its subdifferential $\partial f$ is maximal cyclically monotone. The function is uniquely determined by its subdifferential mapping up to an additive constant.

The integrability of nonconvex Lipschitz functions is very different.

**Example 1.3.4** Given two continuous functions $f, g : \mathbb{R} \to \mathbb{R}$ with $f \neq g$, there exist distinct Lipschitz functions $h_1, h_2$ with $h_1(0) = h_2(0) = 0$ and such that the Clarke subdifferentials satisfy $\partial_c h_1(x) = \partial_c h_2(x) = \text{co}\{g(x), f(x)\}$ for every $x \in \mathbb{R}$. 

1.4 Minimal subdifferentials and partially strict differentiability

To this end, we recall the following fact: on \([a, b]\) with \(a < b\) there exists a Lebesgue measurable set \(E_a\) such that \(0 < \mu(E_a \cap I) < \mu(I)\) for each nondegenerate interval \(I \subset [a, b]\) (see Lemma 3.3.1). Since \(f \not= g\), without loss of any generality we assume there exists \(x_0\) with \(f(x_0) > g(x_0)\). Choose \(\delta > 0\) small enough such that \(f > g\) on \([x_0 - \delta, x_0 + \delta]\). With \([a, b] := [x_0 - \delta, x_0 + \delta]\), we may extend \(E_a\) and \([a, b] \setminus E_a\) periodically to all \(R\), and get Lebesgue measurable sets \(E_1\) and \(E_2 := R \setminus E\) with the property that both meet each nondegenerate interval \(I \subset R\) with positive measures. Define \(\phi_i(x) := \int_0^x \chi_{E_i}(s) \cdot (f - g)(s)ds\). Then \(\phi_i\) is locally Lipschitz on \(R\) and \(\phi_i\) is differentiable almost everywhere: \(\phi_i'(x) = \chi_{E_i}(x) \cdot (f - g)(x)\) for every \(x \in R \setminus N_i\) where \(N_i\) is a Lebesgue null set. Since \(f - g\) is continuous, \(f - g\) may be recovered by any dense subset of \(R\). We have

\[
\partial_c \phi_1(x) = \partial_c \phi_2(x) = \left[\min\{0, f(x) - g(x)\}, \max\{0, f(x) - g(x)\}\right].
\]

Setting \(G(x) := \int_0^x g(s)ds, h_i := G + \phi_i\) for \(i = 1, 2\), we have \(\partial_c h_1(x) = \partial_c h_2(x)\) for every \(x \in R\). Moreover \(h_1(0) = 0 = h_2(0)\) and \(h_1, h_2\) are distinct. \(\square\)

1.4 Minimal subdifferentials and partially strict differentiability

The continuous semi-norm \(p\) defined on \(l^\infty\) in Example 1.1.5 has a minimal subdifferential which is nowhere single-valued, otherwise \(p\) would be Gâteaux differentiable at some point.

We recall that a Banach space \(X\) is an Asplund space if every continuous convex function on \(X\) is generically Fréchet differentiable. Also, a Banach space \(X\) is class \((S)\) if whenever \(Y\) is a Baire space every minimal \(w^*-usco\) (minimal \(w^*-cusco\)) \(F : Y \to X^*\) is generically single-valued.

A Banach space \(X\) is an Asplund space if and only if whenever \(Y\) is a Baire space every locally bounded minimal \(w^*-usco\) (minimal \(w^*-cusco\)) \(F : Y \to X^*\) is generically single-valued and norm-to-norm continuous [10]. In particular, each Asplund space is class \(S\). If \(X\) is a Banach space whose dual norm is strictly convex, then \(X\) is of class \(S\) [12]. All smoothable Banach spaces belong to this class [76]. On a Banach space of class \(S\), Lipschitz functions with minimal subdifferentials do have some nice differentiability properties:

**Definition 1.4.1** Let \(Y\) be a Banach subspace of \(X\). Given a function \(f\) on \(X\), we say that
1.4 Minimal subdifferentials and partially strict differentiability

\( f \) is \( Y \)-strictly \( \beta \)-differentiable at \( x \) and has \( Y \)-strict \( \beta \)-derivative \( \nabla f(x) \) in \( X^* \) if \( f(x) \) is finite and as \( t \to 0, x' \to x \)

\[
\frac{f(x' + tu) - f(x')}{t} - (\nabla f(x), u) \to 0,
\]

uniformly in \( u \in V \) for every \( V \in \beta \) where \( \beta \) is a bornology of \( Y \).

**Proposition 1.4.2** Assume that \( f \) is locally Lipschitz on an open set \( A \subset X, Y \subset X \) is class \((S)\) (Asplund), and \( \partial c f : A \to X^* \) is minimal, then \( f \) is generically \( Y \)-strictly Gâteaux (Fréchet) differentiable on \( A \).

**Proof.** Let \( \tau : Y \to X \) defined by \( \tau(y) = y \) and \( \tau^* \) be the adjoint map of \( \tau \). \( \tau^* \) is the restriction map, that is, \( \tau^*(x^*) = x^*|_Y \) for \( x^* \in X^* \). Then \( Y^* = \tau^*X^* \) and \( \tau^* : X^* \to Y^* \) is linear and \( w^*\)-to-\( w^* \) continuous. Since \( \partial f : A \to X^* \) is minimal \( w^* \)-cusc, \( \tau^* \circ \partial f : A \to Y^* \) is minimal \( w^* \)-cusc [76, 20]. Because \( Y \) is of class \( S \), \( \tau^* \circ \partial f \) is generically single-valued on a set \( G \subset A \). Choosing \( x \in G, x^* := \tau^* \circ \partial c f(x) \), for \( y \in Y \) we consider

\[
Q(x', t, y) := \frac{f(x' + ty) - f(x')}{t} - (x^*, y) \in (\partial c f(\xi), y) - (x^*, y) = (\tau^* \circ \partial c f(\xi) - x^*, y),
\]

with \( \xi \in [x', x' + ty] \). Since \( \tau^* \circ \partial f \) is norm-to-\( w^* \) continuous at \( x \), \( Q(x', t, y) \to 0 \) when \( x' \to x, t \to 0 \). If \( Y \) is Asplund, \( \tau^* \circ \partial f \) is norm-to-norm usc at \( x \), and for every \( \epsilon > 0 \) we have \( \tau^* \circ \partial f(B_{\delta}(x)) \subset B_{\epsilon}(x^*) \) for \( \delta > 0 \) small, thus \( Q(x', t, y) \leq \epsilon \) for every \( \|y\| \leq 1 \) if \( \|x' - x\| \leq \delta/2, 0 < t < \delta/2 \). \( \square \)

Thus, continuous convex functions do have strictly partial differentiability on nice subspaces. Since \( c_0 \) is a norm closed separable subspace of \( l^\infty \), each continuous convex function defined on \( l^\infty \) is, generically on \( l^\infty \), partially strictly Gâteaux differentiable on \( c_0 \). Although the spaces \( l^\infty, L^\infty[0,1], \) and \( l^1(\Gamma) \) (\( \Gamma \) uncountable) do not admit equivalent Gâteaux differentiable norms [72, page 67], every equivalent norm is generically strictly partially Gâteaux differentiable on separable complete subspaces. \( L^1[0,1] \) has no equivalent norms which are Fréchet differentiable at every nonzero point, but every equivalent norm is, generically on \( L^1[0,1] \), strictly partially Fréchet differentiable with respect to each Asplund subspace. Recall \( L^1[0,1] \) contains a isomorphic copy of \( l^2 \).

However, we observe that not every ‘bad’ Banach space has ‘nice’ infinite dimensional subspaces. The Shur space \( l^1[60] \) is not an Asplund space as \( \| \cdot \|_1 \) is nowhere Fréchet differentiable, and in \( l^1 \) every Asplund subspace is finite dimensional.
Chapter 2

The tool-box

In this chapter, we pack the main tools for the development of the theory in this thesis.

2.1 Baire category and Zorn’s lemma

The Baire category theorem and Zorn lemma offer the most basic non-constructive methods for proving existence theorems.

Let $X$ be a topological space. A subset $A \subseteq X$ is nowhere dense in $X$ if $\text{int}(\overline{A}) = \emptyset$. A set $B \subseteq X$ is of first category in $X$ if $B$ is a countable union of nowhere dense sets of $X$. A set $E \subseteq X$ is called residual in $X$ if $X \setminus E$ is of first category in $X$. Thus a closed subset $A$ of $X$ is nowhere dense if and only if $X \setminus A$ is dense in $X$. Roughly speaking, sets of first category play a role in topology analogous to that of null sets in measure, although there is no direct overlap even on $X = [0, 1]$.

Definition 2.1.1 The topological space $X$ is Baire if the complement of every first category subset of $X$ is dense in $X$.

Theorem 2.1.2 (Baire)

(i) A complete metric space is a Baire space;

(ii) An open subset of a Baire space is a Baire space in the relative topology;
2.2 The strong separation theorem

(iii) A nonempty set of type $G_δ$ contained in a complete metric space is a Baire space.

Proofs of Proposition 2.1.2 may be found in [27, 49].

**Proposition 2.1.3** For a complete metric space $X$, every lower semicontinuous real function $f$ on an open subset $A$ of $X$ is continuous on a dense $G_δ$ subset of $A$.

For proof, see [49, page 109].

Let $P$ be a set. A binary relation $\preceq$ on $P$ is called a partial order on $P$ if for all $x, y, z \in P$, we have: (i) $x \preceq x$ (reflexive); (ii) $x \preceq y$ and $y \preceq z$ imply $x \preceq z$ (transitive); (iii) $x \preceq y$ and $y \preceq x$ imply $x = y$ (anti-symmetric). A set together with a partial order is called a partially ordered set. Let $P$ be a partially ordered set. A non-empty subset $C$ of $P$ is called a chain if for all $x, y \in C$, we have either $x \preceq y$ or $y \preceq x$. An element $m \in P$ is said to be maximal if $m \preceq x$ in $P$ implies $m = x$.

**Proposition 2.1.4 (Zorn's lemma)** Let $P$ be a non-empty partially ordered set. If every chain has an upper bound, then every element is dominated by some maximal element. In particular, $P$ has one maximal element.

For the proof, see [23].

2.2 The strong separation theorem

**Theorem 2.2.1** [49] For two disjoint convex sets $A$ and $B$ in a locally convex space $X$, if $A$ is compact and $B$ is closed then $A$ and $B$ can be strongly separated by a closed hyperplane. That is, there is a continuous linear functional $f$ on $X$ such that $\sup_{x \in A} f(x) < \inf_{x \in B} f(x)$.

**Lemma 2.2.2** The convex hull of a finite family of compact, convex subsets of a Hausdorff topological vector space is compact.

**Proof.** If $A_i (i = 1, \ldots, n)$ are non-empty convex subsets of $X$, their convex hull is

$$A := \left\{ \sum_{i=1}^{n} \lambda_i a_i : a_i \in A_i, \lambda_i \geq 0, \sum_{i=1}^{n} \lambda_i = 1 \right\}.$$
Thus if $P$ is the compact set $\{ (\lambda_1, \ldots, \lambda_n) : \lambda_i \geq 0, \sum_{i=1}^{n} \lambda_i = 1 \}$, $A$ is the continuous image of the set $P \times \prod_{i=1}^{n} A_i \subset \mathbb{R}^n \times X^n$; hence $A$ is compact if each $A_i$ is. \hfill \Box

In a finite dimensional normed linear space, $\text{co}A$ is compact if $A$ is compact [52]. For a compact set $A$ in a separated linear topological space $X$, $\text{co}A$ is not necessarily compact or even closed and $\overline{\text{co}}A$ is not necessarily compact [49, page 57].

We also need the converse of the Krein-Milman theorem [49, page 96].

**Proposition 2.2.3** In a separated locally convex space $X$, for a compact set $A$ in $X$ where $\overline{\text{co}}A$ is compact we have $\text{ext} \overline{\text{co}}A \subset A$, where $\text{ext} \overline{\text{co}}A$ denotes the extreme points of $\overline{\text{co}}A$.

**Proof:** Let $x_0 \in \text{ext} \overline{\text{co}}A$, and $V$ be any closed convex $0$-neighborhood. There exist points $y_i \in A$ ($i = 1, \ldots, n$) with $A \subset \bigcup_{i=1}^{n} (y_i + V)$. Let $W_i := \overline{\text{co}}(A \cap (y_i + V))$. Since each $W_i \subset \overline{\text{co}}A$ is compact and convex, by Lemma 2.2.2 we have $\overline{\text{co}}A = \text{co}(W_1 \cup \ldots \cup W_n)$. So $x_0 = \sum_{i=1}^{n} \lambda_i w_i$, where $w_i \in W_i$, $\lambda_i \geq 0$ ($i = 1, \ldots, n$) and $\sum_{i=1}^{n} \lambda_i = 1$. But $x_0$ is an extreme point of $\overline{\text{co}}A$, we have $x_0 = w_i$ for some $i$; hence $x_0 \in W_i \subset y_i + V \subset A + V$. Since $V$ is an arbitrary $0$-neighborhood and $A$ is closed, it follows that $x_0 \in A$. \hfill \Box

### 2.3 Mean-value theorems

**Theorem 2.3.1 (Lebesgue)** Let $f$ be a real-valued locally Lipschitz function defined on a non-empty open subset of the real line, which contains the non-degenerate interval $[a, b]$. Then there exists a Borel subset $M \subseteq [a, b]$, with positive measure, such that for each $t \in M$, $f'(t)$ exists and $f'(t) \leq (f(b) - f(a))/(b - a)$.

**Theorem 2.3.2** Let $x$ and $y$ be in Banach space $X$, suppose that $f$ is Lipschitz on an open set containing the line segment $[x, y]$. Then there exists a point $u \in (x, y)$ such that $f(y) - f(x) \in \langle \partial_x f(u), y - x \rangle$, where $\partial$ stands for ‘c’ or ‘mp’.

In the latter theorem, the ‘c’ case is due to Lebourg [33] while the ‘mp’ case is due to Borwein, Fitzpatrick and Giles [16].
2.4 Generating uscos and cuscos

If $G : X \to Y$ and $F : Y \to Z$ are multifunctions the composition $F \circ G$ is given by $F \circ G(x) := F(G(x)) = \{ z \in Z : z \in F(y), y \in G(x) \}$.

**Proposition 2.4.1** If $G : X \to Y$ and $F : Y \to Z$ are uscos then so is $F \circ G$. If $G$ and $F$ are cuscos and $F$ has a convex graph then $F \circ G$ is a cusco. In particular the sum of $w^*$-cuscos is a $w^*$-cusco.

**Proof.** The fact that $F \circ G$ is an usco directly follows from the definition. Now let $F$ and $G$ be cuscos, to show that $F \circ G$ is a cusco, it suffices to show $F \circ G$ is convex-valued. Suppose $z_1, z_2 \in F \circ G(x)$ and $0 \leq \lambda \leq 1$, then for some $y_1, y_2 \in G(x)$, we have $z_i \in F(y_i)$. As $F$ has a convex graph, we have $\lambda z_1 + (1 - \lambda) z_2 \in \lambda F(y_1) + (1 - \lambda) F(y_2) \subseteq F(\lambda y_1 + (1 - \lambda) y_2)$. As $G(x)$ is convex, $\lambda y_1 + (1 - \lambda) y_2 \in G(x)$. Hence $\lambda z_1 + (1 - \lambda) z_2 \in F \circ G(x)$.

Suppose $G_i : X \to Y$ is a cusco for $i = 1, 2$, taking $G := (G_1, G_2)$ and $F : Y \times Y \to Y$ by $F(y_1, y_2) := y_1 + y_2$, we see that $F \circ G = G_1 + G_2$ is a cusco. $\square$

If $f_1$ and $f_2$ are locally Lipschitz on $A$, by Proposition 2.4.1 we have: $\partial_c f_1 + \partial_c f_2$, $f_1 \partial_c f_2 + f_2 \partial_c f_1$ and $(f_2 \partial_c f_1 - f_1 \partial_c f_2)/f_2^2$ (if $f_2(x) \neq 0$ for $x \in A$) are cuscos. In general, these cuscos may be not Clarke subdifferentials any more, but when either $f_1$ or $f_2$ is strictly differentiable, they are. The following result is from [72].

**Proposition 2.4.2** Assume that $T$ is a $w^*$-usco from an open set $A$ to $2^{X^*}$. For each $x \in A$ define $(\overline{\partial}^w T)(x) := \overline{\partial}^w T(x)$. Then the map $\overline{\partial}^w T$ is a $w^*$-cusco. Hence, if $T_1, \ldots, T_n$ are $w^*$-cuscos, then $\overline{\partial} \{T_1, \ldots, T_n\}$ defined by $\overline{\partial} \{T_1, \ldots, T_n\}(x) := \overline{\partial} \{T_1(x), \ldots, T_n(x)\}$ is a $w^*$-cusco.

**Proof.** Let $W \subset X^*$ be a $w^*$-open neighborhood of $\overline{\partial}^w T(x)$. Because $(X^*, w^*)$ is locally convex, the compact convex set $\overline{\partial}^w T(x)$ has a neighborhood base of the form $U = \overline{\partial}^w T(x) + W$, where the convex sets $W$ form a neighborhood base of 0. Thus, we can assume that $U$ is of the form $U = \overline{\partial}^w T(x) + W$, where $W$ is a convex $w^*$-closed neighborhood of 0. By the upper semicontinuity of $T$ there exists an open neighborhood $V$ of $x$ in $A$ such that $T(V) \subset \overline{\partial}^w T(x) + W$. If follows that $\overline{\partial}^w T(V) \subset \overline{\partial}^w T(x) + W$, so $\overline{\partial}^w T$ is upper semicontinuous. Since $\bigcup_{i=1}^n T_i$ defined by $(\bigcup_{i=1}^n T_i)(x) := \bigcup_{i=1}^n T_i(x)$ is a $w^*$-usco,
and for each $x$ the set $\text{co}\{T_1, \ldots, T_n\}(x)$ is $w^*$-compact by Lemma 2.2.2, so $\text{co}\{T_1, \ldots, T_n\}$ is a $w^*$-cusco.

\[ \square \]

**Proposition 2.4.3** [72] Suppose $F_1$ and $F_2$ are set-valued maps from $X$ into $Y$ such that $\text{Gr}(F_2)$ is closed, $F_2 \subset F_1$ and $F_1$ is usco. Then $F_2$ is usco.

A multifunction $\Omega : A \to 2^{X^*}$ is said to be densely defined if the domain $\text{Dom}(\Omega) = \{x \in A : \Omega(x) \neq \emptyset\}$ is dense in $A$, and $\Omega$ is locally bounded if $\Omega(V)$ is norm bounded for small neighborhoods $V$ of $x$. Recall that a Banach space $X$ is said to be weakly compactly generated (WCG) provided there exists a weakly compact subset $K$ of $X$ whose linear span is dense in $X$. We say that the dual ball $B_{X^*}$ of a Banach space $X$ is $w^*$-sequentially compact if every sequence in $B_{X^*}$ possesses a $w^*$-convergent subsequence. To generate cuscos and uscos from densely defined multifunctions, we need [10]:

**Proposition 2.4.4** (Borwein) Let $\Omega$ be a densely defined set-valued mapping that maps from a topological space $A$ into the dual of a Banach space $X$. If $\Omega$ is locally bounded on $A$, then there exists a unique smallest $w^*$-usco ($w^*$-cusco) containing $\Omega$, denoted $\text{USC}(\Omega)$ ($\text{CSC}(\Omega)$) and given by,

$$\text{CSC}(\Omega)(x) := \bigcap \{w^*\Omega(V) : V \text{ is an open neighborhood of } x\},$$

$$\text{USC}(\Omega)(x) := \bigcap \{w^*\text{cl}\Omega(V) : V \text{ is an open neighborhood of } x\}.$$  

When $A$ is metrizable and $B_{X^*}$ is $w^*$-sequentially compact then $\text{CSC}(\Omega)$ and $\text{USC}(\Omega)$ are given sequentially by

$$\text{CSC}(\Omega)(x) = \overline{\text{co}}^w \{a : \text{ there exist } a_n \to_{w^*} a \text{ and } x_n \to x \text{ with } a_n \in \Omega(x_n)\},$$

$$\text{USC}(\Omega)(x) = w^*\text{cl}\{a : \text{ there exist } a_n \to_{w^*} a \text{ and } x_n \to x \text{ with } a_n \in \Omega(x_n)\}.$$  

If $X$ is WCG, the ‘$w^*\text{cl}$’ is superfluous in the second equation.

Note that $\text{CSC}(\Omega)$ and $\text{USC}(\Omega)$ are obtained by taking the topological upper limits. An usco (cusco) mapping $\Phi : A \to 2^{X^*}$ satisfies $\text{USC}(\Phi) = \Phi$ ($\text{CSC}(\Phi) = \Phi$). The next result reveals one of the key properties enjoyed by minimal uscos (cuscos).

**Lemma 2.4.5** [20] Let $A$ be a non-empty open subset of a Banach space $X$. If $T : A \to 2^{X^*}$ is a $w^*$-usco ($w^*$-cusco) and $S : A \to 2^{X^*}$ is a minimal $w^*$-usco ($w^*$-cusco) and $T(x) \cap S(x) \neq \emptyset$ for each $x \in A$, then $S(x) \subseteq T(x)$ for all $x \in A$. 
2.5 Haar null sets

In this section, we give some elementary properties of Haar null sets. See [20, 32, 85] for more. Assume that \( X \) is a Banach space. A Borel measure on \( X \) is a measure defined on \( B(X) \) - the Borel subsets of \( X \) (the \( \sigma \)-algebra generated by all open subsets of \( X \)). A Radon measure \( \mu \) on \( X \) is a Borel measure on \( X \), extended to its completion on \( X \), which satisfies:

(i) \( \mu(K) < \infty \) for each compact subset \( K \subset X \);

(ii) \( \mu(A) = \sup\{\mu(K) : K \subset A, K \text{ is compact}\} \) for each \( A \in B(X) \).

We call a subset \( A \subseteq X \) universally (Radon) measurable if it belongs to the \( \mu \)-completion of each finite Borel (Radon) measure on \( X \) and we shall denote by \( \mathcal{U}(X) \) (\( \mathcal{U}_R(X) \)) the family of all universally (Radon) measurable subsets on \( X \). Observe that \( B(X) \subset \mathcal{U}(X) \subset \mathcal{U}_R(X) \).

**Definition 2.5.1** A universally Radon measurable set \( N \subset X \) is a Haar-null set if there exists a Radon probability measure \( p \) (not necessarily unique) on \( X \) such that \( p(x + N) = 0 \) for each \( x \in X \). A subset \( N \subset X \) is a Haar-null set if it is contained in a universally Radon measurable Haar-null set. The complement of a Haar null set is called “prevalent”.

When \( X \) is separable, Ulam’s theorem (that is, every probability measure on a complete separable metric space is tight) shows \( \mathcal{U}(X) = \mathcal{U}_R(X) \), this reduces to Christensen’s Haar null sets; When \( X \) is finite dimensional, a universally measurable set \( N \) is a Haar null if and only if \( N \) is of zero measure set for any Haar measure on \( X \) [30, page 116], thus this reduces to Lebesgue null sets. The following lists some facts about null sets. See [8] for the proof.

**Proposition 2.5.2** Let \( X \) be a Banach space. Then

(i) Every subset of a Haar null set in \( X \) is Haar null.

(ii) Every Haar null subset in \( X \) has empty interior.

(iii) If \( \{A_j : j \in \mathbb{N}\} \) are Haar null sets in \( X \), so is \( \bigcup \{A_j : j \in \mathbb{N}\} \).

(iv) If \( X \) is infinite dimensional, then every compact set \( C \subset X \) is Haar null.
When $X$ is infinite dimensional, since every compact set is Haar null, there does not exist a (countably additive) probability Radon measure $p$ on $X$ with $p(A) = 0$ for every Haar zero set $A$, nor one such that every zero set for $p$ is a Haar zero set [30, page 119]. This follows from the tightness of every countably additive probability Radon measure with respect to the paving of compact sets.

**Example 2.5.3** In an infinite dimensional Banach space, every translation invariant measure which is not identically zero has the property that all open sets have infinite measure.

Indeed, when $X$ is infinite dimensional, by Riesz's lemma every open set $O$ contains a countable number of disjoint open balls with the same radius [38, page 3]. Take such a ball $B_r(x)$. Assume $p$ is a translation invariant measure and not identically zero. Since $p(B_r(x)) > 0$, we have $p(O) = +\infty$. 

We shall use the following version of Fubini’s theorem due to Borwein and Moors [8], to whom we refer for the proof.

**Theorem 2.5.4** If $H$ and $T$ are Banach spaces with $T$ being finite dimensional, then for each universally Radon measurable subset $A \subset H \times T$, the following are equivalent:

(i) The set $[A]_H(h) := \{ t \in T : (h, t) \in A \}$ is a Lebesgue null set of $T$, for almost all $h \in H$;

(ii) The set $A$ is a Haar-null set in the product space $H \times T$.

### 2.6 Lipschitz functions with minimal Clarke subdifferentials

Much of the work on the differentiability of convex functions relies on the minimality of the subdifferential map. In contrast, most non-convex locally Lipschitz functions have maximal subdifferentials as shown in Chapters 3, 5, and 6. For such functions it is impossible to use the techniques developed for convex functions to characterize their differentiability. But a large subclass of locally Lipschitz functions do have similar properties to those of convex functions, namely, essentially smooth functions.
2.6 Lipschitz functions with minimal Clarke subdifferentials

\textbf{Definition 2.6.1} A real-valued locally Lipschitz function defined on a nonempty open subset \( A \) of a Banach space \( X \) is called essentially smooth on \( A \) if for each \( y \in X \) the set

\[ A_y := \{ x \in A : f^0(x; y) \neq -f^0(x; -y) \}, \]

is a Haar null subset of \( X \). We use \( S_c(A) \) to denote the family of all real-valued essentially smooth locally Lipschitz functions defined on \( A \).

The following result is due to Borwein and Moor [8], we give a different proof by using Fubini’s theorem.

\textbf{Proposition 2.6.2} Let \( A \) be a non-empty open subset of a Banach space \( X \). Then each member of \( S_c(A) \) possesses a minimal Clarke subdifferential mapping, and if \( A \) is connected then each member of \( S_c(A) \) is integrable.

\textbf{Proof.} Let \( \phi \in S_c(A) \). Suppose that there exists a \( w^* \)-cusc \( T \) such that \( T(y) \subset \partial c \phi(y) \) for every \( y \in A \), but for some \( x \in A \), \( g \in \partial c \phi(x) \setminus T(x) \). Since \( T(x) \) is \( w^* \)-compact, it can be strongly separated from \( g \), so for some \( v \in X \), and \( \alpha \in \mathbb{R} \) we have \( \sigma_{T(x)}(v) < \alpha < \langle g, v \rangle \leq \phi^0(x, v) \). By the \( w^* \)-upper semicontinuity of \( T \) there exists a neighborhood \( U \) of \( x \) such that \( \sigma_{T(y)}(v) < \alpha \) for each \( y \in U \). By definition of \( \phi^0 \), for some \( y \in U \), \( y + \lambda v \in U \) we have \( [\phi(y + \lambda v) - \phi(y)] / \lambda < \alpha \). Since \( \phi \) is continuous, applying Fubini’s theorem we may choose \( y \) near by such that \( [\phi(y + \lambda v) - \phi(y)] / \lambda = \alpha \), and \( \phi \) is strictly differentiable almost everywhere on \( [y, y + \lambda v] \). By Lebesgue’s mean-value theorem there exists \( \xi \in [y, y + \lambda v] \) such that \( \phi \) is strictly differentiable at \( \xi \) and \( \phi'(\xi)(v) > \alpha \). But \( \xi \in U \), \( \phi'(\xi)v = T(\xi)v < \alpha \), a contradiction. We conclude that \( \partial c \phi \) is a minimal cusc.

We now show that \( \phi \) is integrable. Let \( g \) be any real-valued locally Lipschitz function defined on \( A \) such that \( \partial g(x) \subset \partial c \phi(x) \) for all \( x \in A \). Then \( g, \phi - g \in S_c(A) \), and so \( \partial c (\phi - g) \) is minimal. Since \( \phi = (\phi - g) + g \) we have that \( \partial c \phi(x) \subset \partial c (\phi - g)(x) + \partial g(x) \subset \partial c (\phi - g)(x) + \partial c \phi(x) \) and ‘cancellation’ of closed bounded convex sets implies that \( 0 \in \partial c (\phi - g)(x) \) for all \( x \in A \). Hence \( \partial c (\phi - g) \equiv \{0\} \) on \( A \). As \( A \) is connected, \( f - g \) is constant on \( A \). \( \square \)

One problem with the proof above is that it does not show how to recover the function \( f \) from its subdifferential map (see Chapter 7). \( S_c(U) \) contains: \( C^1 \) function, convex function, convex-concave saddle function, pseudo-regular function. \( S_c(U) \) is closed under addition, subtraction, multiplication and division (when this is defined), as well as lattice operations \([20, 10, 8]\). Hence ‘bad’ functions do not show up in \( S_c(U) \)!
2.6 Lipschitz functions with minimal Clarke subdifferentials

An invariant property of minimal subdifferentials

Observe that \( \partial_c f \) is minimal on \( A \) if and only if \( y(\partial_c f) := [-f^0(; ; -y), f^0(; ; y)] \) is minimal on \( A \) for each \( y \in S_X \) by Theorem 2.2.1 (see [68, 20]).

**Definition 2.6.3** Let \( f \) be real valued function defined on a topological space \( A \). \( f \) is called quasi lower (upper) semicontinuous on \( A \) if for each \( t_0 \in A \), \( \epsilon > 0 \) and open neighborhood \( U \) of \( t_0 \) there exists a non-empty open subset \( V \subset U \) such that \( f(t) > f(t_0) - \epsilon \) for all \( t \in V \) \( (f(t) < f(t_0) + \epsilon \) for all \( t \in V \).

**Definition 2.6.4** Let \( A \) be a non-empty subset of a Banach space \( X \). Then a subset \( S \) of \( A \) is 1-D almost everywhere in \( A \), in the direction \( y \), if for each \( x \in A \),

\[
\mu(\{t \in \mathbb{R} : x + ty \in A \text{ and } x + ty \notin S\}) = 0.
\]

By Lebesgue’s mean-value theorem, \([-f^0(; ; -y), f^0(; ; y)] = \text{CSC}(f^+(; ; y)|_S)\), where \( S \) is any 1-D almost everywhere subset, in the direction \( y \), in \( A \).

**Lemma 2.6.5** [20, 76] A \( w^* \)-cusco \( \Phi \) from a topological space \( A \) into subsets of the dual of a Banach space \( X \) is a minimal cusco (minimal usco) on \( A \) if, and only if, given any open subset \( U \) of \( A \) and \( w^* \)-closed convex subset (closed subset) \( K \) of \( X^* \), with \( \Phi(U) \nsubseteq K \), there exists a non-empty open subset \( V \) of \( U \) such that \( \Phi(V) \cap K = \emptyset \).

**Proof.** Suppose there exists an open set \( U \) of \( A \) and a \( w^* \)-closed convex subset \( K \) such that \( \Phi(U) \nsubseteq K \) and \( \{x \in A : \Phi(x) \cap K \neq \emptyset\} \) is dense in \( U \), then \( \Phi(x) \cap K \neq \emptyset \) for each \( x \in U \).

Define another cusco by

\[
S(x) := \left\{ \begin{array}{ll}
\Phi(x) \cap K & \text{for } x \in U \\
\Phi(x) & \text{otherwise}
\end{array} \right.
\]

Then \( S \) is a \( w^* \)-cusco and \( S(x) \subseteq \Phi(x) \) for every \( x \in A \). If \( \Phi \) is minimal we conclude that \( S = \Phi \). Therefore \( \Phi(x) \subseteq K \) for all \( x \in U \), a contradiction.

Conversely, if \( \Phi \) is not minimal then there exists a \( w^* \)-cusco \( S \) such that \( S(y) \subseteq \Phi(y) \) for every \( y \in A \), and there exists \( x \in A \) and \( g \in X^* \) such that \( g \in \Phi(x) \setminus S(x) \). Since \( S(x) \) is \( w^* \)-compact there exists a \( v \in X \) and \( \alpha \in \mathbb{R} \) such that

\[
S(x) \subseteq \{x^* \in X^* : x^*(v) < \alpha\} := W^- \text{ and } g \in \{x^* \in X^* : x^*(v) > \alpha\} := W^+.
\]
2.6 Lipschitz functions with minimal Clarke subdifferentials

By the $w^*$-upper semicontinuity of $S$ there is an open neighbourhood $U$ of $x$ such that $S(y) \subseteq W^-$ for every $y \in U$. If we let $K := \{x^*: x^*(v) \leq \alpha\}$ we see that $\Phi(U) \not\subseteq K$ but $\Phi(y) \cap K \supseteq S(y) \neq \emptyset$ for every $y \in U$, a contradiction. \hfill $\blacksquare$

**Lemma 2.6.6** \cite{20} Let $f$ be a locally Lipschitz function defined on a non-empty open subset $A$ of a Banach space $X$. Then $\partial_c f$ is a minimal $w^*$-cuso on $A$ if, and only if, for each $y \in S_X$ there exists a subset $P_y$ of $A$, which is 1-D almost everywhere in $A$, in the direction $y$, such that the function $f^+(\cdot; y)$ is quasi lower semicontinuous on $P_y$.

**Proof.** Suppose $\partial_c f$ is minimal, then CSC$(f^+(\cdot; y))$ is minimal on $A$ for each $y \in S_X$, thus $f^+(\cdot; y)$ is quasi lower semicontinuous on $A$ by Lemma 2.6.5. Conversely, suppose that for each $y \in S(X)$ there exists a subset $P_y$ of $A$ which is 1-D almost everywhere in $A$, in the direction $y$, such that the mapping $f^+(\cdot; y)$ is quasi lower semicontinuous on $P_y$. Let $S_y := \{t \in A : f'(t; y) = -f'(t; -y)\}$, and define $R_y := P_y \cap S_y \cap P_{-y}$. Since $P_y, P_{-y}$ and $S_y$ are 1-D almost everywhere in $A$, in the direction $y$, so is $R_y$. Now $R_y \subseteq S_y$, therefore $f^+(x; y) = f'(x; y) = -f'(x; -y) = -f^+(x; -y)$ and so $f^+(\cdot; y)$, restricted to $R_y$, is both quasi lower and upper semi-continuous on $R_y$, then CSC$(f^+(\cdot; y)|_{R_y})$ is minimal by Lemma 2.6.5. \hfill $\blacksquare$

**Lemma 2.6.7** Let $f$ be a locally Lipschitz convex function defined on a non-empty open subset $A$ of a Banach space $X$. Then for each $y \in S(X)$,

$$D_y := \{x \in A : f'(x; y) \text{ is continuous at } x\},$$

is 1-D almost everywhere in $A$, in the direction $y$.

**Proof.** Since $f$ is locally Lipschitz, the set $D_y := \{x \in A : f'(x; y) = -f'(x; -y)\}$ is 1-D almost everywhere. As $f$ is convex, $f'(x; y) = f^0(x; y)$ for all $y \in X$, thus at each $x \in D_y$, $f'(x; y)$ is continuous. \hfill $\blacksquare$

**Proposition 2.6.8** Let $f$ be a real-valued locally Lipschitz function defined on a non-empty open subset $A$ of a Banach space $X$. If $\partial_c f$ is a minimal $w^*$-cuso on $A$ and $g$ is a convex function, then $\partial_c(f + g)$ is a minimal $w^*$-cuso on $A$.

**Proof.** Fix $y \in S_X$. By Lemma 2.6.7, there exists a 1-D almost everywhere $D_y \subset A$ such that $g'(x; y)$ is continuous at $x \in D_y$. By Lemma 2.6.6, there exists a 1-D almost everywhere
2.7 Variational principles

$P_y \subset A$ such that $f^+(\cdot; y)$ is quasi lower semicontinuous on $P_y$. Since $(f + g)^+(x; y) = f^+(x; y) + g^+(x; y)$ for all $x \in A$ and the summation of a quasi lower semicontinuous function and a continuous function is quasi lower semicontinuous, we deduce that $(f + g)^+(\cdot; y)$ is quasi lower semicontinuous on $D_y \cap P_y$, which is 1-D almost everywhere in $A$ in the direction $y$. By Lemma 2.6.6, $\partial_c (f + g)$ is minimal on $A$. \hfill \Box

2.7 Variational principles

In infinite dimensional space, compactness is lost, the supremum $\sup \{ \|T(x)\| : x \in U\}$ is not necessarily attained, even when $T$ is a linear functional.

Example 2.7.1 A linear functional $f$ on $c_0$ may be defined by $f(x) := \sum_{n=1}^{\infty} 2^{-n} x(n)$. Clearly $|f(x)| \leq \|x\|$ for all $x$. Also, for each $n$, we have $\|e_1 + \ldots + e_n\| = 1$ and $f(e_1 + \ldots + e_n) = 1 - 1/2^n$, so $\|f\| = 1$. If $x$ is an arbitrary element of the unit ball of $c_0$, then $|x(n)| \leq 1$ for all $n$, and there exists $N$ such that $|x(n)| \leq 1/2$ for all $n > N$. So we have $|f(x)| \leq \sum_{n=1}^{N} 2^{-n} + 1/2 \sum_{n=N+1}^{\infty} 2^{-n} < 1$. This indeed happens in any nonreflexive space [60, page 173].

Variational principles, the key tools in nonconvex analysis, affirm that there is always a “near-by point” which actually minimizes or maximizes a slightly perturbed functional.

Theorem 2.7.2 (Ekeland’s variational principle) [39] Let $V$ be a complete metric space with associated metric $d$, and let $f : V \to R \cup \{+\infty\}$ be a lower semicontinuous function which is bounded below. If $u$ is a point in $V$ satisfying $f(u) \leq \inf_V f + \epsilon$ for some $\epsilon > 0$, then for every $\lambda > 0$ there exists a point $v \in V$ such that (i) $f(v) \leq f(u)$, (ii) $d(u, v) \leq \lambda$, (iii) for all $w \neq v$ in $V$, one has $f(w) + (\epsilon/\lambda)d(w, v) > f(v)$.

Theorem 2.7.3 (Borwein-Press smooth variational principle) [5] Let $X$ be a $\beta$-smoothable Banach space, let $f : X \to R \cup \{+\infty\}$ be a lower semicontinuous function bounded below and let $\epsilon > 0$ and $\lambda > 0$ be given. Suppose that $u$ satisfies $f(u) < \inf_X f + \epsilon$. Then there exists a locally Lipschitz $\beta$-smooth function $\phi$ on $X$ and $v \in X$ such that for all $x \in X$, $f(x) + \phi(x) \geq f(v) + \phi(v)$, while $\|v - u\| < \lambda$, $f(v) < \epsilon + \inf_X f$ and $\|\nabla \phi(v)\| < \frac{2\epsilon}{\lambda}$. 
2.7 Variational principles

When the Borwein-Preiss smooth variational principle is used for concave functions, it provides points of differentiability. The variational principle on finite dimensional subspaces is very useful for studying approximate subdifferentials. For its proof, we need a lemma concerning the differentiability of distance functions associated with finite dimensional subspaces.

Lemma 2.7.4 Let $X$ be a $\beta$-smooth Banach space and $F \subset X$ be non-empty closed convex. If for every $x \in X \setminus F$, there exists $p(x) \in A$ such that $\|x - p(x)\| = d_F(x)$, then $d_F$ has $\beta$-derivative $\nabla \| \cdot \| (x - p(x))$ at $x \notin F$, and $d_F^2$ is $\beta$-differentiable on $X$. In particular, when $F$ is a finite dimensional subspace of $X$, for each $z_0 \in X$ the function $d_{z_0 + F}^2$ is $\beta$-differentiable with $\nabla d_{z_0 + F}^2(x) \|_F \equiv 0$ for every $z \in X$.

Proof. Suppose $x \notin F$. Let $e^*$ be the $\beta$-derivative of the norm at $x - p(x)$. For any $y \in F$ and $t \in [0, 1]$, $\|x - [(1 - t)p(x) + ty]\| \geq \|x - p(x)\|$. Subtracting $\|x - p(x)\|$ from both sides, dividing by $t$, and taking limit as $t \downarrow 0$, produces $e^*(p(x) - y) \geq 0$. Now for every $h \in X$,

$$
\|x - p(x) + h\| - \|x - p(x)\| - \langle e^*, h \rangle \\
\geq d_F(x + h) - d_F(x) - \langle e^*, h \rangle = \inf_{y \in F} \|x + h - y\| - \langle e^*, x - p(x) \rangle - \langle e^*, h \rangle \\
\geq \inf_{y \in F} \langle e^*, x + h - y \rangle - \langle e^*, x + h - p(x) \rangle = \inf_{y \in F} \langle e^*, p(x) - y \rangle \geq 0.
$$

Since $e^*$ is the $\beta$-derivative of the norm at $x - p(x)$, these inequalities imply that $e^*$ is the $\beta$-derivative of $d_F$ at $x$. For $x \notin F$, by chain rule we have $\nabla d_F^2(x) = 2d_F(x)\nabla d_F(x)$; If $x \in F$ then $\nabla d_F^2(x) = 0$. If $x \notin F$, $\langle \nabla d_F^2(x), p(x) - y \rangle \geq 0$ for $y \in F$. When $F$ is a finite dimensional space, we have $\langle \nabla d_F^2(x), y \rangle = 0$ for every $x \in X$ and $y \in F$. Finally, observe that $d_{z_0 + F}^2(x) = d_F^2(x - z_0)$.

Proposition 2.7.5 Assume that $Y$ is a finite dimensional subspace of a $\beta$-smoothable Banach space $X$ and $h : X \to (-\infty, \infty]$ is lower semicontinuous on $X$. Let $\delta, \epsilon > 0$ be given. Suppose that $h$ is bounded below on $B_\delta[z_0]$, and $h(z) - h(z_0) > -\delta \cdot \epsilon$ for all $z \in z_0 + \delta B_Y$. Then there exist $z, z^*$ satisfying $\|z - z_0\| \leq \delta$, $z^* \in \partial h(z)$ and $\|z^*\|_Y \leq 2\epsilon$.

Proof. By assumption, we may assume that $X$ is endowed with a $\beta$-smooth norm. First for $K > 0$ large, we have

$$
\inf\{h(z) + K d_{z_0 + Y}^2(z) : z \in \delta B_\delta[z_0]\} > h(z_0) - \epsilon \cdot \delta. \tag{2.1}
$$
2.8 Some properties of dual balls of Banach spaces

Suppose not, then there exists $K_n \to +\infty$ such that $h(z_n) + K_n d_{z_0+Y}^2(z_n) \leq h(z_0) - \epsilon \cdot \delta$ for some $z_n \in B_{\delta}[z_0]$. Since $h$ is bounded below on $B_{\delta}[z_0]$, say by $L$, we have

$$d_{z_0+Y}^2(z_n) \leq \frac{h(z_0) - h(z_n) - \epsilon \cdot \delta}{K_n} \leq \frac{h(z_0) - L - \epsilon \cdot \delta}{K_n} \to 0,$$

as $n \to \infty$. Since $Y$ is finite dimensional, there exists $y \in Y$ such that $z_{n_k} \to y + z_0$. Since $\|z_n - z_0\| \leq \delta$, we have $\|y\| \leq \delta$. But by the lower semicontinuity of $h$ at $z_0 + y$,

$$0 \leq \limsup_{n_k \to \infty} K_{n_k} d_{z_0+Y}^2(z_{n_k}) \leq h(z_0) - \liminf_{n_k \to \infty} h(z_{n_k}) - \epsilon \cdot \delta \leq h(z_0) - h(z_0 + y) - \epsilon \cdot \delta < 0,$$

a contradiction.

Next applying the Borwein-Preiss principle to equation (2.1) with $\lambda = \delta$, we have $\|z - z_0\| < \delta$, $\|x^*\| \leq 2\epsilon$ such that

$$0 \in \partial h + Kd_{z_0+Y}^2 + I_{B_k(z_0)}(z) + x^* = \partial h(z) + K\nabla d_{z_0+Y}^2(z) + x^*.$$

Choose $z^* \in \partial h(z)$ with $z^* = -(x^* + K\nabla d_{z_0+Y}^2(z))$. By Lemma 2.7.4, $\nabla d_{z_0+Y}^2(z)|_Y = 0$, thus $\|z^*|_Y\| = \|x^*|_Y\| \leq 2\epsilon$. \qed

2.8 Some properties of dual balls of Banach spaces

Proposition 2.8.1 Let $X$ be an infinite dimensional Banach space, then $w^*clS_{X^*} = B_{X^*}$.

Proof. If not, there exists $x^* \in B_{X^*}$ with $\|x^*\| < 1$ such that $x^* \not\in w^*clS_{X^*}$, then there exists a $w^*$-neighbourhood $W$ containing 0 such that $(x^* + W) \cap w^*clS_{X^*} = \emptyset$. Because $X$ is infinite dimensional, $W$ contains a nontrivial linear space [60, page 191]. Choose $d$ with $\|d\| = 1$ such that $\{td : t \in \mathbb{R}\} \subset W$. Define $\phi(t) := \|x^* + td\|$. We have $\phi(0) < 1$ and

$$\phi(1 + \|x^*\|) \geq (1 + \|x^*\|) - \|x^*\| = 1,$$

thus for some $t_0$ we have $\phi(t_0) = 1$. But $x^* + t_0d \in (x^* + W) \cap S_{X^*}$, a contradiction. \qed

Proposition 2.8.2 Assume $X$ is a Banach space with a smooth norm $\| \cdot \|$. Then the set of extreme points of $B_{X^*}$, denoted by $\text{ext } B_{X^*}$, is norm dense in $S_{X^*}$.
2.8 Some properties of dual balls of Banach spaces

**Proof.** Take $x^* \in S_{X^*}$ with $\|x^*\| = 1$. For every $1 > \varepsilon > 0$, by the Bishop-Phelps theorem [72], there exists $y^* \in X^*$ with $\|y^* - x^*\| < \varepsilon/2$ such that $y^*(y) = \|y^*\|$ for some $\|y\| = 1$. Then $1 - \varepsilon/2 < \|y^*\| < 1 + \varepsilon/2$, and $\|y^*/\|y^*\| - x^*\| \leq \|y^*/\|y^*\| - y^*\| + \|y^* - x^*\| < \varepsilon$.

We show that $y^*/\|y^*\|$ is an extreme point of $B_{X^*}$. Indeed, if $y^*/\|y^*\| = \lambda z_1^* + (1 - \lambda)z_2^*$ for some $0 < \lambda < 1$ with $\|z_1^*\| = 1$, $\|z_2^*\| = 1$, then $1 = y^*(y)/\|y^*\| = \lambda z_1^*(y) + (1 - \lambda)z_2^*(y)$ implies $z_1^*(y) = 1 = z_2^*(y)$. We thus have $z_1^*, z_2^* \in \partial \|\cdot\|(y)$. The Gâteaux differentiability of $\|\cdot\|$ at $y$ implies $z_1^* = z_2^*$.

□

**Proposition 2.8.3** The dual unit ball $B_{X^*}$ of a Banach space $X$ is metrizable in the $w^*$-topology if and only if $X$ is separable.

For the proof, see [49, page 83].

**Theorem 2.8.4 (Hagler and Sullivan)** The dual ball of a smoothable Banach space is $w^*$-sequentially compact.

For the proof, see [51].
Chapter 3

Subdifferentials in $\mathbb{R}$

Nonsmooth analysis deals with nondifferentiabilities. Little has been written on the subdifferentiabilities of the classical nondifferentiable examples. In this chapter, we study the subdifferentiabilities of nowhere monotone functions as they provide the best test ground of generalized subdifferentials. We will frequently use the following Dini derivatives:

$$f^+(x) := \limsup_{h \to 0^+} \frac{f(x+h) - f(x)}{h}, \quad f^-(x) := \limsup_{h \to 0^-} \frac{f(x+h) - f(x)}{h},$$

$$f^+(x) := \liminf_{h \to 0^+} \frac{f(x+h) - f(x)}{h}, \quad f^-(x) := \liminf_{h \to 0^-} \frac{f(x+h) - f(x)}{h}.$$ 

$$\mathcal{F}'(x) := \max\{f^+(x), f^-(x)\}, \quad \mathcal{F}'(x) := \min\{f^-(x), f^+(x)\}.$$ 

By “differentiable” we will always mean “having a finite derivative”. Different choices of $w$ in Equation (1.4) provide inequalities linking the Michel-Penot subderivative with the Dini derivatives of $f$ at $x$:

$$f^\circ(x, 1) \geq \max\{f^+(x), f^-(x)\}, \text{ and } f^\circ(x, -1) \geq \max\{-f^-(x), -f^+(x)\}.$$ 

3.1 Nowhere monotone functions

**Definition 3.1.1** We say a finite real function $f$ defined on $[0, 1]$ is nowhere monotone if $f$ is not monotone in any subinterval of $[0, 1]$. A nowhere monotone function $f$ is of the first species in $[0, 1]$ if there exists a real number $r$ such that the function $f(x) + r \cdot x$ becomes
3.1 Nowhere monotone functions

monotone in \([0, 1]\), and is of the second species in \([0, 1]\) provided that for every \(r \in \mathbb{R}\) the function \(f(x) + r \cdot x\) is also nowhere monotone.

If a nowhere monotone \(f\) is not the second species on \([0, 1]\), then for some \(r\) the function \(f(x) + r x\) is monotonic on some subinterval \(I \subset [0, 1]\), thus the complement of the second species need not be the first species. Since every nondifferentiable function \(f\) is nowhere monotone and for every \(r \in \mathbb{R}\) the function \(f(x) + r \cdot x\) is also nowhere monotone in \([0, 1]\), so every nondifferentiable function \(f\) is a nowhere monotone function of the second species.

**Definition 3.1.2** A continuous function \(f\) defined on \([0, 1]\) is said to be nondecreasing at \(x \in [0, 1]\) if there exists a \(\delta > 0\) such that \(f(t) \leq f(x)\) on \((x - \delta, x) \cap [0, 1]\) and \(f(t) \geq f(x)\) on \((x, x + \delta) \cap [0, 1]\), that is, \((f(t) - f(x))/(t - x) \geq 0\) for all \(t \neq x\) in some neighborhood of \(x\). The function \(f\) is nonincreasing at \(x\) if \(-f\) is nondecreasing at \(x\), and \(f\) is monotonic at \(x\) if it is either nondecreasing or nonincreasing at \(x\). We shall say that \(f\) is of monotonic type at \(x\) if there exists \(\nu \in \mathbb{R}\) such that \(f_{\nu}(x) := f(x) - \nu \cdot x\) is monotonic at \(x\). If \(f\) is not of monotonic type at any point of \([0, 1]\), we say \(f\) is of nonmonotonic type[25].

Recall Corollary 4.3[25]: Suppose \(f\) is continuous on \([a, b]\), \(f^+ \geq 0\) almost everywhere and \(f^+ > -\infty\) except, perhaps, on a countable set. Then \(f\) is nondecreasing.

**Proposition 3.1.3** Monotonic type at no point \(\Rightarrow\) monotonic at no point \(\Rightarrow\) nowhere monotone of second species \(\Rightarrow\) nowhere monotone.

**Proof.** Only the second \(\Rightarrow\) needs a proof. Let \(f\) be monotonic at no point. If \(f\) is not nowhere monotone of second species, then there exists \(m\) such that \(f(x) - mx\) is monotone on some subinterval \([a, b]\). Without loss of generality we assume that \(f(x) - mx\) is nondecreasing on \([a, b]\). Then \(f(x) - mx\), therefore \(f\) is differentiable almost everywhere on \([a, b]\). Since \(f\) is monotonic at no point, \(f'(x) = 0\) almost everywhere on \([a, b]\). Moreover, since \(f(x) - mx\) is nondecreasing on \([a, b]\), \(f_+(x) \geq m\) everywhere on \([a, b]\). Thus, \(f\) is nondecreasing on \([a, b]\), which is a contradiction. \(\Box\)

The following shows the nonreversibility of the implications in Proposition 3.1.3.

**Example 3.1.4** (1) Every differentiable nowhere monotone function is not nowhere monotone of second species. Indeed, if \(f\) is nowhere monotone, then

\[D = \{x : f'(x) = 0, f' \text{ is continuous at } x\},\]
is residual. Given $m > 0$, for every $x \in D$ there exists a neighborhood $N_x$ of $x$ in which $|f'(y)| < m/2$ since $f'$ is continuous at $x$, then $f(y) + my$ is increasing on $N_x$.

(2) Theorem 3.3.2 will give an absolutely continuous nowhere monotone function $f$ of second species, but $f$ is monotonic at each $x$ with $f'(x) > 0$ or $f'(x) < 0$.

(3) Let $M \subset [0, 1]$ be a first category $F_\sigma$ set. Then there exists a continuous function $f : [0, 1] \rightarrow \mathbb{R}$ such that $f'(x) = 0$ for $x \in M$, $f$ is nonmonotone at each $x \in M$ and $f$ is of nonmonotone type at each $x \in [0, 1] \setminus M$ [24]. Thus, $f$ is monotonic at no point, but $f$ is monotonic type at each $x \in M$.

### 3.2 Properties of nowhere monotone functions

**Definition 3.2.1** A function $f$ is nondecreasing (nonincreasing) on the right of a point $t$ if there exists a real number $h > 0$ such that $f(x) \geq f(t)$ ($f(x) \leq f(t)$) for $t < x < t + h$. If $f$ is neither nondecreasing nor nonincreasing on the right of $t$, we say $f$ is oscillating on the right of $t$ or is $O_+$ at $t$. The property that $f$ is oscillating on the left of $t$ or is $O_-$ at $t$ is defined in a similar way.

**Lemma 3.2.2** [42] If $f : [0, 1] \rightarrow \mathbb{R}$ is continuous and nowhere monotone, there exists a residual set $G \subset [0, 1]$ such that $f_-(x) \leq 0 \leq f^-(x)$ and $f_+(x) \leq 0 \leq f^+(x)$ if $x \in G$.

**Proof.** We show that $E := \{ x \in [0, 1] : f$ is nondecreasing on the right of $x \}$ is first category. Define $E_n := \{ x \in [0, 1] : f(t) \geq f(x)$ for $x < t < x + 1/n \}$. Then $E = \bigcup_{n=1}^{\infty} E_n$.

First, $E_n$ is closed. Indeed, assume $x_k \in E_n$ and $x_k \rightarrow x$. For $0 < t - x < 1/n$, when $k$ is large we have $0 < t - x_k < 1/n$, so $f(t) \geq f(x_k)$. By the continuity of $f$ we have $f(t) \geq f(x)$. Next, $E_n$ is nondense in $[0, 1]$. Let $I$ be an arbitrary interval contained in $[0, 1]$, and $J = [a, b] \subset I$ with $b - a < 1/n$. Since $f$ is nowhere monotone in $[0, 1]$, it is not nondecreasing in $J$, so there exist points $c, d \in J, c < d$ such that $f(c) > f(d)$. Let $m := \min\{ f(t) : t \in [c, d] \}$. Since $f(c) > f(d) \geq m$, there exists $c' \in [c, d]$ such that $f(x) > m$ if $x \in [c, c']$. Choosing $t \in [c, d]$ with $f(t) = m$ we then have $t > c'$. If $x \in [c, c']$, then $0 < t - x < 1/n$ and $f(t) = m < f(x)$, so $x \not\in E_n$. Therefore $E_n$ is nondense in $[0, 1]$. Similar arguments show that the set of points at which $f$ is nondecreasing on the right is first category. Then $f$ is $O_+$ at a residual subset $G_+ \subset [0, 1]$. If $t \in G_+$, then $(t, t + h)$
contains two points \( t_1, t_2 \) with \( f(t_1) < f(t) < f(t_2) \) for every \( h > 0 \). As \( f \) is continuous, there exists \( x \in (t, t + h) \) such that \( f(x) = f(t) \). Therefore \( f_+(t) \leq 0 \leq f_-(t) \). Similarly, we obtain a residual subset \( G_- \subset [0, 1] \) such that \( f_-(t) \leq 0 \leq f^-(t) \) if \( t \in G_- \). Then the claim holds on \( G_- \cap G_+ \).

\[ \square \]

**Lemma 3.2.3** If \( f \) is continuous and nowhere monotone on \([0, 1]\) then the set of points at which \( f \) attains local minima is dense in \((0, 1)\).

**Proof.** Take an arbitrary \( x \in (0, 1) \) and \( h > 0 \) such that \([x - h, x + h] \subset (0, 1)\). We will show that \( f \) has a local minimum in \([x - h, x + h]\). Since \( f \) is nowhere monotone in \([x, x + h]\), it can not be non-increasing in \([x, x + h]\), and so there exist points \( c, d \in [x, x + h] \) such that \( c < d \) and \( f(c) < f(d) \). There exists \( \delta > 0 \) such that \( f(t) > f(c) \) on \([d - \delta, d]\) and \( d - \delta > c \). On \([x - h, x]\) the same arguments show that there exist \( c' > d' \) with \( c', d' \in [x - h, x] \) such that \( f(c') < f(d') \). There exists \( \delta' > 0 \) such that \( f(t) > f(c') \) on \([d', d' + \delta']\) and \( d' + \delta' < c' \). Hence the minimum of \( f \) on \([d', d]\) is attained in \((d' + \delta', d - \delta) \subset (x - h, x + h)\).

\[ \square \]

### 3.3 Rockafellar type functions

In this section, we construct absolutely continuous functions on \( \mathbb{R} \) such that \( \partial_c f = \partial_c f \equiv \mathbb{R} \). We show that Rockafellar’s function is Dini subdifferentiable only on a first category set. The following classical result is well-known.

**Lemma 3.3.1** The interval \([0, 1]\) can be expressed as a disjoint union of measurable sets, \([0, 1] = \bigcup_{k=1}^{\infty} B_k\), each of which has positive measure in every subinterval of \([0, 1]\).

**Proof.** We reproduce the simple proof given by Bruckner [26]. By a “thick Cantor set” we mean a nowhere dense perfect set of positive measure.

Let \( A_1 \) be a thick Cantor set contained in \([0, 1]\). Let \( A_2 := A_2^0 \cup A_2^1 \) where, for \( i = 0, 1 \), \( A_2^i \) is a thick Cantor set contained in \((i/2, (i + 1)/2)\) and such that \( A_1 \cap A_2 = \emptyset \). Inductively we obtain a sequence of sets \( \{A_k\} \) such that for each \( k \),

(i) \( A_k \cap (A_1 \cup A_2 \cup \cdots \cup A_{k-1}) = \emptyset \).

(ii) \( A_k \) is a union of thick Cantor sets, \( A_k := A_k^0 \cup A_k^1 \cup \cdots \cup A_k^{k-1} \), with, for each \( i = 0, 1, \cdots, k - 1 \), \( A_k^i \subset (i/k, (i + 1)/k) \).
Such a sequence can be defined because for every \( k \), the set \( A_1 \cup A_2 \cup \cdots \cup A_{k-1} \) is nowhere dense in \([0, 1]\). Now let \( A_0 := [0, 1] \setminus (\bigcup_{k=1}^{\infty} A_k) \). Define a sequence of \( B_k \) by

\[
B_1 := A_0 \bigcup_{n=0}^{\infty} A_{2n+1}, \quad \text{and} \quad B_{k+1} := \bigcup_{n=0}^{\infty} A_{2^k(2n+1)} \text{ for } k \geq 1.
\]

By (i) the sequence \( \{A_k\} \) and therefore the sequence \( \{B_k\} \) is a disjoint sequence of sets. Clearly, \([0, 1] = \bigcup_{k=1}^{\infty} B_k \). Let \( I \subset [0, 1] \) be a nondegenerate interval and let \( |I| \) denote its length. Choose \( n_0 \) so that \( 2/n_0 < |I| \). For each \( n \geq n_0 \), there exists a nonnegative integer \( i_n < n \) such that the interval \( (i_n/n, (i_n + 1)/n) \) is contained in \( I \). It follows that the set \( A_n \cap I \) has positive measure for every \( n \geq n_0 \). Since for each \( k \), the set \( B_k \) contains infinitely many of the sets \( A_n \), we infer that the set \( \mu(B_k \cap I) > 0. \)

**Theorem 3.3.2** Let \( A := \{a_1, a_2, \ldots, \} \) be any sequence of real numbers. There exists an absolutely continuous function \( F \) such that for every interval \( I \subset [0, 1] \) and every \( k \), the set \( \{x : F'(x) = a_k\} \cap I \) has positive measure.

**Proof.** Let \( B_k \) be a sequence of sets satisfying the conclusion of Lemma 3.3.1. We may assume that \( |a_k| \mu(B_k) < 1/k^2 \) for each \( k > 1 \). It follows that the function \( f \) defined by \( f(x) := a_k \) if \( x \in B_k \) is Lebesgue integrable, since

\[
\int_0^1 |f(x)| dx \leq |a_1| \mu(B_1) + \sum_{k=2}^{\infty} \frac{1}{k^2} < +\infty.
\]

Let \( F \) be defined by \( F(x) := \int_0^x f(t)dt \). Then \( F \) is absolutely continuous and \( F'(x) = f(x) \) a.e. in \([0, 1] \) [83, pages 107–110]. In particular for each \( k \), \( F' \) takes on the value \( a_k \) at almost all points of \( B_k \). The proof is completed since \( B_k \) has positive measure in \( I \). \( \Box \)

Theorem 3.3.2 is very useful in constructing pathological examples. In the sequel, by “infinitely many” we mean that the pairwise difference of these functions is not a constant.

**Corollary 3.3.3** There exist infinitely many strictly increasing and absolutely continuous functions \( F \) such that \( \partial_a F = \partial_c F \equiv [0, \infty) \). For each such a function \( F \), the inverse function \( F^{-1} \) satisfies \( \partial_a F^{-1} = \partial_c F^{-1} \equiv [0, \infty) \) on the range of \( F \), which is \( F([0, 1]) \).

**Proof.** Let \( A := \{r \in (0, \infty) : r \text{ is a rational number}\} = \{a_k\}_{k=1}^{\infty} \). Note that \( F(x) := \int_0^x f(s)ds \) where \( f(x) := a_k \) if \( x \in B_k \). Let \( x, y \in [0, 1] \) and \( x < y \). Taking any rational
3.3 Rockafellar type functions

\[ a_k > 0, \text{ we have} \]
\[ F(y) - F(x) = \int_x^y f(s)\,ds \geq \int_{(x,y)\cap B_k} f(s)\,ds \geq a_k\mu(B_k \cap (x,y)) > 0. \]

Thus \( F \) is strictly increasing. In particular, \( \partial_a F(x) \subset [0, \infty) \). Theorem 3.3.2 and (1.2) imply \( [0, \infty) \subset \partial_a F(x) \), thus \( \partial_a F(x) = [0, \infty) = \partial_c F(x) \) for every \( x \in [0,1] \). We proceed to compute \( \partial_a F^{-1} \) and \( \partial_c F^{-1} \). Since \( F \) is absolutely continuous, \( F \) maps sets of zero measure onto sets of zero measure and \( F(B_k) \) is measurable. Because \( F \) is strictly increasing on \([0,1]\) and \( F'(x) = a_k \) at almost every \( x \in B_k \), we have
\[ \mu(F(B_k) \cap [F(x), F(y)]) = \mu(F(B_k \cap [x, y])) = a_k\mu(B_k \cap [x, y]) > 0, \]
for any \( x < y \in [0,1] \). This shows that the range of \( F \) is a countable union of disjoint measurable sets \( \{F(B_k)\}_{k=1}^\infty \), each with positive measure in every subinterval of the range of \( F \). On \( F(B_k) \) we have \( (F^{-1})' = 1/a_k \) almost everywhere. The proof is completed by observing that \( \{1/a_k\}_{k=1}^\infty \) is also dense in \([0, \infty) \) and that \( F^{-1} \) is strictly increasing. \( \Box \)

**Corollary 3.3.4** There are infinitely many absolutely continuous functions such that \( \partial_a F = \partial_c F \equiv \mathbb{R} \) on \([0,1]\). For each such a function \( F \), there is a residual set \( G \) such that \( \partial_{mp} F(x) = \mathbb{R} \) if \( x \in G \).

**Proof.** Let \( A := \{r \in \mathbb{R} : r \text{ is rational}\} \). Then for arbitrary rational \( r \in A \), Theorem 3.3.2 and (1.2) imply \( r \in \partial_a F(x) \). Thus \( \mathbb{R} \subset \partial_a F(x) \subset \partial_c F(x) \subset \mathbb{R} \).

We proceed to compute \( \partial_{mp} F \). For every \( r \), the function \( F_r : [0,1] \to \mathbb{R} \) defined by \( F_r(x) := F(x) - rx \) is continuous. In every subinterval of \([0,1]\), there are positive measure sets on which \( F_r' > 0 \) and some positive measure sets on which \( F_r' < 0 \), thus \( F_r \) is a nowhere monotone function. By Lemma 3.2.2 the sets
\[ G_{-n} := \{x : F_-(x) \leq -n < n \leq F^-(x)\}, \text{ and } G_n := \{x : F_+(x) \leq -n < n \leq F^+(x)\}, \]
are residuals. The set \( G := \bigcap_{n=1}^\infty G_n \) is residual in \([0,1]\), and at \( x \in G \) we have
\[ F^+(x) = F^-(x) = +\infty \text{ and } F_+(x) = F_-(x) = -\infty. \]
This implies \( F^o(x; 1) \geq +\infty \) and \( F^o(x; -1) \geq +\infty \), and so \( \partial F(x) = \mathbb{R} \) if \( x \in G \). \( \Box \)

**Corollary 3.3.5** There exist infinitely many Lipschitz functions \( F \) on \([0,1]\) such that \( \partial_a F = \partial_c F \equiv [-1,1] \).
3.3 Rockafellar type functions

Proof. Choose \( A := \{ r \in [-1, 1] : r \text{ is a rational number} \} \). For every \( x \in [0, 1] \) and \( r \in A \), Theorem 3.3.2 and (1.2) imply \( r \in \partial_a F(x) \). Since \( r \) is arbitrary and \( \partial_a F(x) \subset [-1, 1] \) is closed, we have \( \partial_a F(x) = [-1, 1] \). \( \square \)

When \( A = \{-1, 1\} \) the function \( F \) is called *Rockafellar’s function*. The computation both of the approximate subdifferential and the Michel-Penot subdifferential of Rockafellar’s function is not immediately clear. One indirect way to compute its approximate subdifferential is to use the result given by Borwein and Fitzpatrick [19]. Below we give a direct approach by using nowhere monotone functions.

**Theorem 3.3.6** Let \( f \) be Rockafellar’s function. Then

(i) \( \partial_c f = \partial_a f = [-1, 1] \) on \([0, 1]\).

(ii) The set \( G := \{ x : f^+(x) = f^-(x) = 1, f_-(x) = f_+(x) = -1 \} \) is a residual set in \([0, 1]\).

Thus, \( f \) is subdifferentiable at most on a first category subset.

(iii) For \( x \in G \), \( \partial_{mp} f(x) = [-1, 1] \).

Proof. (i) Choose \(-1 < r < 1\). Consider the function \( g \) defined by \( g(x) := f(x) + rx \). Since both \( \{ x : g'(x) = 1 + r > 0 \} \) and \( \{ x : g'(x) = -1 + r < 0 \} \) are dense in \([0, 1]\), \( g \) is nowhere monotone and so \( g \) has local minimizers densely on \([0, 1]\). Let \( S_r \) denote those minimizers. If \( x \in S_r \), we have \( f(y) + ry \geq f(x) + rx \) for \( y \) near by \( x \). Then \(-r \in \partial_f(x)\).

Since \(-1 < r < 1\) is arbitrary, we have \( \partial_f(x) = [-1, 1] \).

(ii) and (iii) For \( n \geq 2 \), both the functions given by \( f(x) + (-1 + 1/n)x \) and \( f(x) + (1-1/n)x \) are continuous and nowhere monotone in \([0, 1]\), thus by Lemma 3.2.2

\[
G_{-n} := \{ x : f_-(x) \leq -1 + 1/n < 1 - 1/n \leq f^-(x) \},
\]

\[
G_n := \{ x : f_+(x) \leq -1 + 1/n < 1 - 1/n \leq f^+(x) \}.
\]

are residuals in \([0, 1] \). If \( x \in G := \bigcap_{n=2}^{\infty} (G_n \cap G_{-n}) \), we have \( f_-(x) \leq -1, f^-(x) \geq 1, f_+(x) \leq -1, f^+(x) \geq 1 \). Since \( f \) has Lipschitz constant \( 1 \), we deduce \( f_-(x) = f_+(x) = -1 \) and \( f^-(x) = f^+(x) = 1 \). Moreover, \( 1 \geq f^0(x;1) \geq \max \{ f^+(x), f^-(x) \} = 1 \), and

\[
1 \geq f^0(x;1) \geq \max \{ -f_-(x), -f_+(x) \} = 1.
\]

Hence \( \partial_{mp} f(x) = [-1, 1] \). \( \square \)
In many cases, Rockafellar’s function is the beginning point for building more pathological Lipschitz functions. In the following, we give one of many such applications.

**Lemma 3.3.7** Let \( F \) be continuously differentiable around \( z \) and \( g \) locally Lipschitz around \( F(z) \). Then \( \partial_a(g \circ F)(z) = F'(z) \cdot \partial_a g(F(z)) \).

**Proof.** By Corollary 5.4 [56] we have \( \partial_a(g \circ F)(z) \subseteq F'(z) \cdot \partial_a g(w) \) where \( w = F(z) \). To prove the reverse inclusion, we consider two cases: (1) if \( F'(z) = 0 \): since \( g \circ F \) is locally Lipschitz, \( \emptyset \neq \partial_a g \circ F(z) \subset \{0\} \), hence \( \partial_a(g \circ F)(z) = \{0\} \); (2) if \( F'(z) \neq 0 \): By the inverse function theorem, \( F \) is locally invertible around \( z \). Write \( g(w) = g(F \circ F^{-1}(w)) \). Then

\[
\partial_a g(w) \subseteq \partial_a(g \circ F)(F^{-1}(w)) \cdot (F^{-1})'(w) = \partial_a(g \circ F)(z) \cdot \frac{1}{F'(z)},
\]

That is, \( \partial_a g(w) \cdot F'(z) \subseteq \partial_a(g \circ F)(z) \). \( \square \)

**Theorem 3.3.8** Suppose \( f_1 \) and \( f_2 \) are continuous on \( \mathbb{R} \). There exists a locally Lipschitz \( h: \mathbb{R} \rightarrow \mathbb{R} \) with \( \partial_a h(x) = \text{co}\{f_1(x), f_2(x)\} \) for every \( x \in \mathbb{R} \).

**Proof.** Let \( f \) denote Rockafellar’s function. Define \( F(x) := \int_0^x (f_1(s) - f_2(s)) \, ds \), \( k(x) := (f \circ F(x) + F(x))/2 \), and \( h(x) := k(x) + \int_0^x f_2(s) \, ds \). By Lemma 3.3.7 we have \( \partial_a k(x) = [0,1] \cdot (f_1(x) - f_2(x)) \), and so \( \partial_a h(x) = \text{co}\{f_1(x), f_2(x)\} \). \( \square \)

### 3.4 The Michel-Penot subdifferential on null sets

We now show that given any null set there exists a Lipschitz function such that its Michel-Penot subdifferential is large on that set. The following lemma is from [64, page 195].

**Lemma 3.4.1** Let \( F \subseteq \mathbb{R} \) be closed, \( T \subseteq \mathbb{R} \) be measurable, \( F \cap T = \emptyset \), and let \( \omega \) be any real, positive increasing function on \((0, +\infty)\). Then there is an open set \( U \) such that

\[
T \subset U \subset (\mathbb{R} \setminus F) \quad \text{and} \quad \mu((x - r, x + r) \cap (U \setminus T)) \leq \omega(r),
\]

whenever \( x \in F \) and \( r > 0 \).

**Proof.** Let \( d_F \) be the distance function associated with the closed set \( F \). For \( n \in \mathbb{N} \) we let \( R_n := \{x \in \mathbb{R} : d_F(x) > 1/n\} \). For each \( n \in \mathbb{N} \) there is an open set \( U_n \subseteq R_n \) such that

\[
T \cap R_n \subseteq U_n \quad \text{and} \quad \mu(U_n \setminus (T \cap R_n)) = \mu(U_n \setminus T) < \epsilon_n,
\]

where \( \epsilon_n \to 0 \) as \( n \to \infty \).
3.4 The Michel-Penot subdifferential on null sets

where \( \{ \epsilon_n \} \) is a sequence of positive numbers satisfying \( \sum_{j=k}^{\infty} \epsilon_j < \omega(1/k) \) for each \( k \in \mathbb{N} \). We set \( U := \bigcup_{n=1}^{\infty} U_n \). Obviously, \( T \subset U \subset \bigcup_{n=1}^{\infty} R_n = \mathbb{R} \setminus F \). Let \( x \in F \) and \( r > 0 \). There is a smallest \( n \in \mathbb{N} \) for which \( 1 \leq nr \). Hence

\[
\mu((x - r, x + r) \cap (U \setminus T)) \leq \sum_{k=n}^{\infty} \mu(U_k \setminus T) \leq \sum_{k=n}^{\infty} \epsilon_k < \omega(1/n) \leq \omega(r). \quad \Box
\]

**Theorem 3.4.2** Let \( N \subset \mathbb{R} \) with \( \mu(N) = 0 \). Then there exists a Lipschitz function \( H \) on \( \mathbb{R} \) such that \( \partial_{mp} H(x) = [0, 1] \) if \( x \in N \).

**Proof.** The proof follows Lemma 1 [54]. Inductively, we define a sequence \( \{ G_n \}_{n=1}^{\infty} \) of open subsets of \( \mathbb{R} \) in the following fashion.

(i) Choose an open set \( G_1 \supset N \) such that \( \mu(G_1) < 1 \);

(ii) Once an open set \( G_n \supset N \) is defined, we choose an open set \( G_n \supset G_{n+1} \supset N \) such that \( \mu(G_{n+1}) < 1/(n+1) \), and whenever \( x \in \mathbb{R} \setminus G_n \) and \( h > 0 \) we have

\[
\mu((x - h, x + h) \cap G_{n+1}) < h/(n+1).
\]

The existence of \( G_{n+1} \) may be deduced as follows. After \( G_n \) has been defined, we set \( F := X \setminus G_n \), \( T := N \) and \( \omega(r) := r/(n+1) \). Applying Lemma 3.4.1, we obtain

\[
N \subset G_{n+1} \subset G_n \quad \text{and} \quad \mu((x - r, x + r) \cap G_{n+1}) \leq \frac{r}{n+1},
\]

whenever \( x \in \mathbb{R} \setminus G_n \) and \( r > 0 \). Moreover,

\[
\mu(G_{n+1}) \leq \sum_{n=1}^{\infty} \mu(U_n) < \sum_{n=1}^{\infty} \epsilon_n < \omega(1) = \frac{1}{n+1}.
\]

Now put \( P := \bigcup_{n=1}^{\infty} (G_{2n-1} \setminus G_{2n}) \) and \( H(x) := \int_{0}^{x} \chi_{P}(t) dt \). Clearly \( H \) is 1-Lipschitz function. We show that \( \partial_{mp} H(x) = [0, 1] \) if \( x \in N \). To this end, we consider a positive integer \( k \). Let \( (a_k, b_k) \) be the component of \( G_k \) which contains \( x \). By (ii), \( b_k - a_k < \frac{1}{k} \) and

\[
\mu(G_{k+1} \cap (x, b_k)) \leq \frac{1}{k+1}(b_k - x), \quad \mu(G_{k+1} \cap (a_k, x)) \leq \frac{1}{k+1}(x - a_k). \quad (3.1)
\]

If \( k \) is odd, then \( G_k \setminus G_{k+1} \subset P \) and therefore (3.1) gives

\[
\frac{H(b_k) - H(x)}{b_k - x} = \frac{\mu(P \cap (x, b_k))}{b_k - x} \geq 1 - \frac{1}{k+1} \quad \text{and}
\]
3.5 The space of nondecreasing continuous functions

\[
\frac{H(a_k) - H(x)}{a_k - x} = \frac{\mu(P \cap (a_k, x))}{x - a_k} \geq 1 - \frac{1}{k + 1}.
\]

If \( k \) is even, then \( P \cap (a_k, b_k) \subset G_{k+1} \) and therefore by (3.1) we have

\[
\frac{H(b_k) - H(x)}{b_k - x} \leq \frac{1}{k + 1} \quad \text{and} \quad \frac{H(a_k) - H(x)}{a_k - x} \leq \frac{1}{k + 1}.
\]

Since \( H \) is nondecreasing and 1-Lipschitz, we have \( H^+(x) \leq 1, H^-(x) \leq 1, H_+(x) \geq 0, H_-(x) \geq 0 \), and so \( H^+(x) = H^-(x) = 1 \) and \( H_+(x) = H_-(x) = 0 \). Then \( H^0(x; 1) = 1 \) and \( H^0(x; -1) = 0 \), and this shows \( \partial_{mp} H(x) = [0, 1] \) if \( x \in N \).

**Corollary 3.4.3** Let \( N \) be dense in \( \mathbb{R} \) with \( \mu(N) = 0 \). Then there exists a Lipschitz function \( H \) on \( \mathbb{R} \) such that \( \partial_{mp} H(x) = [0, 1] \) on a residual set containing \( N \) and \( \partial_c H \equiv [0, 1] \) on \( \mathbb{R} \).

**Proof.** Let \( H \) be the function given in Theorem 3.4.2. The mean-value theorem in Michel-Penot subdifferential form implies \( \partial_c H(x) = \limsup_{y \to x} \partial_{mp} H(y) \) [16]. When \( N \) is dense in \( \mathbb{R} \), we obtain \( \partial_c H(x) \equiv [0, 1] \) for every \( x \in \mathbb{R} \). For every \( n \geq 2 \), both functions given by \( H(x) - x/n \) and \( H(x) - (1 - 1/n)x \) are nowhere monotone in \( \mathbb{R} \). By Lemma 3.2.2 the sets

\[
G_{-n} := \{ x : H_-(x) \leq 1/n \leq 1 - 1/n \leq H(x) \},
\]

\[
G_n := \{ x : H_+(x) \leq 1/n \leq 1 - 1/n \leq H^+(x) \},
\]

are residuals. Since \( H \) is nondecreasing and has Lipschitz constant 1, the set

\[
G := \bigcap_{n=2}^{\infty} (G_{-n} \cap G_n) = \{ x \in \mathbb{R} : H_+(x) = H_-(x) = 0, H^+(x) = H^-(x) = 1 \},
\]

is residual in \( \mathbb{R} \) and \( N \subset G \). If \( x \in G \), we have \( \partial_{mp} H(x) = [0, 1] \).

3.5 The space of nondecreasing continuous functions

Consider \( X := \{ f : f \text{ is continuous and nondecreasing on } [a, b] \} \) with metric

\[
\rho(f, g) := \sup_{x \in [a, b]} |f(x) - g(x)| \text{ for } f, g \in X.
\]

For \( \nu \in \mathbb{R} \), we define \( f_{-\nu} : [a, b] \to \mathbb{R} \) by \( f_{-\nu}(x) := f(x) - \nu \cdot x \).
3.5 The space of nondecreasing continuous functions

Theorem 3.5.1 In \((X, \rho)\), the set \(\{ f \in X : \partial_x f = \partial_y f \equiv [0, +\infty) \}\) is a residual set.

Proof: Let \(I\) denote an open subinterval of \([a, b]\), and let

\[
A^n_I := \{ f \in X : \text{there exists } \nu \in [1/n, n] \text{ such that } f_{-\nu} \text{ is nondecreasing on } I \},
\]

\[
B^n_I := \{ f \in X : \text{there exists } \nu \in [1/n, n] \text{ such that } f_{-\nu} \text{ is nonincreasing on } I \}.
\]

(1) \(A^n_I\) is closed. Assume \(\{ f_n \} \subset A^n_I\) is Cauchy. Then \(f_n \to f\) uniformly for some \(f \in X\). For each \(k\), there exists \(\nu_k \in [1/n, n]\) such that \(f_k(x) - \nu_kx \geq f_k(y) - \nu_ky\) for all \(x \geq y\) with \(x, y \in I\). There exists an increasing sequence \(\{k_i\}\) such that \(\{\nu_{k_i}\}\) converges to some \(\nu \in [1/n, n]\). Taking the limits, we have \(f(x) - \nu x \geq f(y) - \nu y\) for \(x \geq y\) with \(x, y \in I\). Similar arguments show that \(B^n_I\) is closed.

(2) To show that \(A^n_I\) is nowhere dense, with \(f \in X\) we verify that every open ball \(B_{2\epsilon}(f)\) contains points of \(X \setminus A^n_I\). Fix \(x_0 \in I\) and define a nondecreasing \(h\) by \(h(x) := f(x_0) + \epsilon + \min\{x - x_0, 0\}\). Then \(h_1 := \max\{f, h\}\) and \(h_2 := \min\{f + 2\epsilon, h_1\}\) are continuous and nondecreasing. As \(h_1 \geq f\), \(f + 2\epsilon \geq h_2 \geq f\). For \(\delta > 0\) sufficiently small, we have \(f(x_0) - \epsilon \leq f(y) \leq f(x_0) + \epsilon\) for \(|y - x_0| \leq \delta\). For \(x_0 + \delta \geq x_0 + \epsilon\), \(h(y) = f(x_0) + \epsilon\), thus \(h_1(y) = f(x_0) + \epsilon\) for \(x_0 \leq y \leq x_0 + \delta\). But \(f(x_0) + \epsilon \leq f(y) + 2\epsilon \leq f(x_0) + 3\epsilon\) for \(x_0 \leq y \leq x_0 + \delta\), then \(h_2(y) = f(x_0) + \epsilon\) for \(x_0 \leq y \leq x_0 + \delta\). For every \(\nu \in [1/n, n]\), on \([x_0, x_0 + \delta]\) we have \((h_2(y) - \nu \cdot y)^\prime = -\nu < 0\) almost everywhere, thus \(h_2(y) - \nu y\) is decreasing on \([x_0, x_0 + \delta], \) and \(h_2 \not\in A^n_I\).

To show that \(B^n_I\) is nowhere dense, we use similar arguments. Define \(h \in X\) by \(h(x) := \max\{(n + 1)(x - x_0), 0\} + f(x_0) - \epsilon, h_1 := \min\{f, h\}, \) and \(h_2 := \max\{f - 2\epsilon, h_1\}\). Then \(h_2 \in X\) and \(f - 2\epsilon \leq h_2 \leq f\). For \(\delta > 0\) sufficiently small, \(h_2(x) = (n + 1)(x - x_0)\) on \([x_0, x_0 + \delta]\). For every \(\nu \in [1/n, n]\), \((h_2(x) - \nu \cdot x)^\prime = n + 1 - \nu > 0\) almost everywhere, thus \(h_2(x) - \nu \cdot x\) is increasing on \([x_0, x_0 + \delta], \) and \(h_2 \not\in B^n_I\).

(3) Thus both \(A^n_I\) and \(B^n_I\) are nowhere dense and closed. The sets \(A_I := \bigcup_{n=1}^\infty A^n_I\) and \(B_I := \bigcup_{n=1}^\infty B^n_I\) are first category of type \(F_\sigma\) in \(X\). Let \(\{I_k\}\) be all open subintervals of \([a, b]\) having rational endpoints. The sets \(A := \bigcup_{k=1}^\infty A_{I_k}\) and \(B := \bigcup_{k=1}^\infty B_{I_k}\) are first category of type \(F_\sigma\). It follows that the set \(X \setminus (A \cup B)\) is a residual set of type \(G_\delta\). If \(f \in X \setminus (A \cup B)\), then for every \(\nu > 0\), the function \(f_{-\nu}\) is not monotonic on every \(I_k\), thus nowhere monotonic on \([a, b]\). The set of points at which \(f_{-\nu}\) attains local minimum is dense.
3.6 The space of automorphisms

in \([a, b]\). We have \(\nu \in \partial_a f(x)\) for every \(x \in [a, b]\). Since \(\nu \in (0, +\infty)\) is arbitrary, we have 
\([0, +\infty) \subset \partial_a f(x) \subset \partial_c f(x) \subset [0, +\infty)\), completing the proof of the theorem. \(\square\)

Combining Lemma 3.2.2 and Theorem 3.5.1 we see that a typical nondecreasing continuous real-valued function on \([a, b]\) has a finite derivative only on a first category set on \([a, b]\).

Compare this with Lebesgue’s Differentiation Theorem[83, page 100]: If \(f\) is an increasing real-valued function on the interval \([a, b]\), then \(f\) has a finite derivative almost everywhere.

3.6 The space of automorphisms

A function of bounded variation is called singular if it has almost everywhere a zero derivative. As a singular function has almost everywhere a zero derivative, all of its variation is centered at points of the complementary set of measure zero. So it is the set of measure zero which contributes towards the entire structure of a singular function.

**Definition 3.6.1** A homeomorphism \(h\) of an interval \([a, b]\) onto \([a, b]\) that satisfies \(h(a) = a\) and \(h(b) = b\) is called an automorphism on \([a, b]\).

Let \(H\) denote the family of strictly increasing continuous functions on \([0, 1]\) that have the endpoints fixed. Since a uniform limit of functions on \(H\) need not be strictly increasing, \(H\) is not closed in \(C[0, 1]\) (see page 44). But \(H\) is of type \(G_\delta\) in the complete space \(\overline{H}\) (the closure of \(H\) in \(C[0, 1]\)) and therefore topologically complete. Consequently, Baire category arguments can still be applied. The following lemma is from [27, pages 468–471].

**Lemma 3.6.2** Let \(A\) be a first-category subset of \([0, 1]\). Let \(H_1 := \{h \in H : \mu(h(A)) = 0\}\).

Then \(H_1\) is residual in the topologically complete space \(H\).

Now let \(A\) be a first category subset of \([0, 1]\) with \(\mu(A) = 1\). For \(h \in H_1\), \(\mu(h(A)) = 0\). Since \(h\) is differentiable almost everywhere we have \(h'(x) = 0\) for almost every \(x \in [0, 1]\) [86, page 323], so every \(h \in H_1\) is a strictly increasing continuous singular function.

**Lemma 3.6.3** If a singular function \(f\) is continuous and strictly increasing, then, for every real number \(r > 0\), the function \(f(x) - rx\) is nowhere monotone.

**Proof.** Assume the derivatives of \(f\) are bounded from above in some interval \(J \subset [0, 1]\). As \(f\) is increasing, its derivatives are \(\geq 0\) throughout \(J\). For a continuous function, the lower
and upper bounds of each its derivates are the same as those of the difference quotient 
\((f(y) - f(x))/(y - x)\) with \(x, y \in J, x \neq y\) [25]. It follows that \(f\) is Lipschitz on \(J\). Since \(f\) has zero derivative almost everywhere in \(J\), \(f\) is constant in \(J\). This contradicts the fact that \(f\) is strictly increasing in \([0, 1]\). Hence the derivates of \(f\) are unbounded from above in every subinterval of \([0, 1]\). Assume \(r > 0\). Define the function \(F_r\) by 
\[F_r(x) := f(x) - rx.\] 
\(F_r\) has derivates \(\geq 0\) at points everywhere dense in \([0, 1]\). Moreover, since \(f\) is singular, the function \(F_r\) also has a derivative \(-r < 0\) at an everywhere dense set of points in \([0, 1]\). This shows \(F_r\) is nowhere monotone on \([0, 1]\). \(\square\)

**Theorem 3.6.4** The set 
\[H_1 := \{f \in H : \partial_v f = \partial_a f \equiv [0, +\infty)\}\] 
is residual in the topologically complete space \(H\). Moreover, for every \(f \in H_1\) we have 
\[\partial_m f(x) = [0, +\infty)\] on a residual set of \([0, 1]\).

**Proof.** In Lemma 3.6.2, we chose \(A\) to be of first category and \(\mu(A) = 1\). As indicated, 
each \(f \in H_1\) is a continuous and strictly increasing singular function. For fixed \(r > 0\), 
Lemma 3.6.3 shows \(F_r\) is nowhere monotone. Each nowhere monotone continuous function 
have everywhere dense sets of maxima and minima. At each minimal point \(x \in (0, 1)\), we 
have \(0 \in \partial_- F_r(x)\) and so the set \(\{x \in [0, 1] : r \in \partial_- f(x)\}\) is dense in \([0, 1]\). This implies \(r \in \partial_a f(x)\) for every \(x \in [0, 1]\). Since \(r > 0\) is arbitrary and \(\partial_a f(x) \subset [0, +\infty)\), we have 
\([0, +\infty) = \partial_a f(x) = \partial_v f(x)\).

Next, given an \(f \in H_1\) and a natural number \(n\), since the functions \(F_n\) and \(F_{1/n}\) are 
both nowhere monotone in \([0, 1]\), by Lemma 3.2.2 there exists a residual set \(G_n\) in \([0, 1]\) such 
that when \(x \in G_n\) we have 
\[f_-(x) \leq \frac{1}{n} < n \leq f^-(x), \quad f^+(x) \leq \frac{1}{n} < n \leq f^+(x).\]
If \(x \in G := \bigcap_{n=1}^{\infty} G_n\), then \(f_-(x) \leq 0, \quad f^+(x) \leq 0, \quad f^+(x) = f^-(x) = +\infty\). Since \(f\) 
is increasing, its derivates are all non-negative, then \(f^-(x) = f^+(x) = 0\). The proof is 
complete by observing that 
\[f^2(x, 1) \geq \{f^+(x), f^-(x)\} = +\infty, \quad 0 \geq f^o(x, -1) \geq \max\{f^+(x), -f^-(x)\} = 0. \quad \square\]

The Cantor function \(f : [0, 1] \to [0, 1]\) is continuous and nondecreasing [86, pages 129–
130]. Besides, almost everywhere on \([0, 1]\), we have \(f'(x) = 0\). The most usual strictly increasing continuous singular function on \([0, 1]\) or \(R\) is constructed from Cantor’s function
3.7 The space of continuous functions $C[0,1]$

[86, page 210]. It is interesting to compute $\partial_a f$ and $\partial_c f$ on the Cantor ternary set $K$ associated with $f$. Because $f$ is not strictly increasing, Lemma 3.6.3 does not apply.

**Theorem 3.6.5** Let $f$ be the Cantor function $[0,1] \to [0,1]$ associated with the Cantor ternary set $K$. Then $\partial_a f(x) = \partial_c f(x) = [0, +\infty)$ if $x \in K$.

**Proof.** Fix $x \in K$ and $r > 0$. Assume $I \subset [0,1]$ is an arbitrary open subinterval with $x \in I$. Theorem 7.20 [27] shows that the set $\{x : f'(x) = +\infty, x \in I\}$ is uncountable. But it is not true that $f'(x) = +\infty$ at all two-sided limit points of $K$. Morse's theorem [25] shows that for every $\alpha > 0$ the set $\{x : f_+(x) = \alpha, x \in I\}$ has cardinality $c$. Choose $y \in I$ with $f'(y) = +\infty$. Consider $F_r$ defined by $F_r(x) := f(x) - r \cdot x$. Then $F_r'(y) = +\infty$, and for sufficiently small $\delta > 0$ we have $F_r(z) > F_r(y)$ if $z \in (y, y + \delta)$. Since $F_r' = -r$ almost everywhere, we may choose $\hat{y} < y$ with $F_r'(-\hat{y}) = -r$, and so there exists $\hat{\delta} > 0$ such that $F_r(z) > F_r(\hat{y})$ if $z \in (\hat{y} - \hat{\delta}, \hat{y})$. It follows that $F_r$ has a local minimizer in $(\hat{y} - \hat{\delta}, y + \delta) \subset I$. Then $0 \in \partial_- F_r(z)$ for some $z \in (\hat{y} - \hat{\delta}, y + \delta)$, that is, $r \in \partial_- f(z)$. Because $I$ is arbitrary, we have $r \in \partial_a f(x)$. But $r > 0$ is also arbitrary, thus $[0, +\infty) \subset \partial_a f(x)$. Since $f$ is nondecreasing, $\partial_a f(x) \subset [0, +\infty)$, thence $\partial_a f(x) = [0, +\infty)$. Now $\partial_a f(x) \subset \partial_c f(x) \subset [0, +\infty)$ implies $\partial_c f(x) = [0, +\infty)$.

When $f'(x) = +\infty$, $\partial_- f(x) = \emptyset$. Every open interval $I \subset [0,1]$ containing points of the Cantor set $K$ has uncountably many such points. Theorem 3.6.5 shows $\partial_a f(x) = \partial_c f(x) = [0, +\infty)$ at these points of $K$. We see that $f$ is not regular at uncountably many points on every open interval containing points of the Cantor set.

3.7 The space of continuous functions $C[0,1]$

Let $C[0,1]$ denote the Banach space of real-valued continuous functions $f$ defined on $[0,1]$ with the uniform norm $\|f\| := \sup_{0 \leq x \leq 1} |f(x)|$. We will show that a typical $f \in C[0,1]$ is an antiderivative of a constant Clarke, approximate and Michel-Penot subdifferential map, i.e., the set-valued map defined by $T(x) \equiv \mathbb{R}$ for every $x \in \mathbb{R}$. Moreover for every such $f$, its Dini subdifferential is non-empty only on a set which is Lebesgue null and first category, and its minimal Jeyakumar's convexificator may be chosen as the empty set almost everywhere.

For a Lipschitz function, its Clarke subdifferential $\partial_c f$ has a closed graph, but for a
continuous function \( f \), this might fail (see Example 5.8.3). However, the following result helps when we compute the Clarke subdifferential for continuous functions.

**Proposition 3.7.1** Assume \( \{x_k\}_{k=1}^\infty \) are local minimizers of \( g \) on a general Banach space \( X \). If \( x_k \to x \), \( g(x_k) \to g(x) \), and \( g \) is lower semicontinuous around \( x \), then \( 0 \in \partial_c g(x) \).

**Proof.** Suppose \( 0 \notin \partial_c g(x) \), we consider two cases: (1) if \( \partial_c g(x) = \emptyset \), by Theorem 2.9.1[33], \( g^+(x, 0) = -\infty \). (2) if \( \partial_c g(x) \neq \emptyset \), by Proposition 2.2.1, there exists \( h \in X \) such that

\[
g^+(x, h) = \sup \{ \langle x^*, h \rangle : x^* \in \partial_c g(x) \} < 0.
\]

In either case, \( g^+(x, h) < 0 \) for some \( h \in X \). Since

\[
g^+(x, h) = \sup_{\epsilon > 0} \limsup_{y \to x, g(y) \to g(x)} \inf_{\|w\| < \epsilon} \frac{g(y + tw) - g(y)}{t},
\]

then for every \( \epsilon > 0 \) and \( t_k \downarrow 0 \) we have

\[
0 > \limsup_{t_k \to 0} \inf_{\|w\| < \epsilon} \frac{g(x_k + t_kw) - g(x_k)}{t_k}.
\]

Since \( x_k \) is a local minimizer of \( g \) and \( \|w\| \leq \epsilon + \|h\| \) (thus \( w \) is bounded), we may take \( 0 < t_k < 1/k \) such that \( g(x_k + t_kw) \geq g(x_k) \) for every \( \|w - h\| < \epsilon \). For such \( \{t_k\} \) we have

\[
\limsup_{t_k \to 0} \inf_{\|w\| < \epsilon} \frac{g(x_k + t_kw) - g(x_k)}{t_k} \geq 0.
\]

But this contradicts equation (3.2). Hence \( 0 \in \partial_c g(x) \). \( \square \)

**Definition 3.7.2** The function \( f \) is said to have Jeyakumar’s convexifier, \( \partial^* f(x) \), at \( x \) if \( \partial^* f(x) \) is closed and for each \( v \in \mathbb{R} \) we have

\[
f^-(x; v) \leq \sup_{x^* \in \partial^* f(x)} \langle x^*, v \rangle, \quad \text{and} \quad f^+(x; v) \geq \inf_{x^* \in \partial^* f(x)} \langle x^*, v \rangle.
\]

In term of classical Dini derivatives, a closed set \( \partial^* f(x) \) is Jeyakumar’s convexifier of \( f \) at \( x \) if \( \max\{f^+(x), f^-(x)\} = \sup_{x^* \in \partial^* f(x)} x^* \) and \( \min\{f^-(x), f^+(x)\} \geq \inf_{x^* \in \partial^* f(x)} x^* \). Obviously one can always choose \( \partial^* f(x) = \mathbb{R} \). A convexifier, \( \partial^* f(x) \), of \( f \) yields both an upper convex approximation and a lower concave approximation to \( f \) at \( x \). The Clarke subdifferential and Michael-Penot subdifferential are convexifiers when \( f^+(x, \cdot) \) and \( f^-(x, \cdot) \)
are lower semicontinuous. Moreover, if \( f \) is locally Lipschitz then the approximate subdifferential and Treiman linear generalized subdifferential are convexifiers \([59]\). The interesting thing is to find minimal convexifiers.

Define \( E(f) := \{ x : f^+(x) = f^-(x) = +\infty, f^-(x) = f^+(x) = -\infty \} \) for \( f \in C[0, 1] \). Our main result is:

**Theorem 3.7.3** There exists a residual set of functions \( f \in C[0, 1] \) for each of which

1) the set \( E(f) \) is residual in \((0, 1)\) and \( \mu(E(f)) = 1 \).

2) if \( x \in E(f) \), every closed set in \( \mathbb{R} \), including the empty set, may be chosen as \( \partial^* f(x) \).

3) for every \( x \in E(f) \), \( \partial_- f(x) = \emptyset \).

4) for every \( x \in [0, 1] \), we have \( \partial_a f(x) = \partial_c f(x) = \mathbb{R} \).

5) for every \( x \in (0, 1) \), we have \( \partial_{mp} f(x) = \mathbb{R} \) and \( \partial_{mp} f(x) \neq \partial_- f(x) \).

**Proof of Theorem 3.7.3**

We prove Theorem 3.7.3 by piecing together results from \([25, 48, 71]\). Recall that a function \( f \) is called nonangular at \( x \) if \( f^-(x) \leq f^+(x) \) and \( f^+(x) \leq f^-(x) \).

**Lemma 3.7.4** The functions of nonmonotonic type form a dense subset, denoted by \( S_1 \), of type \( G_\delta \) in \( C[0, 1] \).

**Lemma 3.7.5** The nonangular functions form a dense set, denoted by \( S_2 \), of type \( G_\delta \) in \( C[0, 1] \).

The proofs of Lemma 3.7.4 and Lemma 3.7.5 may be found in \([25, \text{pages 212-213}]\). If \( f \in S_1 \cap S_2 \) then \( f \) is nowhere differentiable. Assume \( \partial_- f(x) \neq \emptyset \) at \( x \in (0, 1) \). If \( x^* \in \partial_- f(x) \) then \( f^-(x) \leq x^* \leq f^+(x) \). Since \( f \) is nonangular at \( x \), we have \( f^+(x) \leq f^-(x) \). This means \( \partial_- f(x) = \{ x^* : x^* = f^-(x) = f^+(x) \} \). Hence every \( f \in S_1 \cap S_2 \) is nowhere differentiable and \( \partial_- f(x) \) is either a singleton or empty at \( x \in (0, 1) \).

**Lemma 3.7.6** If \( f \in S_1 \) then \( \bar{f}(x) = +\infty \) and \( \underline{f}(x) = -\infty \) if \( x \in (0, 1) \).
3.7 The space of continuous functions $C[0,1]$

**Proof.** Fix $x \in (0,1)$ and $\nu \in (-\infty, \infty)$. Since $f$ is of nonmonotonic type at $x$, for every $n$ there exists $x_n \in (x - 1/n, x)$ and $y_n \in (x, x + 1/n)$ such that

$$\frac{f(y_n) - f(x)}{y_n - x} \geq \nu \quad \text{and} \quad \frac{f(x_n) - f(x)}{x_n - x} < \nu,$$

or

$$\frac{f(y_n) - f(x)}{y_n - x} < \nu \quad \text{and} \quad \frac{f(x_n) - f(x)}{x_n - x} \geq \nu.$$

As $n \to \infty$, we have either $f^+(x) \geq \nu$ or $f^-(x) \geq \nu$, so that $f'(x) \geq \nu$. Also, either $f^-(x) \leq \nu$ or $f^+(x) \leq \nu$, so that $f'(x) \leq \nu$. Since $\nu$ is arbitrary, it follows that $f'(x) = -\infty$, while $f'(x) = +\infty$. \hfill \Box

**Lemma 3.7.7** There is a residual subset $S_3 \subset C[0,1]$ such that for every $f \in S_3$ we have

$$\mu(\{x \in [0,1] : f^+(x) = f^-(x) = +\infty \quad \text{and} \quad f^+(x) = f^-(x) = -\infty\}) = 1.$$

**Proof.** See [48, page 543]. \hfill \Box

Now we let

$$C_0 := S_1 \cap S_2 \cap S_3. \quad (3.3)$$

**Lemma 3.7.8** For every $f \in C_0$ the set $E(f)$ is residual in $(0,1)$.

**Proof.** Let $f$ be a continuous nowhere monotone function of the second species in $[0,1]$. Given a positive integer $n$, as the functions $f(x) + n \cdot x$ and $f(x) - n \cdot x$ are both nowhere monotone in $[0,1]$, it follows from Lemma 3.2.2 that there exists a residual set $G_n \subset [0,1]$ such that for each $x \in G_n$, $f^+(x) = f^-(x) \leq -n < n \leq f^-(x) = f^+(x)$. Then set $G = \bigcap_{n=1}^{\infty} G_n$ is residual in $[0,1]$ and at each $x \in G$ we have $f^+(x) = f^-(x) = -\infty$, $f^+(x) = f^-(x) = +\infty$. \hfill \Box

**Lemma 3.7.9** For every $r \in \mathbb{R}$, if $f \in C_0$ then the set $D_r = \{x \in (0,1) \mid \partial^- f(x) = \{r\}\}$ is dense in $(0,1)$. In particular, every $f \in C_0$ is Dini subdifferentiable at $c$-dense set of points (i.e., its cardinality is $c$ in each subinterval of $[0,1]$).

**Proof.** For every $r$ we will show that

$$D_r := \{x : f^+(x) = r = f^-(x), f^+(x) = +\infty, f^-(x) = -\infty\},$$
is dense in $[0, 1]$. Since $f$ is nowhere monotone of second species, for every $r \in \mathbb{R}$ the function $g : [0, 1] \to \mathbb{R}$ defined by $g(x) := f(x) - r \cdot x$ is continuous and nowhere monotone and so by Lemma 3.2.3 has minima at a set $S$ being everywhere dense in $(0, 1)$. At $x \in S$ we have $f^-(x) \leq r \leq f^+(x)$. Lemma 3.7.6 shows that $[f^-(x), f^-(x)] \cup [f^+(x), f^+(x)] = [-\infty, +\infty]$, whence $f^-(x) = -\infty$, $f^+(x) = +\infty$, and $f^+(x) \leq f^-(x)$. This shows $f^-(x) = r = f^+(x)$, so $S \subseteq D_r$. Fixing an arbitrary nondegenerate subinterval $I \subset [0, 1]$, for every $r \in \mathbb{R}$ we have $D_r \cap I \neq \emptyset$ because $D_r$ is dense. The $\epsilon$-dense result follows from the observation that $D_{r_1} \cap D_{r_2} = \emptyset$ if $r_1 \neq r_2$. \hfill \square

To finish the proof of Theorem 3.7.3, we observe that Lemma 3.7.7 and Lemma 3.7.8 give parts 1), 2), and 3). By Lemma 3.7.9 for every $r \in \mathbb{R}$ we have $r \in \partial_a f(x)$, this means $\mathbb{R} \subset \partial_a f(x) \subset \partial_c f(x) \subset \mathbb{R}$, which is part 4). By Lemma 3.7.6, $f^\circ(x, 1) = +\infty$ and $f^\circ(x, -1) = +\infty$ for every $x \in (0, 1)$, thus $\partial_{mp} f(x) = \mathbb{R}$, but $\partial_f (x)$ is singleton whenever it exists, this gives 5).

\hfill \square

Comments and examples

Katriel has shown that for every lower semicontinuous function $f$ defined on $\mathbb{R}$ the approximate subdifferential and the Clarke subdifferential agree on a $G_\delta$ set of $\mathbb{R}$ [62]. Our result shows that in $C[0, 1]$ the functions which share the same trivial Clarke subdifferential and approximate subdifferential map form a dense $G_\delta$ subset of $C[0, 1]$. There are many results on the integrability of subdifferentials of non-locally Lipschitz functions [73, 78, 90, 77]. Unless one assume stringent conditions on the function or the subdifferential map, one can not recover the function from its subdifferential uniquely up to an additive constant.

In order to study the integration of proximal subdifferentials, Poliquin has introduced a class of “primal lower-nice” functions which can be uniquely determined, up to a constant, by their proximal subdifferentials. If $f$ is primal lower-nice at $x$, then $\partial_p f(x) = \partial_c f(x)$ [73]. If $f \in C_0$ (see Equation (3.3)) we see that $\partial_p f(x)$ is either empty or a singleton, whereas $\partial_c f(x) \equiv \mathbb{R}$, thus each function $f \in C_0$ is not primal lower-nice at any $x \in (0, 1)$.

If $f \in C_0$ then $\partial_- f(x)$ is either a singleton or empty, whereas $\partial_a f(x) = \partial_c f(x) = \mathbb{R}$. Hence each $f \in C_0$ is neither Clarke nor approximate subdifferentially regular at each point in $(0, 1)$. Furthermore, every $f \in C_0$ is not directionally Lipschitz at each $x \in (0, 1)$.

Let $h$ be any $C^1$ function on $\mathbb{R}$. Suppose $\partial_a f = \partial_c f = \partial_{mp} f \equiv \mathbb{R}$. Then $\partial_a (f + h) =$
3.7 The space of continuous functions $C[0, 1]$

$\partial_a f + h', \partial_c (f + h) = \partial_c f + h'$, $\partial_{mp} (f + h) = h' + \partial_{mp} f$, it follows that

$$
\partial_a (f + h) = \partial_{mp} (f + h) = \partial_c (f + h) \equiv \mathbb{R}.
$$

Hence, adding any $C^1$ map to such a function will not change its $\partial_c f$, $\partial_a f$, and $\partial_{mp} f$.

**Example 3.7.10** Let $S$ be a nonempty closed subset of a finite dimensional Banach space $X$ and $x \in S$. Then $d_S$ is regular at $x$ if and only if $S$ is regular at $x$ [15]. If $f \in C_0$ then $epi f$ is not regular at any $(x, f(x))$ for $x \in (0, 1)$. With $S := epi f$, $d_S$ is not regular at any point of its boundary.

**Example 3.7.11** (1) The nowhere differentiable Weierstrass function $W : [0, 1] \to \mathbb{R}$ is defined by $W(x) := \sum_{n=0}^{+\infty} a^n \cos(b^n \pi x)$ where $0 < a < 1$, $b$ is an odd positive integer, and $ab > 1 + 3\pi/2$. Set $E(W) := \{x : W^+(x) = W^-(x) = +\infty, W_+(x) = W_-(x) = -\infty\}$, and

$$
E_{c1} := \{x : W'_+(x) = +\infty, W'_-(x) = -\infty\}, \quad E_{c2} := \{x : W'_+(x) = -\infty, W'_-(x) = +\infty\},
$$

$$
E_1 := \{x : W'_+(x) = +\infty, W'_-(x) = +\infty, W_+(x) = W_-(x) = -\infty\},
$$

$$
E_2 := \{x : W'_+(x) = -\infty, W'_-(x) = +\infty, W_-(x) = -\infty\},
$$

$$
E_3 := \{x : W'_+(x) = +\infty, W'_-(x) = +\infty, W_+(x) = -\infty\},
$$

$$
E_4 := \{x : W'_-(x) = -\infty, W'_+(x) = +\infty, W_+(x) = -\infty\}.
$$

For $r \in \mathbb{R}$, we define $E_{1r} := \{x : W^+(x) = r, W_+(x) = -\infty, W^-(x) = +\infty, W_-(x) = -\infty\}$,

$$
E_{2r} := \{x : W^+(x) = +\infty, W_+(x) = r, W^-(x) = +\infty, W_-(x) = -\infty\},
$$

$$
E_{3r} := \{x : W^+(x) = +\infty, W_+(x) = -\infty, W^-(x) = r, W_-(x) = -\infty\},
$$

$$
E_{4r} := \{x : W^+(x) = +\infty, W_+(x) = -\infty, W^-(x) = +\infty, W_-(x) = r\}.
$$

In [47] Garg has shown that the sets $E(W)$, $E_{ci}$ $(i = 1, 2)$, $E_i$ $(i = 1$ to $4$), and $E_{ir}$ $(i = 1$ to $4, r \in \mathbb{R})$ cover all the points of $(0, 1)$, and that the points of these sets are distributed in the interval in the following manner:

(i) $E(W)$ is residual in $(0, 1)$ with $\mu(E(W)) = 1$.

(ii) $E_{ci}$ $(i = 1, 2)$ are both enumerable and everywhere dense in $(0, 1)$.
(iii) each of the sets $E_i$ ($i = 1$ to 4) and $E_{ir}$ ($i = 1$ to 4, $r \in \mathbb{R}$) is of the first category with measure equal to zero and has the power of the continuum in every subinterval of $(0,1)$.

Then $\partial_W(x) = \mathbb{R} = \partial_p W(x)$ if $x \in E_{c1}$ and $\partial_W(x) = \emptyset = \partial_p W(x)$ if $x \in (0,1) \setminus E_{c1}$, while $\partial_-(\neg W)(x) = \mathbb{R} = \partial_p (\neg W)(x)$ if $x \in E_{c2}$ and $\partial_-(\neg W)(x) = \emptyset = \partial_p (\neg W)(x)$ if $x \in (0,1) \setminus E_{c2}$. This means both $W$ and $\neg W$ are only countably Dini or proximally subdifferentiable on $(0,1)$.

Now equation (1.2) shows $\partial_a W(x) = \mathbb{R} = \partial_a W(x)$ and $\partial_a (\neg W)(x) = \mathbb{R} = \partial_a (\neg W)(x)$ for every $x \in [0,1]$. Since $W'(x) = +\infty$ and $W'(x) = -\infty$ for every $x \in (0,1)$, $\partial_{mp} W(x) = \mathbb{R}$ for each $x \in (0,1)$. Thus $W$ is only subdifferentiably regular on $E_{c1}$ and $\neg W$ is only subdifferentiably regular on $E_{c2}$ and nowhere else. Let $\#$ stand for $-$ or $p$. For every $k > 0$,

$$
\partial^k(kW)(x) = \partial_p kW(x) = \mathbb{R} \text{ if } x \in E_{c1}, \quad \partial^k(kW)(x) = \partial_p kW(x) = \emptyset \text{ if } x \in (0,1) \setminus E_{c1},
$$

$$
\partial^k(\neg kW)(x) = \partial_p (\neg kW)(x) = \emptyset \text{ if } x \in (0,1) \setminus E_{c2}.
$$

But $kW - \neg W = (k - 1)W$ is not constant if $k \neq 1$. This answers negatively the following question:

**Let $f$ and $g$ both be continuous on $(0,1)$. Assume $\partial_x f(x) = \partial_x g(x)$ and $\partial^k (-f)(x) = \partial^k (-g)(x)$ for every $x \in (0,1)$. Is $\neg f - g$ constant on $(0,1)$?**

Let $\partial_m^* W$ denote the minimal convexificator map of $W$.

(i) If $x \in E(W)$ then every closed set, including the empty set, may be chosen as $\partial_m^* W(x)$, thus $\partial_m^* W(x) = \emptyset$.

(ii) If $x \in E_{c1} \cup E_{c2}$ then every nonempty closed set, unbounded from above and below, may be chosen as $\partial_m^* W(x)$, thus there is no minimal convexificator.

(iii) If $x \in E_1 \cup E_3$ then every nonempty closed set unbounded above may be chosen as $\partial_m^* W(x)$, so there is no minimal convexificator. If $x \in E_2 \cup E_4$ then every nonempty closed set unbounded below may be chosen as $\partial_m^* W(x)$, so there is no minimal convexificator.
3.8 Such a pathological behavior is prevalent too!

(iv) Fix $r \in \mathbb{R}$, if $x \in E_{1r} \cup E_{3r}$ then every closed nonempty set with infimum less than or equal to $r$ may be chosen as $\partial^* W(x)$, so $\partial^*_m W(x) = \{ \bar{r} \}$ as long as $\bar{r} \leq r$; if $x \in E_{2r} \cup E_{4r}$ then every nonempty closed set with supremum greater than or equal to $r$ may be chosen as $\partial^* W(x)$, so $\partial^*_m W(x) = \{ \bar{r} \}$ as long as $\bar{r} \geq r$.

Observe that the minimal convexificator on $E_{ir}$ ($i = 1$ to $4$, $r \in \mathbb{R}$) is not unique.

(2) Let $\phi$ be the function on $\mathbb{R}$ defined by $\phi(x) = |x|$ if $|x| \leq 2$ and $\phi(x + 4p) = \phi(x)$ if $x \in \mathbb{R}$ and $p \in \mathbb{Z}$. $\phi$ is in fact the distance function $\phi(x) = d_A(x)$ where $A := \{ 4m | m \in \mathbb{Z} \}$. Setting $f_n(x) := 4^{-n}\phi(4^nx)$, the van der Waerden function is defined by $f(x) := \sum_{n=1}^{\infty} f_n(x)$, and $f$ is continuous and nowhere differentiable [86, pages 174–175], hence nowhere monotone of the second species. Therefore $\partial_d f(x) = \partial_d f(x) = \mathbb{R}$ for every $x \in \mathbb{R}$.

(3) Choosing any nondifferentiable function $f : \mathbb{R} \to \mathbb{R}$ we define $F(x,y) := f(x) + f(y)$. Then $\partial_d F(x,y) = \partial_d f(x) \times \partial_d f(y)$. Since $\partial_d f(x) = \mathbb{R}$ for every $x \in \mathbb{R}$, we have $\partial_d F(x,y) = \mathbb{R}^2 = \partial_d F(x,y)$ for every $(x,y) \in \mathbb{R}^2$.

3.8 Such a pathological behavior is prevalent too!

What happens measure theoretically if we consider the nondifferentiable functions in $C[0,1]$ with supremum norm? The set of nowhere differentiable functions in the metric space $C[0,1]$ forms a set that is co-analytic, that is, the complement of an analytic set, and not Borel, but universally measurable [27].

Definition 3.8.1 A function $f \in C[0,1]$ is $M$-Lipschitz at a point $x \in [0,1]$ if there exists a constant $M$ such that for all $y \in [0,1]$, $|f(y) - f(x)| \leq M|y - x|$. We say $f$ is Lipschitz at $x$ if it is $M$-Lipschitz for some $M$.

Define $A_n := \{ f \in C[0,1] : f \text{ is } n\text{-Lipschitz at some } x \in [0,1] \}$, $A_n$ is closed and nowhere dense. The nowhere Lipschitz functions $A := \cap_{n=1}^{\infty} C[0,1] \setminus A_n$ are a dense $G_\delta$ in $C[0,1]$. Let $g(x) := \sum_{k=1}^{\infty} 1/k^2 \cos 2^k \pi x$, and $h(x) := \sum_{k=1}^{\infty} 1/k^2 \sin 2^k \pi x$. Hunt showed [55]:

Proposition 3.8.2 For all $f \in C[0,1]$, the set $\{(\alpha, \beta) : (\alpha g + \beta h) \in f + \cup_{n=1}^{\infty} A_n \}$ has Lebesgue measure zero.
3.8 Such a pathological behavior is prevalent too!

From this we see that $\bigcup_{n=1}^{\infty} A_n$ is Haar null. Since the set of nowhere differentiable functions $B$ contains $A$, we have $C[0, 1] \setminus B \subset \bigcup_{n=1}^{\infty} A_n$, so $C[0, 1] \setminus B$ is Haar null. One may now say almost every function in $C[0, 1]$ has trivial Clarke and approximate subdifferentials. A self contained arguments, using nowhere monotone functions, come as follows:

If $f$ is not nowhere monotone of the second species on $[0, 1]$, then for some $r$ we have $f(x) + rx$ monotone on some subinterval $I \subset [0, 1]$. Let $r \in \mathbb{R}$ and define $f_r$ by $f_r(x) := f(x) + rx$. Let $I$ be a subinterval of $[0, 1]$. Define

$$A_f := \{ f \in C[0, 1] : \text{there exists a } r \in \mathbb{R} \text{ with } f_r \text{ being nondecreasing on } I \}.$$ 

For each $n \in N$, let $A_n$ denote those functions $f \in C[0, 1]$ for which there exists $r \in [-n, n]$ such that $f_r$ is nondecreasing on $I$. Then $A_f = \bigcup_{n=1}^{\infty} A_n$. We show that for each $n \in N$ the set $A_n$ is closed and $C[0, 1] \setminus A_n$ is dense.

To verify that $A_n$ is closed, let $f_k$ be a sequence of functions in $A_n$ such that $f_k \to f$ uniformly. Then $f \in C[0, 1]$, and we must show that $f \in A_n$. For each $k$, there exists $r_k \in [-n, n]$ such that $f_k(x) + r_kx \geq f_k(y) + r_ky$ if $x \geq y$ and $x, y \in I$. There exists an increasing sequence $k_i$ from $N$ such that $\{r_{k_i}\}$ converges to some $r \in [-n, n]$. Then $f(x) + rx \geq f(y) + ry$. Thus $f \in A_n$, and $A_n$ is closed in $C[0, 1]$. To show that $A_n$ is nowhere dense, we verify that $A_n$ has no interior. Take a continuous nowhere differentiable function $g$ defined on $[0, 1]$. For every $\epsilon > 0$, we claim $f + \epsilon g \notin A_n$ if $f \in A_n$. Suppose $f + \epsilon g \in A_n$, then for some $r_1$ we have $h(x) := f(x) + \epsilon g + r_1 x$ being monotone on $I$. Since $f \in A_n$, there exists another $r_2$ with $f(x) + r_2 x$ being monotone on $I$. But

$$h(x) - r_1 x + r_2 x = (f(x) + r_2 x) + \epsilon g(x),$$

implies $g(x) = [h(x) - (f(x) + r_2x) - r_1 x + r_2 x]/\epsilon$. Hence $g$ is differentiable almost everywhere on $I$, a contradiction. Thus $A_n$ is nowhere dense and closed.

Now we show that $A_n$ is Haar null. Let $g$ be a nowhere differentiable function. Define a Borel probability measure as: $\mu(E) = \lambda\{ t \in [0, 1] : tg \in E \}$. We will verify $\mu(f + A_n) = 0$ for every $f \in C[0, 1]$. In fact, the set $\{ t \in [0, 1] : tg \in A_n + f \}$ is either empty or a singleton. If not, we may find $t_1 \neq t_2$ such that $t_1 g \in f + A_n$ and $t_2 g \in f + A_n$. Then there exists $r_1, r_2 \in [-n, n]$ such that $h_1(x) := t_1 g(x) - f(x) + r_1 x$ and $h_2(x) := t_2 g(x) - f(x) + r_2 x$ are nondecreasing on $I$. It follows that $g(x) = [h_1(x) - h_2(x) - (r_1 - r_2)x]/(t_1 - t_2)$ is differentiable almost everywhere on $I$, a contradiction.
3.9 Typical Lipschitz functions have constant subdifferentials

Since \( A_I = \bigcup_{n=1}^{\infty} A_n \), \( A_I \) is Haar null and a countable union of nowhere dense closed sets. The same is true of the set \( B_I := \{ f \in C[0,1] : -f \in A_I \} \).

Let \( \{ I_k \} \) be all the subintervals of \( [0, 1] \) with rational endpoints. Define \( A := \bigcup_k A_{I_k} \) and \( B := \bigcup_k B_{I_k} \). It follows that each of \( A \) and \( B \) is Haar null and a countable union of nowhere dense closed subsets in \( C[0,1] \). Then \( C[0,1] \setminus (A \cup B) \) is a residual set of type \( G_\delta \) and \( A \cup B \) is Haar null. If \( f \in C[0,1] \setminus (A \cup B) \), then for every \( r \in \mathbb{R} \) the function \( f_r \) is not monotonic at any subinterval of \( [0, 1] \), thus it is nowhere monotonic of the second species. Each nowhere monotonic function of the second species \( f \) has \( \partial_a f = \partial_c f = \mathbb{R} \), and \( \partial_- f \) exists only on a first category set of \( [0, 1] \). Hence, we have proved the following theorem:

**Theorem 3.8.3** The set

\[
D := \{ f \in C[0,1] : \partial_a f = \partial_c f = \mathbb{R} \text{ and } \partial_- f \text{ exists only on a first category set} \}.
\]

is prevalent and residual in \( C[0,1] \).

This follows from that \( C[0,1] \setminus (A \cup B) \subset D \). Nondifferentiable functions constitute a proper subclass of the class of continuous nowhere monotone functions of the second species.

3.9 Typical Lipschitz functions have constant subdifferentials

How should we consider the subdifferentials of Lipschitz functions instead of nowhere monotone functions of the second species? Three spaces come into mind right away:

1. The space of all Lipschitz functions with supremum norm. Because nowhere differentiable functions are uniform limits of polynomials, the space is not complete.

2. The space of all Lipschitz functions with the norm given by

\[
\| f \| := |f(0)| + \sup \{ |f(y) - f(x)| / |y - x| : x, y \in [0, 1], x \neq y \},
\]

is a Banach space [65]. It is too big in the following sense: (i) the differentiable functions are not dense. Under the Lipschitz norm, if \( f_n \to f \), for every \( \varepsilon > 0 \) we have

\[
\partial f_n \subset \partial (f_n - f) + \partial f \subset \varepsilon B + \partial f,
\]

and

\[
\partial f \subset \partial (f_n - f) + \partial f_n \subset \varepsilon B + \partial f_n.
\]

Let \( f \) be the Rockafellar function. Then \( \partial f_n \subset \varepsilon B + [-1,1], \ [-1,1] \subset \varepsilon B + \partial f_n \). If \( f_n \) is differentiable, we may take \( x_0 \) such that \( \partial f_n (x_0) = \{ f'_n(x_0) \} \). Then \( 1 = \varepsilon + f'_n(x_0) \) and
3.9 Typical Lipschitz functions have constant subdifferentials

\[-1 = f'_n(x_0) - \varepsilon.\] If \(\varepsilon < 1\), we obtain a contradiction; (ii) Lipschitz functions with constant subdifferential maps are not dense. To see this, we define \(f(x) = 0\) if \(0 \leq x \leq 1/2\), and \(f(x) = x - 1/2\) if \(1/2 \leq x \leq 1\). If \(f_n \rightarrow f\) in Lipschitz norm and \(\partial_c f_n \equiv [a_n, b_n]\) on \([0, 1]\), then \(\|f_n\| \rightarrow 0\) on \([0, 1/2]\) and \(\|f_n\| \rightarrow 1\) on \([1/2, 1]\), a contradiction.

(3) It is in the space of Lipschitz functions with uniformly Lipschitz constant in the topology of uniform convergence that we show typical functions have constant Clarke and approximate subdifferential map. More precisely, we consider the space

\[\text{Lip}_M := \{f : [0, 1] \rightarrow \mathbb{R} : |f(x) - f(y)| \leq M|x - y| \text{ for all } x, y \in [0, 1]\},\]

with the metric \(\rho(f, g) := \max_{x \in [0, 1]} |f(x) - g(x)|\) for \(f, g \in \text{Lip}_M\). The following lemma may be found in [86, page 165].

**Lemma 3.9.1** Suppose the metric space \(Y\) is complete and that \((f_n)_{n=1}^\infty\) is an equicontinuous sequence in \(C(X, Y)\) that converges at each point of a dense subset \(D\) of the topological space \(X\). Then there is a function \(f \in C(X, Y)\) such that \((f_n)_{n=1}^\infty\) converges to \(f\) uniformly on each compact subset \(K\) of \(X\).

As functions in \(\text{Lip}_M\) are equicontinuous, Lemma 3.9.1 shows in \(\text{Lip}_M\) the topology of pointwise convergence and the topology of uniform convergence are the same.

**Lemma 3.9.2** Let \(E \subset [0, 1]\) be an \(F_\sigma\) set of measure 0. Then there is a residual set \(S \subset \text{Lip}_M\) such that for every \(f \in S\) and \(x \in E\) we have

\[
\limsup_{y \to x} \frac{f(y) - f(x)}{y - x} = M \quad \text{and} \quad \liminf_{y \to x} \frac{f(y) - f(x)}{y - x} = -M.
\]

**Proof.** (1) Let \(E\) be a nonempty closed set of measure zero. Let \(G_k\) be the set of those \(f \in \text{Lip}_M\) for which one can find \(\delta > 0\) with the property that for every \(x \in E\) there is \(y \in [0, 1]\) such that \(\delta < |y - x| < 1/k\) and

\[
\frac{f(y) - f(x)}{y - x} > M - \frac{1}{k} + \delta.
\]

We will show that \(G_k\) is open in \(\text{Lip}_M\). Assume \(f_0 \in G_k\). By definition, for some \(\delta > 0\), for each \(x \in E\), there exists \(1/k > |y - x| > \delta\) such that

\[
\frac{f_0(y) - f_0(x)}{y - x} > M - \frac{1}{k} + \delta.
\]
3.9 Typical Lipschitz functions have constant subdifferentials

For this $y_i$, there exists $0 < \eta_{k_i} < \delta$ and $|y-x| + \eta_{k_i} < 1/k$ such that for each $z \in [x-\eta_{k_i}, x+\eta_{k_i}]$ we have

\[ \frac{f_0(y) - f_0(z)}{y-z} > M - \frac{1}{k} + \delta. \]

Then $\{(x - \eta_{k_i}, x + \eta_{k_i}) : x \in E\}$ covers $E$. By compactness, we may take a finite number of them to cover $E$, say $\{(x_i - \eta_{k_i}, x_i + \eta_{k_i})\}^{m_i}_{i=1}$. Set $\eta := \max\{\eta_{k_i}\}$. For every $x \in E$, there exists $x_i$ with $x \in [x_i - \eta_{k_i}, x_i + \eta_{k_i}]$ such that

\[ \frac{1}{k} > \eta_{k_i} + |y_i - x_i| > |y_i - x_i| + |x - x_i| \geq |y_i - x| \geq |y_i - x_i| - |x_i - x| > \delta - \eta, \]

\[ \frac{f_0(y_i) - f_0(x)}{y_i - x} > M - \frac{1}{k} + \delta. \]

Since $(f_0(y_i) - f_0(z))/|y_i - z|$ is continuous on $[x_i - \eta_{k_i}, x_i + \eta_{k_i}]$, its minimum exists denoted by $m_i > M - 1/k + \delta$. Setting $m := \min\{m_i\}$, we have $m > M - 1/k + \delta$. Now, assuming $\rho(f, f_0) \leq \epsilon$, for $x \in (x_i - \eta_{k_i}, x_i + \eta_{k_i})$ with $y = y_i$ we have $1/k > |y - x| > \delta - \eta$ and

\[ \frac{f(y) - f(x)}{y - x} = \frac{f(y) - f(y_i) + f(y_i) - f(x) + f_0(y) - f_0(x)}{y - x} \]

\[ > \frac{-2\epsilon}{\delta} + \frac{f_0(y) - f_0(x)}{y - x} > \frac{-2\epsilon}{\delta} + m. \]

If $\epsilon$ is sufficiently small, then $-2\epsilon/\delta + m > M - 1/k + \delta > M - 1/k + \delta - \eta$. For this $\epsilon$, we have $B_{r_0}(f_0) \subset G_k$, thus $G_k$ is an open set.

To prove $G := \bigcap_{k=1}^{\infty} G_k$ is a residual subset of $Lip_M$, it suffices to show that it is dense. Whenever $f \in Lip_M$, let $f_j(x) := f(0) + \int_0^x \phi_j(t) dt$, where $\phi_j(t) = \phi(t)$ if $d_E(t) > 1/j$ and $\phi_j(t) = M$ if $d_E(t) \leq 1/j$. Since $E$ is a closed subset of $[0,1]$ with measure 0, $\bigcap_{j=1}^{\infty} E_j = E$ and $E_j \subset E_{j-1}$, we have $\lim_{j \to \infty} \mu(\{t \in [0,1] : d_E(t) \leq 1/j\}) = \mu(E) = 0$. Now

\[ |f_j(x) - f(x)| = \left| \int_0^x \phi_j - \phi(t) dt \right| \leq 2M \mu(\{t \in [0,1] : d_E(t) \leq 1/j\}), \]

shows $f_j$ uniformly converges to $f$. For fixed $j$, $f_j \in G_k$ for every $k$ because if $k < j$, we set $\delta_k = 1/(2j)$; if $k \geq j$, we set $\delta_k = 1/(2k)$. Thus $f_j \in G$, and $G$ is residual in $Lip_M$. Then the set $S := G \cap \{f \in Lip_M : -f \in G\}$ is also residual in $Lip_M$. If $f \in S$ and $x \in E$, then for every $k$ there exists $\delta_k < |y_k - x| < 1/k$ such that $(f(y_k) - f(x))/(y_k - x) > M - 1/k + \delta_k$. Letting $k \to \infty$, together with $f \in Lip_M$, we have $\lim_{y \to x} \sup_{y \to x} (f(y) - f(x))/(y - x) = M$. Applying the same arguments to $-f$, we obtain $\lim_{y \to x} \inf_{y \to x} (f(y) - f(x))/(y - x) = -M$. (2) Let $E = \bigcup_{n=1}^{\infty} E_n$ with $E_n$ being closed sets measure 0. We may apply (1) on each $E_n$ to get a residual set $S_n$. Then $\bigcap_{n=1}^{\infty} S_n$ is the desired residual set. \qed
3.9 Typical Lipschitz functions have constant subdifferentials

Since \( f^+(x, 1) \geq \max \{ f^+(x), f^-(x) \} = M \) and \( f^-(x, -1) \geq \max \{ -f_-(x), -f_+(x) \} = M \), we have \( f^+(x, 1) = f^+(x, -1) = M \) for every \( x \in E \). Then \( \partial^o f(x) = [-M, M] \) at \( x \in E \). Define \( E := \{ r : r \in (0, 1) \cap Q \} \). Then \( E \) is countable and dense in \([0, 1]\), in particular, of measure zero and \( F_\sigma \). Thus \( \partial_c f = \partial_a f = [-M, M] \). We have proved the following:

**Theorem 3.9.3** The typical \( f \in Lip_M \) has the following property:

1. \( \partial_c f = \partial_a f = [-M, M] \) on \([0, 1]\).
2. \( \partial_{mp} f(x) = [-M, M] \) for every \( x \in (0, 1) \cap Q \).

Clearly, the same arguments apply on \( R \). One may also deduce Theorem 3.9.3 via nowhere monotone functions:

**Theorem 3.9.4** In \( Lip_1 \), the set

\[ \{ f : f(x) - r \cdot x \text{ is nowhere monotone on } [0, 1] \text{ for every } |r| < 1 \} \text{ is residual.} \]

**Proof.** Let \( I \) denote an open subinterval of \([a, b]\), and let

\[ A_I := \{ f \in Lip_1 : \text{ there exists some } r \in [-1 + 1/n, 1 - 1/n] \text{ with } f(x) - r \cdot x \]

being nondecreasing on \( I \}. \]

To verify that \( A_I \) is closed, let \( \{f_k\} \) be a sequence of functions in \( A_I \) such that \( f_k \rightarrow f \)
uniformly. Then \( f \in Lip_1 \), and we must show that \( f \in A_I \). For each \( k \in N \), there exists
\( r_k \in [-1 + 1/n, 1 - 1/n] \) such that \( f_k(x) - r_k x \geq f_k(y) - r_k y \) for \( x \geq y \in I \). There exists an
increasing sequence \( \{k_i\} \) from \( N \) such that \( \{r_{k_i}\} \) converges to some \( r \in [-1 + 1/n, 1 - 1/n] \),
then \( f(x) - rx \geq f(y) - ry \) for \( x \geq y \in I \). Then \( f \in A_I \) and \( A_I \) is closed in \( Lip_1 \).

To show that \( A_I \) is nowhere dense, we verify that every ball in \( Lip_1 \) contains points of
\( Lip_1 \setminus A_I \). Let \( B_r(f) \) be an open ball in \( Lip_1 \). If \( f \not\in A_I \), there is nothing to prove, so assume
\( f \in A_I \). Let \( (x_0 - \epsilon, x_0 + \epsilon) \subset I \). We define

\[ \phi_r(t) := \begin{cases} 
-1 & \text{if } t \in (x_0 - \epsilon, x_0], \\
1 & \text{if } t \in (x_0, x_0 + \epsilon], \\
f'(t) & \text{otherwise and provided } f'(t) \text{ exists.} 
\end{cases} \]
3.9 Typical Lipschitz functions have constant subdifferentials

Define \( f_\epsilon(x) := f(0) + \int_0^x \phi_\epsilon(t)dt \). Then \( f_\epsilon \in Lip_1 \) and

\[
|f(x) - f_\epsilon(x)| = \left| \int_0^x f'(t) - \phi_\epsilon(t)dt \right| \leq \int_0^1 |f'(t) - \phi_\epsilon(t)|dt = 4\epsilon.
\]

On \( I \), for every \( r \in [-1 + 1/n, 1 - 1/n] \), the function \( f_\epsilon(x) - rx \) is not nondecreasing on \( I \) because on \((x_0 - \epsilon, x_0)\) the function \( f_r \) has derivative \(-1 - r \leq -1/n\). Thus \( A_I \) is nowhere dense. Now let \( \{I_k\} \) be an enumeration of those open subintervals of \([0, 1]\) having rational endpoints. Set \( A := \bigcup_{k=1}^\infty A_{I_k} \). Then \( A \) is a first category set. Similarly, we show that

\[
B := \{ f \in Lip_1 : f(x) - rx \text{ is nonincreasing on some open subinterval of } [0, 1] \text{ for some } r \in (-1, 1) \},
\]

is of first category in \( Lip_1 \). If \( f \in Lip_1 \setminus (A \cup B) \), then for every \( r \in (-1, 1) \) the function \( f(x) - r \cdot x \) is nowhere monotone on \([0, 1]\). \( \square \)

This naive result shows: a typical \( f \in Lip_1 \) has \( \partial_0 f = \partial_\epsilon f = [-1, 1] \). For every such \( f \), \( \partial_- f \) exists only on a first category set by Lemma 3.2.2. Hence, a typical function in \( Lip_1 \) is only differentiable on a first category set. This generalizes the classical known fact (exercise 7.9.2[27]): There exists a Lipschitz function for which the set of points of differentiability is first category.

Consider \( X := \{ f : |f(x) - f(y)| \leq |x - y| \text{ for } x, y \in [a, b] \text{ and } f \text{ is nondecreasing} \} \), endowed with the supremum metric \( \rho \).

**Theorem 3.9.5** In \((X, \rho)\), the set \( \{ f \in X : \partial_0 f = \partial_\epsilon f = [0, 1] \text{ and } f \text{ is strictly increasing} \} \) is residual.

**Proof.** Fix \( x \in (a, b) \). Consider

\[
G_k := \{ f \in X : \frac{f(x + t_1) - f(x)}{t_1} - 1 > -\frac{1}{k} \text{ and } \frac{f(x + t_2) - f(x)}{t_2} < \frac{1}{k} \text{ for some } 0 < t_1, t_2 < \frac{1}{k} \}
\]

(1) \( G_k \) is open. Let \( f_0 \in G_k \). If \( \epsilon > 0 \) is sufficiently small, for every \( f \in X \) satisfying \( \rho(f, f_0) < \epsilon \), we have

\[
\frac{f(x + t_1) - f(x)}{t_1} - 1 > \frac{-2\epsilon}{t_1} + \frac{f_0(x + t_1) - f_0(x)}{t_1} - 1 > -\frac{1}{k}.
\]
for the same $t_1, t_2$ associated with $f_0$.

(2) $G_k$ is dense. Given $f \in X$ and $\epsilon > 0$. Define $\tilde{f}(x) := f(0) + \int_0^x \phi_\delta(t)dt$ with

$$
\phi_\delta(t) := \begin{cases}
    f'(t) & \text{if } t \not\in [x, x + \delta] \\
    0 & \text{if } t \in (x, x + \delta) \\
    1 & \text{if } t \in (x + \delta, x + \delta),
\end{cases}
$$

where $\min\{\epsilon/2, 1/k\} > \delta > \tilde{\delta} > 0$ such that $\delta^{-1}[\tilde{f}(x + \delta) - \tilde{f}(x)] = 1 - 1/k^2$ and $\tilde{\delta}^{-1}[\tilde{f}(x + \tilde{\delta}) - \tilde{f}(x)] = 0$. Then $\tilde{f}$ is nondecreasing, 1-Lipschitz, $\tilde{f} \in G_k$ and

$$
|f(x) - \tilde{f}(x)| = \left| \int_0^x f'(s) - \phi_\delta(s)ds \right| \leq 2\delta < \epsilon.
$$

Then $G_x := \cap_{k=1}^\infty G_k$ is a dense $G_\delta$ set in $X$. If $f \in G_x$, then $f^+(x) = 1, f^-(x) = 0$, so

$$
1 \geq f^0(x, 1) \geq f^\circ(x, 1) \geq f^+(x) = 1 \text{ and } 0 \leq -f^0(x, -1) \leq -f^\circ(x, -1) \leq f^+(x) = 0,
$$

thus $\partial_{mp}f(x) = \partial_c f(x) = [0, 1]$. Let $\{x_k\}$ be dense in $[a, b]$ and set $G := \cap_{k=1}^\infty G_{x_k}$. Then $G$ is a dense $G_\delta$ in $X$. If $f \in G$ we have $\partial_{mp}f(x_k) = \partial_c f(x_k) = [0, 1]$ for every $x_k$, so $\partial_c f(x) = \partial_a f(x) = [0, 1]$ for each $x \in [a, b]$. Moreover, every $f \in G$ must be strictly increasing, otherwise $f$ would be constant on some subinterval $I$, hence $\partial_c f = \{0\}$ on $I$, a contradiction. \qed 

3.10 Can the pseudo-regular points generate the subdifferential?

One of the open problems in Sciffer’s thesis [84] is: “For a locally Lipschitz function $\phi$ on a separable Banach space, do the pseudo-regular points generate the subdifferential?” The answer is seen to be ‘no’ by using nowhere monotone differentiable functions. Observe that a Gâteaux differentiable function $\phi$ is pseudo-regular at $x$ if and only if $\partial_c\phi$ is a singleton.

The space of bounded derivatives on $[0, 1]$, denoted by $M\Delta'$, with metric $\rho(f, g) := \sup_{x \in [0, 1]} |f(x) - g(x)|$ is complete. Let $M\Delta'_o = \{f \in M\Delta' : f = 0$ on a dense set}. 

**Lemma 3.10.1 (Weil)** The set of functions in $M\Delta'_o$ which are positive on one dense subset of $[0, 1]$ and negative on another dense subset of $[0, 1]$ forms a residual subset of $(M\Delta'_o, \rho)$. 

3.11 A comparison to convex analysis

For the proof of Lemma 3.10.1, see [86]. For \( f \in M\Delta'_\alpha \), we define \( F(x) := \int_0^x f(s)ds \). Then \( F \) is globally Lipschitz and \( F' = f \) on \([0, 1]\). Lemma 3.10.1 shows:

**Proposition 3.10.2** Let \( \Delta_\alpha \) denote the set of differentiable functions \( F \) on \([0, 1]\) such that \( F(0) = 0 \) and \( F' \in M\Delta'_\alpha \). For \( F, G \in \Delta_\alpha \), let \( \rho(F, G) = \sup_{x \in [0, 1]} |F'(x) - G'(x)| \). Then

(i) \((\Delta_\alpha, \rho)\) is a complete metric space.

(ii) If \( F \in \Delta_0 \), then \( 0 \in \partial_d F(x) = \partial_e F(x) \) for every \( x \in [0, 1] \).

(iii) A typical \( F \in \Delta_\alpha \) has a \( \partial_e F \) which is not singleton on a positive measure subset of each nondegenerate subinterval \( I \subseteq [0, 1] \).

Choose \( F \in \Delta_\alpha \) satisfying (iii). As \( 0 \in \partial_d F(x) \) for each \( x \in [0, 1] \), \( F \) is pseudo-regular at \( x \) if and only if \( \partial_e F(x) = \{0\} \). The cuso generated by pseudo-regular points is identically \( \{0\} \). Since \( \partial_e F \neq \{0\}, \partial_d F \) can not be generated by the pseudo-regular points.

3.11 A comparison to convex analysis

For a sequence of convex functions \( \{f_i\} \) defined on \( A \subseteq \mathbb{R}^n \), if \( \sup_i \{f_i(x) : i \in \mathbb{N}\} < M_x \) for every \( x \in A \), then \( \{f_i\} \) are locally equi-Lipschitz, thus \( f_i \) converges to \( f \) on each compact convex subset of \( A \) when \( \{f_i\} \) converges to \( f \) pointwise on \( A \) [79, page 90]. Our typical results may be compared with the following result in convex analysis [79, page 233]:

**Proposition 3.11.1 (Rockafellar)** Let \( f \) be a convex function on \( \mathbb{R}^n \), and let \( A \) be an open convex set on which \( f \) is finite. Let \( f_1, f_2, \ldots \) be a sequence of convex functions finite on \( A \) and converging pointwise to \( f \) on \( A \). Let \( x \in A \), and let \( x_1, x_2, \ldots \) be a sequence of points in \( A \) converging to \( x \). Then, for any \( y \in \mathbb{R}^n \) and any sequence \( y_1, y_2, \ldots \), converging to \( y \), one has \( \lim_{i \to \infty} f'_i(x_i; y_i) \leq f'(x; y) \). Moreover, given any \( \epsilon > 0 \), there exists an index \( i_0 \) such that \( \partial f_i(x_i) \subset \partial f(x) + \epsilon B_{\mathbb{R}^n} \) for all \( i \geq i_0 \).

**Proof.** Given any \( \mu > f'(x; y) \), there exists a \( \lambda > 0 \) such that \( x + \lambda y \in A \) and \( (f(x + \lambda y) - f(x))/\lambda < \mu \). As \( f_i \) converges to \( f \) uniformly on compact sets, \( f_i(x_i + \lambda y_i) \) tends to \( f(x + \lambda y) \) and \( f_i(x_i) \) tends to \( f(x) \). Hence, for all sufficiently large index \( i \), one has

\[
[f_i(x_i + \lambda y_i) - f_i(x_i)]/\lambda < \mu.
\]
3.11 A comparison to convex analysis

By convexity of \( f_i \), \( f'_i(x_i; y_i) \leq [f_i(x_i + \lambda y_i) - f_i(x_i)]/\lambda \), we have \( \limsup_{i \to \infty} f'_i(x_i; y_i) \leq \mu \). This is true for any \( \mu > f'(x; y) \), so the "lim sup" inequality in the theorem is valid. We may conclude in particular (by taking \( y_i = y \) for every \( i \)) that

\[
\limsup_{i \to \infty} f'_i(x_i; y) \leq f'(x; y) \quad \text{for all } y \in \mathbb{R}^n.
\]

The convex functions \( f'_i(x_i; \cdot) \) and \( f'(x; \cdot) \) are the support functions of the non-empty closed bounded convex sets \( \partial f_i(x_i) \) and \( \partial f(x) \), respectively, hence they are finite throughout \( \mathbb{R}^n \). Therefore, given any \( \epsilon > 0 \), there exists an index \( i_0 \) such that \( f'_i(x_i; y) \leq f'(x; y) + \epsilon \) for all \( y \in B_{\mathbb{R}^n} \) and \( i \geq i_0 \). By positive homogeneity we have \( f'_i(x_i; y) \leq f'(x; y) + \epsilon\|y\| \) for all \( y \in \mathbb{R}^n \) and \( i \geq i_0 \). This implies that \( \partial f_i(x_i) \subseteq \partial f(x) + \epsilon B_{\mathbb{R}^n} \) for \( i \geq i_0 \). \( \square \)

Because every \( C^1 \) function is a uniform limit of nondifferentiable functions from \( A \), Theorem 3.7.3 shows that Proposition 3.11.1 fails dramatically for nonconvex continuous functions. In order to pose open questions, we recall:

**Theorem 3.11.2 (Denjoy-Young-Saks)** Let \( f \) be an arbitrary finite function defined on \([a, b]\). Then almost every \( x \in [a, b] \) is in one of the following four sets:

(i) \( A_1 \) on which \( f \) has a finite derivative;

(ii) \( A_2 \) on which \( f^+ = f^- \) (finite), \( f^- = \infty, f^+ = -\infty \);

(iii) \( A_3 \) on which \( f^- = f^+ \) (finite), \( f^+ = \infty, f^- = -\infty \);

(iv) \( A_4 \) on which \( f^- = f^+ = \infty, f_- = f_+ = -\infty \).

**Problem 3.11.3** What is the analogue of the Denjoy-Young-Saks theorem in nonsmooth analysis?

**Problem 3.11.4** Let \( A \) be an open subset of \( \mathbb{R}^n \) with \( n > 1 \). For each continuous or locally Lipschitz function \( f : A \to \mathbb{R} \), is \( \{ x \in A : \partial_\alpha f(x) = \partial_\varepsilon f(x) \} \) residual in \( A \)?
Chapter 4

Prescribed Hölder subdifferentials

In [35] Clarke, Ledyaev and Wolenski construct a $C^1$ function, $f$, on $\mathbb{R}$ such that both $\partial_p f$ and $\partial_p (-f)$ are nonempty only on a set that is small in both the sense of measure and category. Correspondingly, in [3] Benoist shows that for every countable dense set $D \subseteq \mathbb{R}$ there exist infinitely many Lipschitz functions $f$ such that $\partial_p f(x) = (-1, 1)$ if $x \in D$ and $\partial_p f(x) = \emptyset$ if $x \notin D$. Correspondingly, Borwein, Girgensohn and Wang show that a much simpler construction exists when $D$ are dyadic rationals [18]. It is, however, not clear to us what the proximal subdifferentiability properties of $-f$ are in either case. Thus, it is natural to ask: If $\partial_{h,s} f(x) \subset \partial_{h,s} g(x)$ and $\partial_{h,s} (-f)(x) \subset \partial_{h,s} (-g)(x)$ for all $x \in \mathbb{R}$, does $f$ differ from $g$ by a constant? In this chapter we answer this question negatively as a consequence of explicitly constructing Lipschitz functions, $f$, such that both $f$ and $-f$ are $s$-Hölder subdifferentiable only on countable dense sets and nowhere else. On one hand this, in some sense, strengthens Clarke, Ledyaev and Wolenski’s example. On the other hand, this also means that although we are given both $\partial_{h,s} f$ and $\partial_{h,s} (-f)$, we still can not recover $f$ uniquely up to a constant. Our construction method is a modification of Benoist’s [3].

4.1 Domains of Hölder super-sub-differentials may be countable

Theorem 4.1.1 Suppose $S_1$ and $S_2$ are two arbitrary countable dense sets in $\mathbb{R}$ with $S_1 \cap S_2 = \emptyset$. Then there exist two countable sets $D_1 \subset S_2$ and $D_2 \subset S_1$ with $D_1$ and $D_2$ dense
in \( \mathbb{R} \) such that there exists a Lipschitz function \( f : \mathbb{R} \to \mathbb{R} \) satisfying: for every \( s > 0 \)

(i) \( \partial_{hs} f(x) = (-1, 1) \) if \( x \in D_2 \) and \( \partial_{hs} f(x) = \emptyset \) if \( x \in \mathbb{R} \setminus D_2 \).

(ii) \( \partial_{hs} (-f)(x) = (-1, 1) \) if \( x \in D_1 \) and \( \partial_{hs} (-f)(x) = \emptyset \) if \( x \in \mathbb{R} \setminus D_1 \).

In particular, when \( s = 1 \), Theorem 4.1.1 holds for the proximal subdifferential. Choosing \( d_1 \in D_1 \) and \( d_2 \in D_2 \), we may apply Theorem 4.1.1 on the countable dense sets \( D_1 \setminus \{d_1\} \) and \( D_2 \setminus \{d_2\} \) to get a Lipschitz function \( g \) such that for all \( s > 0 \), \( \partial_{hs} g \subset \partial_{hs} f \) and \( \partial_{hs} (-g) \subset \partial_{hs} (-f) \). However the difference of \( f \) and \( g \) is not constant. Hence:

**Corollary 4.1.2** There exists a Lipschitz function \( f \) on \( \mathbb{R} \) satisfying: for all \( s > 0 \),

(i) Both \( f \) and \( -f \) are only countably \( s \)-Hölder subdifferentiable.

(ii) There exist infinitely many Lipschitz functions \( g \) on \( \mathbb{R} \), differing by more than a constant, such that \( \partial_{hs} g \subset \partial_{hs} f \) and \( \partial_{hs} (-g) \subset \partial_{hs} (-f) \).

### 4.2 Basic construction

Let \( \mathbb{N} \) denote the nonnegative integers, \( Z \) the integers, and \( Q \) the rationals of \( \mathbb{R} \). For \( n \in \mathbb{N} \) and \( k \in Z \) we use \( r_{n,k} \) to denote the remainder of \( k \) divided by \( 3^n \). Suppose \( S_1 \) and \( S_2 \) are countable and dense in \( \mathbb{R} \) with \( S_1 \cap S_2 = \emptyset \), we consider two cases:

**Case 1.** \( (S_1 \cup S_2) \cap Z = \emptyset \).

**Case 2.** \( (S_1 \cup S_2) \cap Z \neq \emptyset \). Since \( Z - (S_1 \cup S_2) \) is countable, we choose \( x_0 \notin Z - (S_1 \cup S_2) \) and so \( ((S_1 \cup S_2) + x_0) \cap Z = \emptyset \). Suppose that Theorem 4.1.1 holds in Case 1, that is, there exists a Lipschitz function \( g \) associated with \( S_1 + x_0 \) and \( S_2 + x_0 \) satisfying Theorem 4.1.1, then \( f : \mathbb{R} \to \mathbb{R} \) defined by \( f(x) := g(x + x_0) \) satisfies Theorem 4.1.1.

We now show that Theorem 4.1.1 holds in Case 1. Define \( P^+_0 := \{\frac{1}{2}\} \), \( P^-_0 := \emptyset \) and for \( n \geq 1 \)

\[
P^+_n := \{1 - \frac{1}{n+2}\} + [-\frac{1}{2(n+3)^3}, \frac{1}{2(n+3)^3}],
\]

\[
P^-_n := \{-1 + \frac{1}{n+2}\} + [-\frac{1}{2(n+3)^3}, \frac{1}{2(n+3)^3}].
\]
Lemma 4.2.1 For $n \in \mathbb{N}$, we have $\sup P_n^+ < \inf P_{n+1}^+$ and $\inf P_n^- > \sup P_{n+1}^-$. Moreover, $P_n^+ \subset \left[ \frac{1}{2}, 1 \right]$ and $P_n^- \subset \left[ -1, -\frac{1}{2} \right]$.

Proof. Noting that $P_n^- = -P_n^+$ for $n \geq 1$, we only prove the claim for $P_n^+$. To show $\sup P_n^+ < \inf P_{n+1}^+$ for $n \geq 1$, it is equivalent to show
\[
1 - \frac{1}{n+2} + \frac{1}{2(n+3)^3} < 1 - \frac{1}{n+3} - \frac{1}{2(n+4)^3}.
\]
Letting $N = n + 3$, we have $1 - \frac{1}{N-1} + \frac{1}{2N^3} < 1 - \frac{1}{N} - \frac{1}{2(N+1)^3}$. That is,
\[
\frac{1}{2N^3} + \frac{1}{2(N+1)^3} < \frac{1}{(N-1)N}.
\]
But $\frac{1}{N^3} + \frac{1}{2(N+1)^3} < \frac{1}{N^2}$ and $\frac{1}{(N-1)N} \geq \frac{1}{N^2}$. Noting that $N \geq 3$, we conclude that (4.1) is true. When $n = 0$, we have $\frac{1}{2} \leq \frac{1}{2} + \frac{1}{23^3} < 1 - \frac{1}{3} - \frac{1}{24^3} = \inf P_1^+$.

Now we prove $P_n^+ \subset \left[ \frac{1}{2}, 1 \right]$ by induction. When $n = 0$, $P_0^+ = \{ \frac{1}{2} \}$ we have $P_0^+ \subset \left[ \frac{1}{2}, 1 \right]$. Assume $P_{n-1}^+ \subset \left[ \frac{1}{2}, 1 \right]$ and $n \geq 1$. Since $P_n^+ = [\inf P_n^+, \sup P_n^+]$, we have
\[
\inf P_n^+ > \sup P_{n-1}^+ \geq \frac{1}{2}, \quad \text{and} \quad \sup P_n^+ < \inf P_{n+1}^+ = 1 - \frac{1}{n+3} - \frac{1}{2(n+4)^3} < 1.
\]
That is, $P_n^+ \subset \left[ \frac{1}{2}, 1 \right]$. \hfill \Box

We will construct a sequence of functions $\{f_n\}$ and a sequence of “partitions” $\{a_{n,k} : k \in Z\}$ of $\mathbb{R}$ such that for all $n \in \mathbb{N}$:

(C1) Conditions on $\{a_{n,k} : k \in Z\}$:

(i) For every $k \in Z$, $a_{n,k} < a_{n,k+1}$.

(ii) If $n \geq 1$, $a_{n,0,k} = a_{n-1,k}$.

(iii) If $r_{n,k}$ is odd, then $a_{n,k} \in S_2$; If $r_{n,k}$ is even and not zero, then $a_{n,k} \in S_1$; If $r_{n,k} = 0$, then $a_{n,k} \not\in S_1 \cup S_2$.

(iv) $|a_{n,k+1} - a_{n,k}| \leq \frac{1}{2^n}$.

(C2) Conditions on $\{f_n\}$:

(i) $f_n$ is Lipschitz on $\mathbb{R}$. 

4.2 Basic construction

(ii) \( f_n \) is affine on \([a_{n,k}, a_{n,k+1}]\) with slope \( p_{n,k} \). If \( r_{n,k} \) is odd, \( p_{n,k} \in P_n^- \); if \( r_{n,k} \) is even, \( p_{n,k} \in P_n^+ \).

(iii) For \( x \in \{a_{n,k} : k \in \mathbb{Z}\} \), \( f_m(x) = f_n(x) \) for all \( m \geq n \).

(iv) Case 1: if \( p_{n-1,k} \in P_{n-1}^+ \) and \( n \geq 1 \):

(a) On \([a_{n,3k}, a_{n,3k+1}]\), \( f_n \geq f_{n-1} \).

(b) On \([a_{n,3k+1}, a_{n,3k+2}]\), \( f_n(x) = f_n(a_{n,3k+2}) + p_{n,3k+1}(x-a_{n,3k+2}) \) with \( p_{n,3k+1} \in P_n^- \).

(c) On \([a_{n,3k+2}, a_{n,3k+3}]\), \( f_n \leq f_{n-1} \).

(d) For every \( n \geq 2 \), if \( x \in [a_{n-1,k}, a_{n-1,k+1}] \) we have:

\[
  f_n(x) \geq f_{n-1}(a_{n-1,k}) + \sup P_{n-2}(x-a_{n-1,k}) \quad \text{and} \quad f_n(x) \leq f_{n-1}(a_{n-1,k+1}) + \sup P_{n-2}(x-a_{n-1,k+1}).
\]

Case 2: if \( p_{n-1,k} \in P_{n-1}^- \) and \( n \geq 1 \):

(a) On \([a_{n,3k}, a_{n,3k+1}]\), \( f_n \leq f_{n-1} \).

(b) On \([a_{n,3k+1}, a_{n,3k+2}]\), \( f_n(x) = f_n(a_{n,3k+1}) + p_{n,3k+1}(x-a_{n,3k+1}) \) with \( p_{n,3k+1} \in P_n^+ \).

(c) On \([a_{n,3k+2}, a_{n,3k+3}]\), \( f_n \geq f_{n-1} \).

(d) For every \( n \geq 2 \), if \( x \in [a_{n-1,k}, a_{n-1,k+1}] \) we have:

\[
  f_n(x) \geq f_{n-1}(a_{n-1,k+1}) + \inf P_{n-2}(x-a_{n-1,k+1}) \quad \text{and} \quad f_n(x) \leq f_{n-1}(a_{n-1,k}) + \inf P_{n-2}(x-a_{n-1,k}).
\]

Construction of \( \{f_n\} \) and \( \{a_{n,k}\} \)

Define \( f_0 \) by \( f_0(x) := \frac{1}{2}x \) and \( \{a_{0,k} : a_{0,k} = k, k \in \mathbb{Z}\} \). Clearly \( f_0 \) satisfies C1 and C2.

Suppose \( f_0, \ldots, f_{n-1} \) and \( \{a_{0,k}, \ldots, a_{n-1,k}\} \) satisfying C1 and C2 have been chosen, we proceed to construct \( f_n \) and \( \{a_{n,k}\} \):

**Lemma 4.2.2** \( \{[a_{n-1,k}, a_{n-1,k+1}] : k \in \mathbb{Z}\} \) forms a partition of \( \mathbb{R} \).
4.2 Basic construction

Proof By C1 (i), $a_{n-1,k}$ is strictly increasing with respect to $k$, so $L := \lim_{k \to +\infty} a_{n-1,k}$ and $L' := \lim_{k \to -\infty} a_{n-1,k}$ exist. Since $a_{n-1,3^v-1,k} = a_{0,k} = k$, we have $L = +\infty$ and $L' = -\infty$. □

Now set $a_{n,3k} := a_{n-1,k}$ and $a_{n,3k+3} := a_{n-1,k+1}$. We will specify $a_{n,3k+1}$ and $a_{n,3k+2}$ in the construction later. Suppose $p_{n-1,k} \in D^+_n$, we define $g$ on $[a_{n,3k}, a_{n,3k+3}]$ as follows (see Figure 4.1 Case 1):

![Figure 4.1: The function of g on [a_{n,3k}, a_{n,3k+3}]](image)

(i) If $x \in [a_{n,3k}, a_{n,3k+1}]$, $g(x) := f_{n-1}(a_{n-1,k}) + (1 - \frac{1}{n+2})(x - a_{n-1,k})$. Since $1 - \frac{1}{n+2} \in D^n$, by Lemma 4.2.1 we have $1 - \frac{1}{n+2} > \sup D^n_{n-1}$, so $g \geq f_{n-1}$ on $[a_{n,3k}, a_{n,3k+1}]$.

(ii) If $x \in [a_{n,3k+1}, a_{n,3k+2}]$,

$$g(x) := \frac{1}{2} (f_{n-1}(a_{n-1,k}) + f_{n-1}(a_{n-1,k+1})) + (-1 + \frac{1}{n+2})(x - \frac{1}{2}(a_{n-1,k} + a_{n-1,k+1})).$$

(iii) If $x \in [a_{n,3k+2}, a_{n,3k+3}]$, $g(x) := f_{n-1}(a_{n-1,k+1}) + (1 - \frac{1}{n+2})(x - a_{n-1,k+1})$. For the same reason as in (i), we have $g \leq f_{n-1}$ on $[a_{n,3k+2}, a_{n,3k+3}]$.

Suppose $p_{n-1,k} \in D^n_{n-1}$, we define $g$ on $[a_{n,3k}, a_{n,3k+3}]$ as follows (see Figure 4.1 Case 2):
4.2 Basic construction

(i') If $x \in [a_{n,3k}, a_{n,3k+1}]$, $g(x) := f_{n-1}(a_{n-1,k}) + \left(-1 + \frac{1}{n+2}\right)(x - a_{n-1,k})$. Since $-1 + \frac{1}{n+2} \in P_n^-$, by Lemma 4.2.1 we have $-1 + \frac{1}{n+2} < \inf P_{n-1}^-$, so $g \leq f_{n-1}$ on $[a_{n,3k}, a_{n,3k+1}]$.

(ii') If $x \in [a_{n,3k+1}, a_{n,3k+2}]$,
\[
g(x) := \frac{1}{2} \left(f_{n-1}(a_{n-1,k}) + f_{n-1}(a_{n-1,k+1})\right) + \\
\left(1 - \frac{1}{n+2}\right)(x - \frac{1}{2}(a_{n-1,k} + a_{n-1,k+1})).
\]

(iii') If $x \in [a_{n,3k+2}, a_{n,3k+3}]$, $g(x) := f_{n-1}(a_{n-1,k+1}) + \left(-1 + \frac{1}{n+2}\right)(x - a_{n-1,k+1})$. For the same reason as (i') we have $g \geq f_{n-1}$ on $[a_{n,3k+2}, a_{n,3k+3}]$.

Now we summarize the properties of $g$.

**Proposition 4.2.3** On $[a_{n,3k}, a_{n,3k+3}]$, $g$ satisfies: If $p_{n-1,k} \in P_{n-1}^+$,

(i) On $[a_{n,3k}, a_{n,3k+1}]$, $g \geq f_{n-1}$, and the slope of $g$ belongs to $P_n^+$.

(ii) On $[a_{n,3k+1}, a_{n,3k+2}]$, the slope of $g$ belongs to $P_n^-$.

(iii) On $[a_{n,3k+2}, a_{n,3k+3}]$, $g \leq f_{n-1}$, and the slope of $g$ belongs to $P_n^+$.

If $p_{n-1,k} \in P_{n-1}^-$,

(i') On $[a_{n,3k}, a_{n,3k+1}]$, $g \leq f_{n-1}$, and the slope of $g$ belongs to $P_n^-$.

(ii') On $[a_{n,3k+1}, a_{n,3k+2}]$, the slope of $g$ belongs to $P_n^+$.

(iii') On $[a_{n,3k+2}, a_{n,3k+3}]$, $g \geq f_{n-1}$, and the slope of $g$ belongs to $P_n^-$.

**Proposition 4.2.4** For every $x \in [a_{n,3k}, a_{n,3k+3}]$ and $n \geq 2$, we have:

(i) If $p_{n-1,k} \in P_{n-1}^+$, then
\[
g(x) \geq f_{n-1}(a_{n-1,k}) + \sup P_{n-2}^+(x - a_{n-1,k}), \quad (4.2)
\]
\[
g(x) \leq f_{n-1}(a_{n-1,k+1}) + \sup P_{n-2}^+(x - a_{n-1,k+1}). \quad (4.3)
\]
(ii) If \( p_{n-1,k} \in P_{n-1}^− \), then

\[
g(x) \geq f_{n-1}(a_{n-1,k+1}) + \inf P_{n-2}^−(x - a_{n-1,k+1}),
\]

\[
g(x) \leq f_{n-1}(a_{n-1,k}) + \inf P_{n-2}^−(x - a_{n-1,k}).
\]

**Proof.** We only prove (i). The proof of (ii) is similar. Since \( I_3I_2 = I_3I_4 < I_3I_5 \) from Figure 4.2, we see that in order to show (4.2) it suffices to show \( I_3I_5 < I_3I_1 \). Now

![Figure 4.2: The upper and lower bounds of function g on \([a_{n,3k}, a_{n,3k+3}]\)](image)

\[
I_5I_3 = [(1 - \frac{1}{n+2}) - (1 - \frac{1}{n+1} + \epsilon_{n-1})]K_0I_0 \quad \text{and,}
\]

\[
I_3I_1 = [(1 - \frac{1}{n+1} + \epsilon_{n-1}) - \sup P_{n-2}^+]K_0I_0.
\]

Either \( n = 2 \) or \( n > 2 \), it suffices to show: \( -2\epsilon_{n-1} + \frac{1}{2(n+1)^2} < \frac{1}{n+2} + \frac{1}{n} - \frac{2}{n+1} \). Setting \( N = n + 1 \), we have

\[
-2\epsilon_{n-1} + \frac{1}{2N^2} < \frac{1}{N+1} + \frac{1}{N-1} - \frac{2}{N} = \frac{2}{(N^2 - 1)N}.
\]
4.2 Basic construction

Since \( |\epsilon_{n-1}| \leq \frac{1}{2(n+2)^2} = \frac{1}{2(N+1)^3} \) and \( N \geq 3 \), it follows that

\[
\frac{1}{(N+1)^3} + \frac{1}{2N^3} \leq \frac{3}{2N^3} \leq \frac{2}{N(N^2 - 1)}.
\]

Since \( K_3K_1 \) and \( K_0K_2 \) are parallel, (4.3) is true. \( \square \)

**Proposition 4.2.5** Let \( x_0, x_1, x_2 \) and \( x_3 \) be the \( x \)-coordinate of \( K_0, K_1, K_2 \) and \( K_3 \) respectively. Then

\[
\frac{3x_0 + x_3}{4} < x_1 < \frac{x_0 + x_3}{2}, \text{ and } \frac{x_0 + x_3}{2} < x_2 < \frac{3x_3 + x_0}{4}.
\]

**Proof.** By Lemma 4.2.1, \( \frac{L_5I_6}{I_5K_0} = (1 - \frac{1}{n+1} + \epsilon_{n-1})/(1 - \frac{1}{n+2}) > \frac{1}{2} \), we have \( \frac{I_5I_3}{I_5K_0} < \frac{1}{2} \). Because \( \frac{I_5I_6}{I_5K_0} = \frac{I_5I_6}{I_5K_0} < \frac{1}{2} \), we have

\[
x_1 > \frac{x_0 + x_3}{2} - \frac{1}{2} = \frac{3x_0 + x_3}{4}.
\]

Using the fact that \( \frac{x_0 + x_3}{2} = \frac{x_1 + x_2}{2} \), we have \( x_2 < \frac{3x_3 + x_0}{4} \). \( \square \)

Now we modify \( g \). If \( p_{n-1,k} > 0 \): Since \( S_1 \) and \( S_2 \) are dense and noting that \( P^+_n \) is an interval, we may slightly perturb \( K_1 \) and \( K_2 \) toward \( I_3 \) to get \( K_1' \) and \( K_2' \) such that the line segments connecting \( K_0K_1'K_2'K_3 \) still satisfy Propositions 4.2.3, 4.2.4, and 4.2.5. If \( p_{n-1,k} < 0 \): similar arguments apply. Furthermore, we may and do require that

\[
\begin{cases}
    a_{n,3k+1} \in S_2 \text{ and } a_{n,3k+2} \in S_1 & \text{if } p_{n-1,k} \in P^+_n, \\
    a_{n,3k+1} \in S_1 \text{ and } a_{n,3k+2} \in S_2 & \text{if } p_{n-1,k} \in P^-_n.
\end{cases}
\]

Define the restriction of \( f_n \) on \([a_{n,3k}, a_{n,3k+3}]\) to be the perturbed \( g \), that is, \( K_0K_1'K_2K_3 \).

**Verification of C1 and C2**

C1 (i) and (ii) follow from the construction.

**Proposition 4.2.6** The function \( f_n \) is affine on \([a_{n,k'}, a_{n,k'+1}]\) with slope \( p_{n,k'} \) defined by

\[
p_{n,k'} := \begin{cases}
    \in P^+_n & \text{if } r_{n,k'} \text{ is even}, \\
    \in P^-_n & \text{if } r_{n,k'} \text{ is odd}.
\end{cases}
\] (4.6)
4.2 Basic construction

Proof. By construction, \( f_n \) is clearly affine on \([a_{n,k'}, a_{n,k'+1}]\). We now prove (4.6) by induction. When \( n = 0 \), we have \( r_{0,k'} = 0 \) and \( p_{0,k'} = 1/2 \in P_0^+ \), so the property is true at order \( n = 0 \). For \( k' \) there exist \( k \in Z \) and \( r \in \{0, 1, 2\} \) such that \( k' = 3k + r \), so \( r_{n,k'} = 3r_{n-1,k} + r \). Assume (4.6) is true at order \( n - 1 \) and \( n \geq 1 \).

**Case 1.** If \( r_{n,k'} \) is even, we consider two situations: (1) If \( p_{n-1,k} \in P_{n-1}^+ \). Since the property is true at order \( n - 1 \), \( r_{n-1,k} \) is even, and so \( r \) is even. By Proposition 4.2.3 (i) and (iii), \( p_{n,k'} \in P_n^+ \). (2) If \( p_{n-1,k} \in P_{n-1}^- \), then \( r_{n-1,k} \) is odd, and so \( r \) is odd. By Proposition 4.2.3 (ii'), \( p_{n,k'} \in P_n^+ \).

**Case 2.** If \( r_{n,k'} \) is odd, the proof is similar in which we use Proposition 4.2.3 (i'), (iii'), and (ii).

**Proposition 4.2.7** If \( r_{n,k} \) is odd, then \( a_{n,k} \in S_2 \). If \( r_{n,k} \) is even and not 0, then \( a_{n,k} \in S_1 \). If \( r_{n,k} = 0 \), \( a_{n,k} \notin S_1 \cup S_2 \).

Proof. We prove this by induction. Consider \( a_{n,k'} \). There exists \( k \) and \( r = 0, 1, 2 \) such that \( k' = 3k + r \) and so \( r_{n,k'} = 3r_{n-1,k} + r \). When \( n = 0 \), we have \( r_{0,k'} = 0 \). By assumption \( a_{0,k'} \notin S_1 \cup S_2 \). Assume the property is true at order \( n - 1 \). If \( r = 0 \), then \( r_{n,k'} \) and \( r_{n-1,k} \) have the same parity, \( a_{n,k'} = a_{n,3k} = a_{n-1,k} \) shows that the claim holds by assumption. If \( r \neq 0 \), then \( r_{n,k'} \neq 0 \). Applying Proposition 4.2.6, we consider two cases:

**Case 1.** Assume \( r_{n,k'} \) is even. If \( r_{n-1,k} \) is even, then \( p_{n-1,k} \in P_{n-1}^+ \) and \( r \) is even. By construction \( a_{n,3k+2} \in S_1 \). If \( r_{n-1,k} \) is odd, then \( p_{n-1,k} \in P_{n-1}^- \) and \( r \) is odd. By construction \( a_{n,3k+1} \in S_1 \).

**Case 2.** Assume \( r_{n,k'} \) is odd. If \( r_{n-1,k} \) is even, then \( p_{n-1,k} \in P_{n-1}^+ \) and \( r \) is odd. By construction \( a_{n,3k+1} \in S_2 \). If \( r_{n-1,k} \) is odd, then \( p_{n-1,k} \in P_{n-1}^- \) and \( r = 2 \), the construction shows \( a_{n,3k+2} \in S_2 \).

**Proposition 4.2.8** For every \( n \in \mathbb{N} \) and \( k' \in Z \) we have \( |a_{n,k'} - a_{n,k'+1}| \leq \frac{1}{2^n} \).

Proof. We prove this by induction. When \( n = 0 \), for every \( k' \) we have \( |a_{0,k'} - a_{0,k'+1}| = 1 \leq \frac{1}{2^0} \). Assume \( |a_{n-1,k'} - a_{n-1,k'+1}| \leq \frac{1}{2^n - r} \) for \( n \geq 1 \) and \( k' \in Z \). For \( k' \in Z \) there exists \( k \in Z \) and \( r \in \{0, 1, 2\} \) such that \( k' = 3k + r \). By C1 (i) and (ii), we have \( a_{n,k'}, a_{n,k'+1} \in [a_{n-1,k}, a_{n-1,k+1}] \). By Proposition 4.2.5 we have: If \( r = 0 \), then

\[
|a_{n,k'} - a_{n,k'+1}| \leq |a_{n-1,k} - \frac{a_{n-1,k} + a_{n-1,k+1}}{2}| \leq \frac{1}{2^n}.
\]
4.2 Basic construction

If \( r = 1 \), then \( \frac{3a_{n-1,k} + a_{n-1,k+1}}{4} < a_{n,k'} < a_{n,k'+1} < \frac{3a_{n-1,k+1} + a_{n-1,k}}{4} \). This means:

\[ |a_{n,k'} - a_{n,k'+1}| \leq \frac{1}{2} |a_{n-1,k} - a_{n-1,k+1}| \leq \frac{1}{2^n}. \]

If \( r = 2 \), then \( |a_{n,k'} - a_{n,k'+1}| \leq |a_{n-1,k+1} - a_{n-1,k} + a_{n-1,k+1}| \leq \frac{1}{2^n}. \)

C2 (iii) follows from the construction. C2 (iv) follows from Proposition 4.2.3 and 4.2.4.

Properties of \( f_n \) and \( \{a_{n,k}\} \)

**Lemma 4.2.9** For every \( n \in \mathbb{N} \) and \( x \in \mathbb{R} \), there exists \( k \in \mathbb{Z} \) such that \( x \in [a_{n,k}, a_{n,k+1}] \).

**Proof.** This follows from C1 (i), (iii), and the fact \( \{a_{0,k}\} = \{k : k \in \mathbb{Z}\} \).

**Lemma 4.2.10** \( f_n \) converges uniformly on \( \mathbb{R} \) to a 1-Lipschitz function.

**Proof.** By Lemma 4.2.9 for every \( x \) and \( n \geq 2 \) there exists \( k \in \mathbb{Z} \) such that \( x \in [a_{n-1,k}, a_{n-1,k+1}] \).

Case 1. If \( p_{n-1,k} \in P_{n-1}^+ \), by C2 (iv) Case 1 (d) we have

\[ f_{n-1}(a_{n-1,k}) \leq f_n(x) \leq f_{n-1}(a_{n-1,k+1}), \]

\[ f_{n-1}(a_{n-1,k}) \leq f_n(x) \leq f_{n-1}(a_{n-1,k+1}). \]

Thus by Lemma 4.2.1,

\[ |f_{n-1}(x) - f_n(x)| \leq |f_{n-1}(a_{n-1,k}) - f_{n-1}(a_{n-1,k+1})| \]

\[ = |p_{n-1,k}(a_{n-1,k} - a_{n-1,k+1})| \leq |a_{n-1,k} - a_{n-1,k+1}| \leq \frac{1}{2^n}. \quad (4.7) \]

Case 2. If \( p_{n-1,k} \in P_{n-1}^- \), the proof of equation (4.7) is similar by using C2 (iv) Case 2 (d). But \( \{a_{n-1,k}, a_{n-1,k+1} : k \in \mathbb{Z}\} \) is a partition of \( \mathbb{R} \), so \( \sup_{x \in \mathbb{R}} |f_n(x) - f_{n-1}(x)| \leq \frac{1}{2^n} \). By the Weierstrass M-test, \( \sum_{n=2}^{\infty} (f_n - f_{n-1}) + f_1 \) converges uniformly on \( \mathbb{R} \). Since \( |f_{n-1}(x) - f_{n-1}(y)| \leq |x - y| \) for every \( x, y \in \mathbb{R} \), we have \( |f(x) - f(y)| \leq |x - y| \) for every \( x, y \in \mathbb{R} \).

Properties of \( f \)
4.3 Basic construction

Lemma 4.2.11 For every $n \geq 2$, $k_1, k_2 \in Z$ with $k_1 < k_2$. Assume $\alpha \in [\inf P_{n-1}^-, \sup P_{n-1}^+]$ and $\beta \in R$. Then

(1) If $f_n(x) \geq \alpha x + \beta$ for each $x \in [a_{n,k_1}, a_{n,k_2}]$, then $f(x) \geq \alpha x + \beta$ on $[a_{n,k_1}, a_{n,k_2}]$.

(2) If $f_n(x) \leq \alpha x + \beta$ for each $x \in [a_{n,k_1}, a_{n,k_2}]$, then $f(x) \leq \alpha x + \beta$ on $[a_{n,k_1}, a_{n,k_2}]$.

In particular, setting $\alpha = 0$ we have:

(1') If $f_n \geq \beta$ on $[a_{n,k_1}, a_{n,k_2}]$, then $f \geq \beta$ on $[a_{n,k_1}, a_{n,k_2}]$.

(2') If $f_n \leq \beta$ on $[a_{n,k_1}, a_{n,k_2}]$, then $f \leq \beta$ on $[a_{n,k_1}, a_{n,k_2}]$.

Proof. We prove (2), the proof of (1) is similar. By Lemma 4.2.10, it suffices to prove that for all $q \geq n$ we have $f_q(x) \leq \alpha x + \beta$ for each $x \in [a_{n,k_1}, a_{n,k_2}]$. When $q = n$, by assumption the claim is true. Suppose the claim is true at order $q - 1 \geq n$. That is,

$$f_{q-1}(x) \leq \alpha x + \beta \quad \text{for each } x \in [a_{n,k_1}, a_{n,k_2}]. \quad (4.8)$$

By Cl (ii), $[a_{n,k_1}, a_{n,k_2}] = [a_{q-1,3^{n-1}k_1}, a_{q-1,3^{n-1}k_2}] = \bigcup_{k=3^{n-1}k_1}^{3^{n-1}k_2-1} [a_{q-1,k}, a_{q-1,k+1}]$.

Case 1. Suppose $p_{q-1,k} \in P_{q-1}^+$. By C2 (iv) Case 1 (d) for each $x \in [a_{q-1,k}, a_{q-1,k+1}]$ we have $f_q(x) \leq f_{q-1}(a_{q-1,k+1}) + \sup P_{q-2}^+(x - a_{q-1,k+1})$. Since $q - 2 \geq n - 1$, by Lemma 4.2.1 we have $\sup P_{q-2}^+ \geq \sup P_{n-1}^+ \geq \alpha$. By (4.8),

$$f_q(x) \leq \alpha a_{q-1,k+1} + \beta + \sup P_{q-2}^+(x - a_{q-1,k+1}) \leq \alpha a_{q-1,k+1} + \beta + \alpha (x - a_{q-1,k+1}) = \alpha x + \beta.$$

Case 2. Suppose $p_{q-1,k} \in P_{q-1}^-$. By C2 (iv) Case 2 (d) for each $x \in [a_{q-1,k}, a_{q-1,k+1}]$ we have $f_q(x) \leq f_{q-1}(a_{q-1,k}) + \inf P_{q-2}^-(x - a_{q-1,k})$. Since $q - 2 \geq n - 1$, by Lemma 4.2.1 we have $\inf P_{q-2}^- \leq \inf P_{n-1}^\leq \alpha$. By (4.8),

$$f_q(x) \leq \alpha a_{q-1,k} + \beta + \inf P_{q-2}^-(x - a_{q-1,k}) \leq \alpha a_{q-1,k} + \beta + \alpha (x - a_{q-1,k}) = \alpha x + \beta. \quad \Box$$

Lemma 4.2.12 For every $x \in \{a_{n,k} : k \in Z\}$ and $n \in N$, we have $f(x) = f_n(x)$.

Proof. By C2 (iii), $f_m(x) = f_n(x)$ for every $m \geq n \in N$ and $x \in \{a_{n,k} : k \in Z\}$. By Lemma 4.2.10, $f(x) = f_n(x)$.
4.3 Computing the Hölder super-sub-differentials of $f$

Define $D := \{a_{n,k} : n \in \mathbb{N}, k \in \mathbb{Z}\}$ and $D_0 := \{a_{n,k} : n \in \mathbb{N}, k \in \mathbb{Z}, \text{ and } r_{n,k} = 0\}$,

$$D_1 := \{a_{n,k} : n \in \mathbb{N}, k \in \mathbb{Z}, \text{ and } r_{n,k} \text{ is odd}\},$$

$$D_2 := \{a_{n,k} : n \in \mathbb{N}, k \in \mathbb{Z}, \text{ and } r_{n,k} \text{ is even but not 0}\}.$$  

Clearly $D = \bigcup_{i=0}^{2} D_i$. By Proposition 4.2.7, we know $D_1 \subset S_2$ and $D_2 \subset S_1$.

**Lemma 4.3.1** Both $D_1$ and $D_2$ are dense in $\mathbb{R}$.

**Proof.** Suppose $x \in \mathbb{R}$ and $V$ is an open neighborhood of $x$. By Lemma 4.2.9 for every $n \geq 1$ there exists $k_n$ such that $x \in [a_{n-1,k_n}, a_{n-1,k_n+1}]$. By C1 (iv) for $n$ large we have $[a_{n-1,k_n}, a_{n-1,k_n+1}] \subset V$. By C1 (ii), $[a_{n-1,k_n}, a_{n-1,k_n+1}] = [a_{n,3k_n}, a_{n,3k_n+3}]$. Consider $a_{n,3k_n+r} \in [a_{n,3k_n}, a_{n,3k_n+3}]$ with $r = 1, 2$. Since $r_{n,3k_n+r} = 3r_{n-1,k_n} + r$,

Case 1. If $r_{n-1,k_n}$ is even, then $r_{n,3k_n+1}$ is odd and $r_{n,3k_n+2} \neq 0$ is even, so $a_{n,3k_n+1} \in D_1$ and $a_{n,3k_n+2} \in D_2$.

Case 2. If $r_{n-1,k_n}$ is odd, then $r_{n,3k_n+1} \neq 0$ is even and $r_{n,3k_n+2}$ is odd, so $a_{n,3k_n+1} \in D_2$ and $a_{n,3k_n+2} \in D_1$. 

$$\partial_{hs} f(x) = \emptyset \text{ if } x \notin D$$

Suppose $\xi \in \partial_{hs} f(x)$, then there exists an open neighbourhood $V$ of $x$ and some $\sigma > 0$ such that

$$f(y) \geq f(x) + \xi(y - x) - \sigma|y - x|^{1+s} \text{ for all } y \in V. \quad (4.9)$$

By Lemma 4.2.9, for every $n \geq 1$ there exists $k_n \in \mathbb{Z}$ such that

$$x \in [a_{n-1,k_n}, a_{n-1,k_n+1}] = [a_{n,3k_n}, a_{n,3k_n+3}]. \quad (4.10)$$

By C1 (iv) we have $[a_{n-1,k_n}, a_{n-1,k_n+1}] \subset V$ for $n$ large. There exists $r \in \{0, 1, 2\}$ such that $x \in [a_{n,3k_n+r}, a_{n,3k_n+r+1}]$. Consider the following six situations:

$S1^+_n: p_{n-1,k_n} \in P^+_{n-1}$ and $r = 0, \quad S1^-_n: p_{n-1,k_n} \in P^-_{n-1}$ and $r = 2,$

$S2^+_n: p_{n-1,k_n} \in P^+_{n-1}$ and $r = 1, \quad S2^-_n: p_{n-1,k_n} \in P^-_{n-1}$ and $r = 1,$
4.3 Computing the Hölder super-sub-differentials of $f$

$S3^+_n$: $p_{n-1,k_n} \in P^+_n$ and $r = 2$,  $S3^-_n$: $p_{n-1,k_n} \in P^-_n$ and $r = 0$.

Case 1. $S1^+_n$ holds for infinitely many $n$'s. Without loss of generality we assume $S1^+_n$ holds for all $n \geq 3$. By (4.9) and (4.10), we have

$$f(a_{n-1,k_n}) \geq f(x) + \xi(a_{n-1,k_n} - x) - \sigma|a_{n-1,k_n} - x|^{1+s},$$

$$f(a_{n-1,k_n+1}) \geq f(x) + \xi(a_{n-1,k_n+1} - x) - \sigma|a_{n-1,k_n+1} - x|^{1+s}.$$  

Since $f_{n-1}$ is affine on $[a_{n-1,k_n}, a_{n-1,k_n}+1]$, by Lemma 4.2.12 we have

$$f(a_{n-1,k_n}) = f_{n-1}(a_{n-1,k_n}) = f_{n-1}(x) + p_{n-1,k_n}(a_{n-1,k_n} - x),$$

$$f(a_{n-1,k_n+1}) = f_{n-1}(a_{n-1,k_n+1}) = f_{n-1}(x) + p_{n-1,k_n}(a_{n-1,k_n+1} - x).$$

Then

$$f_{n-1}(x) + p_{n-1,k_n}(a_{n-1,k_n} - x) \geq f(x) + \xi(a_{n-1,k_n} - x) - \sigma|a_{n-1,k_n} - x|^{1+s}, \quad (4.11)$$

$$f_{n-1}(x) + p_{n-1,k_n}(a_{n-1,k_n+1} - x) \geq f(x) + \xi(a_{n-1,k_n+1} - x) - \sigma|a_{n-1,k_n+1} - x|^{1+s}. \quad (4.12)$$

By C2 (iv) Case 1 (a) we have $f_n \geq f_{n-1}$ on $[a_n, 3k_n, a_{n,3k_n+1}]$, so does $f \geq f_{n-1}$ on $[a_n, 3k_n, a_{n,3k_n+1}]$ by Lemma 4.2.11. In particular $f(x) \geq f_{n-1}(x)$. (4.11) and (4.12), together with the fact $x \not\in D$, show that:

$$p_{n-1,k_n} - \xi \leq \sigma|a_{n-1,k_n}|^s; \quad (4.13)$$

$$\xi - p_{n-1,k_n} \leq \sigma|a_{n-1,k_n+1} - x|^s. \quad (4.14)$$

Letting $n \to \infty$, we get $\xi = 1$. By (4.14) and C1 (iv) we have

$$\frac{1}{n+1} - \frac{1}{2(n+2)^3} \leq \sigma\left(\frac{1}{2n+1}\right)^s.$$  

This is a contradiction for $n$ sufficiently large.

Case 1'. $S1^+_n$ holds for infinitely many $n$'s. Without loss of generality we assume $S1^+_n$ holds for all $n \geq 3$. By (4.9) and (4.10), we have

$$f(a_{n-1,k_n}) \geq f(x) + \xi(a_{n-1,k_n} - x) - \sigma|a_{n-1,k_n} - x|^{1+s},$$

$$f(a_{n-1,k_n+1}) \geq f(x) + \xi(a_{n-1,k_n+1} - x) - \sigma|a_{n-1,k_n+1} - x|^{1+s}.$$
4.3 Computing the H"older super-sub-differentials of $f$

Since $f_{n-1}$ is affine on $[a_{n-1,k_n},a_{n-1,k_n+1}]$, by Lemma 4.2.12 we have

$$f(a_{n-1,k_n}) = f_{n-1}(a_{n-1,k_n}) = f_{n-1}(x) + p_{n-1,k_n}(a_{n-1,k_n} - x),$$

$$f(a_{n-1,k_n+1}) = f_{n-1}(a_{n-1,k_n+1}) = f_{n-1}(x) + p_{n-1,k_n}(a_{n-1,k_n+1} - x).$$

Then

$$f_{n-1}(x) + p_{n-1,k_n}(a_{n-1,k_n} - x) \geq f(x) + \xi(a_{n-1,k_n} - x) - \sigma|a_{n-1,k_n} - x|^{1+s}, \quad (4.15)$$

$$f_{n-1}(x) + p_{n-1,k_n}(a_{n-1,k_n+1} - x) \geq f(x) + \xi(a_{n-1,k_n+1} - x) - \sigma|a_{n-1,k_n+1} - x|^{1+s}. \quad (4.16)$$

By C2 (iv) Case 2 (c) we have $f_n \geq f_{n-1}$ on $[a_{n,3k_n+2},a_{n,3k_n+3}]$, so does $f \geq f_{n-1}$ on $[a_{n,3k_n+2},a_{n,3k_n+3}]$ by Lemma 4.2.11. In particular $f(x) \geq f_{n-1}(x)$. (4.15) and (4.16), together with the fact $x \notin D$, show that:

$$\xi - p_{n-1,k_n} \geq -\sigma|x - a_{n-1,k_n}|^s, \quad (4.17)$$

$$\xi - p_{n-1,k_n} \leq \sigma|a_{n-1,k_n+1} - x|^s. \quad (4.18)$$

Letting $n \to \infty$, we get $\xi = -1$. By (4.17) and C1 (iv) we have

$$-\frac{1}{n+1} + \frac{1}{2(n+2)^3} \geq -\sigma\left(\frac{1}{2n-1}\right)^s.$$ 

This is a contradiction for $n$ sufficiently large.

Case 2. $S2_n^+$ holds for infinitely many $n$'s. Without loss of generality we assume that $S2_n^+$ holds for all $n \geq 3$. By (4.9) we have

$$f(a_{n,3k_n}) \geq f(x) + \xi(a_{n,3k_n} - x) - \sigma|x - a_{n,3k_n}|^{1+s}, \quad (4.19)$$

$$f(a_{n,3k_n+2}) \geq f(x) + \xi(a_{n,3k_n+2} - x) - \sigma|x - a_{n,3k_n+2}|^{1+s}. \quad (4.20)$$

By C2 (iv) Case 1 (d) on $[a_{n-1,k_n},a_{n-1,k_n+1}]$, we have

$$f_n \geq f_{n-1}(a_{n-1,k_n}) = f(a_{n-1,k_n}),$$

and so $f \geq f(a_{n-1,k_n}) = f(a_{n,3k_n})$ by Lemma 4.2.11. On $[a_{n,3k_n+1},a_{n,3k_n+2}]$ we have $f_n \geq f_n(a_{n,3k_n+2})$, and so $f \geq f_n(a_{n,3k_n+2}) = f(a_{n,3k_n+2})$ on $[a_{n,3k_n+1},a_{n,3k_n+2}]$ by Lemma 4.2.11. In particular, $f(x) \geq f(a_{n,3k_n})$ and $f(x) \geq f(a_{n,3k_n+2})$. From (4.19) and (4.20) we have

$$-\xi \leq \sigma|x - a_{n,3k_n}|^s \quad \text{and} \quad \xi \leq \sigma|a_{n,3k_n+2} - x|^s.$$
Letting $n \to \infty$, by C1 (iv) we have $\xi = 0$. Since $p_{n,3k_n+1} \in P_n^-$, by Lemma 4.2.1 we have $p_{n,3k_n+1} \leq \inf P_{n-1}^+ < -1 + \frac{1}{n+1}$. By C2 (iv) Case 1 (b) on $[a_{n,3k_n+1}, a_{n,3k_n+2}]$,

\[ f_n(y) = f_n(a_{n,3k_n+1} + p_{n,3k_n+1}(y - a_{n,3k_n+1}) \geq f_n(a_{n,3k_n+1} + (-1 + \frac{1}{n+1})(y - a_{n,3k_n+1}). \]

Then by Lemma 4.2.11 on $[a_{n,3k_n+1}, a_{n,3k_n+2}]$ we have

\[ f(y) \geq f_n(a_{n,3k_n+1}) + (-1 + \frac{1}{n+1})(y - a_{n,3k_n+1}). \]

In particular, $f(x) \geq f_n(a_{n,3k_n+2} + (-1 + \frac{1}{n+1})(x - a_{n,3k_n+1})$. Together with (4.20) we get: \( \sigma a_{n,3k_n+2} - x \|^{1+s} \geq (-1 + \frac{1}{n+1})(x - a_{n,3k_n+1}). \) That is, \( \sigma a_{n,3k_n+2} - x \|^{1+s} \geq (1 - \frac{1}{n+1}). \) Letting $n \to \infty$, we get $0 \geq 1$, which is a contradiction.

**Case 2.** $S2_n^-$ holds for infinitely many $n$’s. Without loss of generality we assume that $S2_n^-$ holds for all $n \geq 3$. By (4.9) we have

\[ f(a_{n,3k_n+3}) \geq f(x) + \xi (a_{n,3k_n+3} - x) - \sigma |x - a_{n,3k_n+3}|^{1+s}, \]

(4.21)

\[ f(a_{n,3k_n+1}) \geq f(x) + \xi (a_{n,3k_n+1} - x) - \sigma |x - a_{n,3k_n+1}|^{1+s}. \]

(4.22)

By C2 (iv) Case 2 (d) on $[a_{n-1,k_n}, a_{n-1,k_n+1}]$, we have

\[ f_n \geq f_{n-1}(a_{n-1,k_n+1}) = f(a_{n-1,k_n+1}), \]

and so $f \geq f(a_{n-1,k_n+1}) = f(a_{n,3k_n+3})$ by Lemma 4.2.11. On $[a_{n,3k_n+1}, a_{n,3k_n+2}]$ we have $f_n \geq f_n(a_{n,3k_n+1})$, and so $f \geq f_n(a_{n,3k_n+1}) = f(a_{n,3k_n+1})$ on $[a_{n,3k_n+1}, a_{n,3k_n+2}]$ by Lemma 4.2.11. In particular, $f(x) \geq f(a_{n,3k_n+3})$ and $f(x) \geq f(a_{n,3k_n+1})$. From (4.21) and (4.22) we have

\[ \xi \leq \sigma |a_{n,3k_n+3} - x|^{s} \quad \text{and} \quad -\xi \leq \sigma |a_{n,3k_n+1} - x|^{s}. \]

Letting $n \to \infty$, by C1 (iv) we have $\xi = 0$. Since $p_{n,3k_n+1} \in P_n^+$, by Lemma 4.2.1 we have $p_{n,3k_n+1} \geq \sup P_{n-1}^+ > 1 - \frac{1}{n+1}$. By C2 (iv) Case 2 (b) on $[a_{n,3k_n+1}, a_{n,3k_n+2}]$,

\[ f_n(y) = f_n(a_{n,3k_n+1}) + p_{n,3k_n+1}(y - a_{n,3k_n+1}) \geq f_n(a_{n,3k_n+1}) + (1 - \frac{1}{n+1})(y - a_{n,3k_n+1}). \]

Then by Lemma 4.2.11 on $[a_{n,3k_n+1}, a_{n,3k_n+2}]$ we have

\[ f(y) \geq f_n(a_{n,3k_n+1}) + (1 - \frac{1}{n+1})(y - a_{n,3k_n+1}). \]
4.3 Computing the Hölder super-sub-differentials of $f$

In particular, $f(x) \geq f_n(a_n, 3k_n + 1) + (1 - \frac{1}{n+1})(x - a_n, 3k_n + 1)$. Together with (4.22) we get:

$$
\sigma |a_n, 3k_n + 1 - x|^{1+s} \geq (1 - \frac{1}{n+1})(x - a_n, 3k_n + 1).
$$

That is, $\sigma |x - a_n, 3k_n + 1|^{s} \geq (1 - \frac{1}{n+1})$. Letting $n \to \infty$, we get $0 \geq 1$, which is a contradiction.

Now Case 1, Case 1', Case 2, and Case 2' together show that there exists $n_0 \geq 3$ such that for all $n \geq n_0$ either $S^+_n$ or $S^-_n$ holds.

Case 3. Suppose at $n_0$ we have $S^+_n$: We will show that $S^+_n$ holds for $n \geq n_0$. More precisely, for $n \geq n_0$,

$$
\begin{align*}
\begin{cases}
 p_{n-1,k_n} \in P^{+}_{n-1}, \\
 k_n = 3^{n-n_0}(k_{n_0} + 1) - 1, \\
 x \in [a_n, 3k_n + 2, a_n, 3k_n + 3],
\end{cases}
\end{align*}
$$

(4.23)

By assumption, this is true for $n = n_0$. Assume this is true at order $n$ and $n \geq n_0$. Since $x \in [a_n, 3k_n + 2, a_n, 3k_n + 3]$, we have $k_{n+1} = 3k_n + 2$, so

$$
k_{n+1} = 3(3^{n-n_0}(k_{n_0} + 1) - 1) + 2 = 3^{n+1-n_0}(k_{n_0} + 1) - 1.
$$

Since $p_{n-1,k_n} \in P^{+}_{n-1}$, by C2 (ii) we know $r_{n-1,k_n}$ is even, and so $r_{n,k_n+1} = 3r_{n-1,k_n} + 2$ is even. By C2 (ii) we have

$$
p_{n,k_n+1} \in P^{+}_n.
$$

(4.24)

By (4.23) and C1 (ii) we have $x \in [a_n, k_n+1, a_n, k_n+1] = [a_n, 3k_n + 1, a_n, 3k_n + 1 + 3]$. But for $n \geq n_0$ we are only only in situation $S^+_n$ or $S^-_n$, by (4.24) we know

$$
x \in [a_n, 3k_n+1 + 2, a_n, 3k_n+1 + 3].
$$

Thus (4.23) is true at order $n + 1$. It follows that for $n \geq n_0$,

$$
x \in [a_n, 3k_n+2, a_n, 3k_n+3] = [a_n, 3k_n+2, a_n, 3k_n+1 + 3] = [a_n, 3k_n+2, a_n, 3k_n+1 + 3].
$$

By C1 (iv), we have $|a_n, 3k_n+2 - a_n, 3k_n+3| \leq \frac{1}{2^{k_n+1}}$. Letting $n \to \infty$, we get $x = a_{n_0-1,k_{n_0}+1}$, which is a contradiction since $x \notin D$.

Case 3'. Suppose at $n_0$ we have $S^-_n$: We will show that $S^-_n$ holds for $n \geq n_0$. More precisely, for $n \geq n_0$,

$$
\begin{align*}
\begin{cases}
p_{n-1,k_n} \in P^{-}_{n-1}, \\
k_n = 3^{n-n_0}k_{n_0}, \\
x \in [a_n, 3k_n, a_n, 3k_n+1],
\end{cases}
\end{align*}
$$

(4.25)
4.3 Computing the Hölder super-sub-differentials of $f$

By assumption, this is true for $n = n_0$. Assume this is true at order $n$ and $n \geq n_0$. Since $x \in [a_{n,3k_n}, a_{n,3k_n+1}]$, we have $k_{n+1} = 3k_n$, so $k_{n+1} = 3(3^{n-n_0}k_{n_0}) = 3^{n+1-n_0}k_{n_0}$. Since $p_{n-1,k_n} \in P_{n-1}^-$, by C2 (ii) we also know $r_{n-1,k_n}$ is odd, and so $r_{n,k_n+1} = 3r_{n-1,k_n}$ is odd. By C2 (ii) we have

$$p_{n,k_n+1} \in P_n^-.$$  \hspace{1cm} (4.26)

By (4.25) and C1 (ii) we have $x \in [a_{n,k_n+1}, a_{n,k_n+1+1}] = [a_{n+1,3k_n+1}, a_{n+1,3k_n+1+3}]$. But for $n \geq n_0$ we are in situation $S3^n_+$ or $S3^n_-$, by (4.26) we know $x \in [a_{n+1,3k_n+1}, a_{n+1,3k_n+1+1}]$. Thus (4.25) is true at order $n + 1$. It follows that for $n \geq n_0$,

$$x \in [a_{n,3k_n}, a_{n,3k_n+1}] = [a_{n,3^{n+1-n_0}k_{n_0}}, a_{n,3k_n+1}].$$

By C1 (iv), we have $|a_{n,3k_n} - a_{n,3k_n+1}| \leq \frac{1}{2^n}$. Letting $n \to \infty$, we get $x = a_{n_0-1,k_0}$, which is a contradiction since $x \not \in D$.

**What is $\partial_{hs} f(x)$ if $x \in D$?**

In this case there exists $m \in \mathbf{N}$ and $k \in \mathbb{Z}$ such that $x = a_{m,k}$. By C1 (ii) we have $x = a_{n,3^{n-m}k}$ for $n \geq m$, and this also shows that we may assume $m \geq 2$. Let $y_1 := a_{n,3^{n-m}k-1}, y_2 := a_{n,3^{n-m}k+1}, p_1 := p_{n,3^{n-m}k-1}$, and $p_2 := p_{n,3^{n-m}k}$. Then $x \in [y_1, y_2]$ and $|y_1 - y_2| \leq |y_1 - x| + |x - y_2| \leq \frac{1}{2^{n-1}}$. Suppose $\xi \in \partial_{hs} f(x)$, by (4.9) there exists $\sigma > 0$ and an open neighborhood $V$ of $x$ such that $f(y) \geq f(x) + \xi(y-x) - \sigma|y-x|^{1+s}$ for all $y \in V$. Now for $n$ large we have $[y_1, y_2] \subset V$. Then for $i = 1, 2$,

$$f(y_i) \geq f(x) + \xi(y_i-x) - \sigma|y_i-x|^{1+s}. \hspace{1cm} (4.27)$$

By C2 (ii) and (iii),

$$f(y_i) = f_n(y_i) = f_n(x) + p_i(y_i-x) = f(x) + p_i(y_i-x). \hspace{1cm} (4.28)$$

(4.27) and (4.28) show that

$$\xi - p_1 \geq -\sigma|y_1-x|^s, \hspace{1cm} (4.29)$$

$$\xi - p_2 \leq \sigma|y_2-x|^s. \hspace{1cm} (4.30)$$

**Case 1.** If $x \in D_0$, we know $r_{m,k} = 0$. Then $r_{n,3^{n-m}k} = 0$ and $r_{n,3^{n-m}k-1} = 3^n - 1$, so $p_1, p_2 \in P_n^+$. Letting $n \to \infty$, we get $\xi = 1$. By (4.30), $\frac{1}{n+2} - \frac{1}{2(n+3)^s} \leq \sigma(\frac{1}{2^n})^s$, and this is absurd if $n$ is large. This means $\partial_{hs} f(x) = \emptyset$ if $x \in D_0$. 
4.3 Computing the Hölder super-sub-differentials of $f$

Case 2. If $x \in D_1$, we know $r_{m,k}$ is odd. Then $r_{n,3n^m-m_k} = 3^{n-m}r_{m,k}$ is odd and $r_{n,3n^m-m_k-1} = 3^{n-m}r_{m,k} - 1$ is even. Hence $p_1 \in P_n^+$ and $p_2 \in P_n^-$. Letting $n \to \infty$, we get $\xi \geq 1$ and $\xi \leq -1$ which is equally absurd. This means $\partial_{hs} f(x) = \emptyset$ if $x \in D_1$.

Case 3. If $x \in D_2$, we know $r_{m,k}$ is even not zero. Then $r_{n,3n^m-m_k} = 3^{n-m}r_{m,k}$ is even and $r_{n,3n^m-m_k-1} = 3^{n-m}r_{m,k} - 1$ is odd, so $p_1 \in P_n^-$ and $p_2 \in P_n^+$. Consider $f$ on $[y_1, y_2]$. Letting $n \to \infty$, it follows from (4.29) and (4.30) that $-1 \leq \xi \leq 1$. If $\xi = 1$ we have from (4.30) that $\frac{1}{n+2} - \frac{1}{2(n+3)^2} \leq \sigma(\frac{1}{2^r})^s$, and this is absurd if $n$ is large. If $\xi = -1$, from (4.29) we have $\frac{1}{n+2} + \frac{1}{2(n+3)^2} \geq -\sigma(\frac{1}{2^r})^s$ which is also absurd if $n$ is large. Hence $\partial_{hs} f(x) \subset (-1, 1)$.

We proceed to show $(-1, 1) \subset \partial_{hs} f(x)$. For each $y \in [a_{n,3n^m-m_k-1}, a_{n,3n^m-m_k}]$ by C2 (iv) Case 2 (d), we have $f_{n+1}(y) \geq f_n(a_{n,3n^m-m_k}) + \inf P_{n-1}^- (y - a_{n,3n^m-m_k})$. Since $\inf P_{n-1}^- \in [\inf P_n^-, \sup P_n^+]$, Lemma 4.2.11 shows that

$$f(y) \geq f_n(a_{n,3n^m-m_k}) + \inf P_{n-1}^- (y - a_{n,3n^m-m_k}) \geq f(x) + (-1 + \frac{1}{n+1})(y - x),$$

for all $y \in [a_{n,3n^m-m_k-1}, a_{n,3n^m-m_k}]$. For each $y \in [a_{n,3n^m-m_k}, a_{n,3n^m-m_k+1}]$ by C2 (iv) Case 1 (d), we have $f_{n+1}(y) \geq f_n(a_{n,3n^m-m_k}) + \sup P_{n-1}^+ (y - a_{n,3n^m-m_k})$. Since $\sup P_{n+1}^+ \in [\inf P_n^-, \sup P_n^+]$, Lemma 4.2.11 also shows that

$$f(y) \geq f_n(a_{n,3n^m-m_k}) + \sup P_{n-1}^+ (y - a_{n,3n^m-m_k}) \geq f(x) + (1 - \frac{1}{n+1})(y - x),$$

for every $y \in [a_{n,3n^m-m_k}, a_{n,3n^m-m_k+1}]$. Now (4.31) and (4.32) show that for every fixed $n \geq m$ on $[a_{n,3n^m-m_k-1}, a_{n,3n^m-m_k+1}]$ we have

$$f(y) \geq f(x) + t(y - x) \quad \text{for every } t \in [-1 + \frac{1}{n+1}, 1 - \frac{1}{n+1}].$$

This shows $[-1 + \frac{1}{n+1}, 1 - \frac{1}{n+1}] \subset \partial_{hs} f(x)$ for every $n \geq m$. Hence

$$(-1, 1) = \bigcup_{n=m} \left[ -1 + \frac{1}{n+1}, 1 - \frac{1}{n+1} \right] \subset \partial_{hs} f(x).$$

$$\partial_{hs} (-f)(x) = \emptyset \text{ if } x \not\in D$$

Suppose $\xi \in \partial_{hs} (-f)(x)$. Then there exists an open neighborhood $V$ of $x$ and some $\sigma > 0$ such that

$$-f(y) \geq -f(x) + \xi(y - x) - \sigma|y - x|^{1+s} \quad \text{for all } y \in V.$$

By Lemma 4.2.9, for every $n \geq 1$ there exists $k_n \in Z$ such that

$$x \in [a_{n-1,k_n}, a_{n-1,k_n+1} = [a_{n,3k_n}, a_{n,3k_n+3}[.$$
4.3 Computing the Hölder super-sub-differentials of $f$

By C1 (iv) we have $[a_{n-1,k_n}, a_{n-1,k_n+1}] \subset V$ for $n$ large. There exists $r \in \{0,1,2\}$ such that $x \in]a_{n,3k_n+r}, a_{n,3k_n+r+1}[$. Consider the following six situations:

$$S1^+_n: \ p_{n-1,k_n} \in P^+_{n-1} \text{ and } r = 2, \quad S1^-_n: \ p_{n-1,k_n} \in P^-_{n-1} \text{ and } r = 0,$$

$$S2^+_n: \ p_{n-1,k_n} \in P^+_{n-1} \text{ and } r = 1, \quad S2^-_n: \ p_{n-1,k_n} \in P^-_{n-1} \text{ and } r = 1,$$

$$S3^+_n: \ p_{n-1,k_n} \in P^+_{n-1} \text{ and } r = 0, \quad S3^-_n: \ p_{n-1,k_n} \in P^-_{n-1} \text{ and } r = 2.$$

Case 1. $S1^+_n$ holds for infinitely many $n$'s. Without loss of generality we assume $S1^+_n$ holds for all $n \geq 3$. By (4.34) and (4.35), we have

$$-f(a_{n-1,k_n}) \geq -f(x) + \xi(a_{n-1,k_n} - x) - \sigma|a_{n-1,k_n} - x|^{1+s},$$

$$-f(a_{n-1,k_n+1}) \geq -f(x) + \xi(a_{n-1,k_n+1} - x) - \sigma|a_{n-1,k_n+1} - x|^{1+s}.$$

Since $f_{n-1}$ is affine on $[a_{n-1,k_n}, a_{n-1,k_n+1}]$, by Lemma 4.2.12 we have

$$f(a_{n-1,k_n}) = f_{n-1}(a_{n-1,k_n}) = f_{n-1}(x) + p_{n-1,k_n} (a_{n-1,k_n} - x),$$

$$f(a_{n-1,k_n+1}) = f_{n-1}(a_{n-1,k_n+1}) = f_{n-1}(x) + p_{n-1,k_n} (a_{n-1,k_n+1} - x).$$

Then

$$-f_{n-1}(x) - p_{n-1,k_n} (a_{n-1,k_n} - x) \geq -f(x) + \xi(a_{n-1,k_n} - x) - \sigma|a_{n-1,k_n} - x|^{1+s}, \quad (4.36)$$

$$-f_{n-1}(x) - p_{n-1,k_n} (a_{n-1,k_n+1} - x) \geq -f(x) + \xi(a_{n-1,k_n+1} - x) - \sigma|a_{n-1,k_n+1} - x|^{1+s}. \quad (4.37)$$

By C2 (iv) Case 1 (c) we have $f_n \leq f_{n-1}$ on $[a_{n,3k_n+2}, a_{n,3k_n+3}]$, so does $f \leq f_{n-1}$ on $[a_{n,3k_n+2}, a_{n,3k_n+3}]$ by Lemma 4.2.11. In particular $f(x) \leq f_{n-1}(x)$. (4.36) and (4.37), together with the fact $x \notin D$, show that:

$$-\xi - p_{n-1,k_n} \leq \sigma|a_{n-1,k_n} - x|^s, \quad (4.38)$$

$$\xi + p_{n-1,k_n} \leq \sigma|a_{n-1,k_n+1} - x|^s. \quad (4.39)$$

Letting $n \to \infty$, we get $\xi = -1$. By (4.38) and C1 (iv) we have

$$\frac{1}{n+1} - \frac{1}{2(n+2)^3} \leq \sigma\left(\frac{1}{2n+1}\right)^s.$$

This is a contradiction for $n$ sufficiently large.
4.3 Computing the Hölder super-sub-differentials of $f$

Case 1’. $S1^-_n$ holds for infinitely many $n$’s. Without loss of generality we assume $S1^-_n$ holds for all $n \geq 3$. By (4.34) and (4.35), we have

$$-f(a_{n-1,k_n}) \geq -f(x) + \xi(a_{n-1,k_n} - x) - \sigma|a_{n-1,k_n} - x|^{1+s},$$

$$-f(a_{n-1,k_n+1}) \geq -f(x) + \xi(a_{n-1,k_n+1} - x) - \sigma|a_{n-1,k_n+1} - x|^{1+s}.$$  (4.40)

Since $f_{n-1}$ is affine on $[a_{n-1,k_n}, a_{n-1,k_n+1}]$, by Lemma 4.2.12 we have

$$f(a_{n-1,k_n}) = f_{n-1}(a_{n-1,k_n}) = f_{n-1}(x) + p_{n-1,k_n}(a_{n-1,k_n} - x),$$

$$f(a_{n-1,k_n+1}) = f_{n-1}(a_{n-1,k_n+1}) = f_{n-1}(x) + p_{n-1,k_n}(a_{n-1,k_n+1} - x).$$

Then

$$-f_{n-1}(x) - p_{n-1,k_n}(a_{n-1,k_n} - x) \geq -f(x) + \xi(a_{n-1,k_n} - x) - \sigma|a_{n-1,k_n} - x|^{1+s},$$  (4.41)

By C2 (iv) Case 2 (a) we have $f_n \leq f_{n-1}$ on $[a_{n,3k_n}, a_{n,3k_n+1}]$, so does $f \leq f_{n-1}$ on $[a_{n,3k_n}, a_{n,3k_n+1}]$ by Lemma 4.2.11. In particular $f(x) \leq f_{n-1}(x)$. Now (4.40) and (4.41), together with the fact $x \not\in D$, show that:

$$-\xi - p_{n-1,k_n} \leq \sigma|a_{n-1,k_n} - x|^s,$$  (4.42)

$$\xi + p_{n-1,k_n} \leq \sigma|a_{n-1,k_n+1} - x|^s.$$  (4.43)

Letting $n \to \infty$, we get $\xi = 1$. By (4.43) and C1 (iv) we have

$$\frac{1}{n+1} - \frac{1}{2(n+2)^3} \leq \sigma\left(\frac{1}{2n-1}\right)^s.$$  (4.43)

This is a contradiction for $n$ sufficiently large.

Case 2. $S2^+_n$ holds for infinitely many $n$’s. Without loss of generality we assume that $S2^+_n$ holds for all $n \geq 3$. By (4.34) we have

$$-f(a_{n,3k_n+3}) \geq -f(x) + \xi(a_{n,3k_n+3} - x) - \sigma|x - a_{n,3k_n+3}|^{1+s},$$  (4.44)

$$-f(a_{n,3k_n+1}) \geq -f(x) + \xi(a_{n,3k_n+1} - x) - \sigma|x - a_{n,3k_n+1}|^{1+s}.$$  (4.45)

By C2 (iv) Case 1 (d) on $[a_{n-1,k_n}, a_{n-1,k_n+1}]$, we have

$$f_n \leq f_{n-1}(a_{n-1,k_n+1}) = f(a_{n-1,k_n+1}),$$
and so \( f \leq f(a_{n-1,k_n+1}) = f(a_{n,3k_n+3}) \) by Lemma 4.2.11. On \([a_{n,3k_n+1}, a_{n,3k_n+2}]\) we have \( f_n \leq f_n(a_{n,3k_n+1}) \), and so \( f \leq f_n(a_{n,3k_n+1}) = f(a_{n,3k_n+1}) \) by Lemma 4.2.11 on \([a_{n,3k_n+1}, a_{n,3k_n+2}]\). In particular, \( f(x) \leq f(a_{n,3k_n+3}) \) and \( f(x) \leq f(a_{n,3k_n+1}) \). From (4.44) and (4.45) we have

\[
\xi \leq \sigma |a_{n,3k_n+3} - x|^{s} \quad \text{and} \quad -\xi \leq \sigma |a_{n,3k_n+1} - x|^{s}.
\]

Letting \( n \to \infty \), by C1 (iv) we have \( \xi = 0 \). Since \( p_{n,3k_n+1} \in P_n^- \), by Lemma 4.2.1 we have \( p_{n,3k_n+1} \leq \inf P_n^- = -1 + \frac{1}{n+1} \). By C2 (iv) Case 1 (b) on \([a_{n,3k_n+1}, a_{n,3k_n+2}]\),

\[
f_n(y) = f_n(a_{n,3k_n+1}) + p_{n,3k_n+1}(y - a_{n,3k_n+1}) \\
\leq f_n(a_{n,3k_n+1}) + (-1 + \frac{1}{n+1})(y - a_{n,3k_n+1}).
\]

Then by Lemma 4.2.11 on \([a_{n,3k_n+1}, a_{n,3k_n+2}]\) we have

\[
f(y) \leq f_n(a_{n,3k_n+1}) + (-1 + \frac{1}{n+1})(y - a_{n,3k_n+1}).
\]

In particular, \( f(x) \leq f_n(a_{n,3k_n+1}) + (-1 + \frac{1}{n+1})(x - a_{n,3k_n+1}) \). Together with (4.45) we get:

\[
\sigma |a_{n,3k_n+1} - x|^{1+s} \geq (1 - \frac{1}{n+1})(x - a_{n,3k_n+1}).
\]

That is, \( \sigma |a_{n,3k_n+1} - x|^{s} \geq (1 - \frac{1}{n+1}) \). Letting \( n \to \infty \), we get \( \sigma \geq 1 \), which is a contradiction.

**Case 2'**. \( S_{2n}^- \) holds for infinitely many \( n \)'s. Without loss of generality we assume that \( S_{2n}^- \) holds for all \( n \geq 3 \). By (4.34) we have

\[
-f(a_{n,3k_n}) \geq -f(x) + \xi(a_{n,3k_n} - x) - \sigma|x - a_{n,3k_n}|^{1+s}, \tag{4.46}
\]

\[
-f(a_{n,3k_n+2}) \geq -f(x) + \xi(a_{n,3k_n+2} - x) - \sigma|x - a_{n,3k_n+2}|^{1+s}. \tag{4.47}
\]

By C2 (iv) Case 2 (d) on \([a_{n-1,k_n}, a_{n-1,k_n+1}]\), we have

\[
f_n \leq f_{n-1}(a_{n-1,k_n}) = f(a_{n-1,k_n}),
\]

and so \( f \leq f(a_{n-1,k_n}) = f(a_{n,3k_n}) \) by Lemma 4.2.11. On \([a_{n,3k_n+1}, a_{n,3k_n+2}]\) we have \( f_n \leq f_n(a_{n,3k_n+2}) \), and so \( f \leq f_n(a_{n,3k_n+2}) = f(a_{n,3k_n+2}) \) on \([a_{n,3k_n+1}, a_{n,3k_n+2}]\) by Lemma 4.2.11. In particular, \( f(x) \leq f(a_{n,3k_n}) \) and \( f(x) \leq f(a_{n,3k_n+2}) \). From (4.46) and (4.47) we have

\[
-\xi \leq \sigma |a_{n,3k_n} - x|^{s} \quad \text{and} \quad \xi \leq \sigma |a_{n,3k_n+2} - x|^{s}.
\]
4.3 Computing the Hölder super-sub-differentials of $f$

Letting $n \to \infty$, by C1 (iv) we have $\xi = 0$. Since $p_{n,3k_n+1} \in P_n^+$, by Lemma 4.2.1 we have $p_{n,3k_n+1} \geq \sup P_{n-1}^+ > 1 - \frac{1}{n+1}$. By C2 (iv) Case 2 (b) on $[a_{n,3k_n+1}, a_{n,3k_n+2}]$, 

$$f_n(y) = f_n(a_{n,3k_n+2}) + p_{n,3k_n+1}(y - a_{n,3k_n+2})$$

$$\leq f_n(a_{n,3k_n+2}) + (1 - \frac{1}{n+1})(y - a_{n,3k_n+2}).$$

Then by Lemma 4.2.11 on $[a_{n,3k_n+1}, a_{n,3k_n+2}]$ we have 

$$f(y) \leq f_n(a_{n,3k_n+2}) + (1 - \frac{1}{n+1})(y - a_{n,3k_n+2}).$$

In particular, $f(x) \leq f_n(a_{n,3k_n+2}) + (1 - \frac{1}{n+1})(x - a_{n,3k_n+2})$. Together with (4.47) we get: 

$$\sigma|a_{n,3k_n+2} - x|^{1+s} \geq (1 - \frac{1}{n+1})(a_{n,3k_n+2} - x).$$

That is, $\sigma|a_{n,3k_n+2} - x|^s \geq (1 - \frac{1}{n+1})$. Letting $n \to \infty$, we get $0 \geq 1$, which is a contradiction.

Now again Case 1, Case 1', Case 2, and Case 2' together show that there exists $n_0 \geq 3$ such that for all $n \geq n_0$ either $S3^+_n$ or $S3^-_n$ holds.

Case 3. Suppose at $n_0$ we have $S3^+_n$. We will show that $S3^+_n$ holds for $n \geq n_0$. More precisely, for $n \geq n_0$,

$$\begin{cases}
  p_{n-1,k_n} \in P_{n-1}^+,
  
  k_n = 3^{n-n_0}k_{n_0},
  
  x \in [a_{n,3k_n}, a_{n,3k_n+1}],
\end{cases} \quad (4.48)$$

By assumption, this is true for $n = n_0$. Assume this is true at order $n$ and $n \geq n_0$. Since $x \in [a_{n,3k_n}, a_{n,3k_n+1}]$, we have $k_{n+1} = 3k_n$, so $k_{n+1} = 3(3^{n-n_0}k_{n_0}) = 3^{n+1-n_0}k_{n_0}$. Since $p_{n-1,k_n} \in P_{n-1}^+$, by C2 (ii) we know $r_{n-1,k_n}$ is even, and so $r_{n,k_{n+1}} = 3r_{n-1,k_n}$ is even. By C2 (ii) we have 

$$p_{n,k_n+1} \in P_n^+ \quad (4.49)$$

By (4.48) and C1 (ii) we have $x \in [a_{n,k_{n+1}}, a_{n,k_{n+1}+1}] = [a_{n+1,3k_{n+1}+1}, a_{n+1,3k_{n+1}+3}]$. But for $n \geq n_0$ we are only in situation $S3^+_n$ or $S3^-_n$, by (4.49) we know 

$$x \in [a_{n+1,3k_{n+1}+1}, a_{n+1,3k_{n+1}+1+1}].$$

Thus (4.48) is true at order $n + 1$. It follows that for $n \geq n_0$,

$$x \in [a_{n,3k_n}, a_{n,3k_n+1}] = [a_{n,3^{n+1-n_0}k_{n_0}}, a_{n,3k_{n+1}}] = [a_{n-1,k_{n_0}}, a_{n,3k_{n+1}}].$$

By C1 (iv), we have $|a_{n,3k_n} - a_{n,3k_n+1}| \leq \frac{1}{2^n}$. Letting $n \to \infty$, we get $x = a_{n_0-1,k_{n_0}}$, which is a contradiction since $x \notin D$.
4.3 Computing the Hölder super-sub-differentials of $f$

Case 3'. Suppose at $n_0$ we have $S3^-_{n_0}$: We will show that $S3^-_n$ holds for $n \geq n_0$. More precisely, for $n \geq n_0$,

\[
\begin{align*}
& \quad p_{n-1,k_n} \in P_{n-1}^-, \\
& k_n = 3^{n-n_0}(k_{n_0} + 1) - 1, \\
& x \in [a_n,3k_n+2, a_n,3k_n+3].
\end{align*}
\] (4.50)

By assumption, this is true for $n = n_0$. Assume this is true at order $n$ and $n \geq n_0$. Since $x \in [a_n,3k_n+2,a_n,3k_n+3]$, we have $k_{n+1} = 3k_n + 2$, so

\[
k_{n+1} = 3(3^{n-n_0}(k_{n_0} + 1) - 1) + 2 = 3^{n+1-n_0}(k_{n_0} + 1) - 1.
\]

Since $p_{n-1,k_n} \in P_{n-1}^-$, by C2 (ii) we know $r_{n-1,k_n}$ is odd, and so $r_{n,k_{n+1}} = 3r_{n-1,k_n} + 2$ is odd. By C2 (ii) we have

\[
p_{n,k_{n+1}} \in P_n^-.
\] (4.51)

By (4.50) and C1 (ii) we have $x \in [a_{n,k_{n+1}}, a_{n,k_{n+1}+1}] = [a_{n+1,3k_{n+1}}, a_{n+1,3k_{n+1}+3}]$. But for $n \geq n_0$ we are only in situation $S3^+_n$ or $S3^-_n$, by (4.51) we have

\[
x \in [a_{n+1,3k_{n+1}+2}, a_{n+1,3k_{n+1}+3}].
\]

Thus (4.50) is true at order $n+1$. It follows that for $n \geq n_0$,

\[
x \in [a_{n,3k_n+2}, 3k_n+3] = [a_n,3k_n+2, a_n,3^{n+1-n_0}(k_{n_0}+1)] = [a_n,3k_n+2, a_{n_0-k_{n_0}+1}].
\]

By C1 (iv), we have $|a_{n,3k_n+2} - a_{n,3k_n+3}| \leq \frac{1}{2^{n+1}}$. Letting $n \to \infty$, we get $x = a_{n_0-k_{n_0}+1}$, which is a contradiction since $x \notin D$.

What is $\partial_{hs}(-f)(x)$ if $x \in D$?

In this case there exists $m \in \mathbb{N}$ and $k \in \mathbb{Z}$ such that $x = a_{m,k}$. By C1 (ii) we have $x = a_{n,3^{n-m}k}$ for $n \geq m$, and this also shows that we may assume $m \geq 2$. Let $y_1 := a_{n,3^{n-m}k-1}$, $y_2 := a_{n,3^{n-m}k+1}$, $p_1 := p_{n,3^{n-m}k-1}$, and $p_2 := p_{n,3^{n-m}k}$. Then $x \in [y_1,y_2]$ and $|y_1-y_2| \leq |y_1-x|+|x-y_2| \leq \frac{1}{2^{n+1}}$. Suppose $\xi \in \partial_{hs}(-f)(x)$. By (4.34) there exists $\sigma > 0$ and an open neighborhood $V$ of $x$ such that $-f(y) \geq -f(x) + \xi(y-x) - \sigma|y-x|^{1+s}$ for all $y \in V$. Now for $n$ large we have $[y_1,y_2] \subset V$. Then for $i = 1, 2$,

\[
-f(y_i) \geq -f(x) + \xi(y_i-x) - \sigma|y_i-x|^{1+s}.
\] (4.52)
4.3 Computing the H"older super-sub-differentials of $f$

By C2 (ii) and (iii),

$$f(y_i) = f_n(y_i) = f_n(x) + p_i(y_i - x) = f(x) + p_i(y_i - x). \quad (4.53)$$

Equations (4.52) and (4.53) show that

$$-(\xi + p_1) \leq \sigma|y_i - x|^s, \quad (4.54)$$

$$\xi + p_2 \leq \sigma|y_2 - x|^s. \quad (4.55)$$

**Case 1.** If $x \in D_0$, we know $r_{m,k} = 0$. Then $r_{n,3^m-m_k} = 0$ and $r_{n,3^m-m_k-1} = 3^n - 1$, so $p_1, p_2 \in P_n^+$. Letting $n \to \infty$, we get $\xi = -1$. By (4.54), $\frac{1}{n+2} - \frac{1}{2(n+3)^{3s}} \leq \sigma \left(\frac{1}{2^n}\right)^s$, and this is absurd if $n$ is large. This means $\partial_{hs}(-f)(x) = \emptyset$ if $x \in D_0$.

**Case 2.** If $x \in D_2$, we know $r_{m,k}$ is even and not 0. Then $r_{n,3^m-m_k} = 3^n-m_r_{m,k}$ is even and $r_{n,3^m-m_k-1} = 3^n-m_r_{m,k} - 1$ is odd. Hence $p_1 \in P_n^-$ and $p_2 \in P_n^+$. Letting $n \to \infty$, we get $\xi \geq 1$ and $\xi \leq -1$ which is absurd. This means $\partial_{hs}(-f)(x) = \emptyset$ if $x \in D_2$.

**Case 3.** If $x \in D_1$, we know $r_{m,k}$ is odd. Then $r_{n,3^m-m_k} = 3^n-m_r_{m,k}$ is odd and $r_{n,3^m-m_k-1} = 3^n-m_r_{m,k} - 1$ is even, so $p_1 \in P_n^+$ and $p_2 \in P_n^-$. Consider $f$ on $[y_1, y_2]$. Letting $n \to \infty$, it follows from (4.54) and (4.55) that $-1 \leq \xi \leq 1$. If $\xi = 1$ we have from (4.55) that $\frac{1}{n+2} - \frac{1}{2(n+3)^{3s}} \leq \sigma \left(\frac{1}{2^n}\right)^s$, and this is false if $n$ is large. If $\xi = -1$, from (4.54) we have $\frac{1}{n+2} - \frac{1}{2(n+3)^{3s}} \leq \sigma \left(\frac{1}{2^n}\right)^s$ which is also false if $n$ is large. Hence $\partial_{hs}(-f)(x) \subset (-1, 1)$.

We proceed to show $(-1, 1) \subset \partial_{hs}(-f)(x)$. By C2 (iv) Case 1 (d),

$$f_{n+1}(y) \leq f_n(a_{n,3^m-m_k}) + \sup P_{n-1}^+(y - a_{n,3^m-m_k}),$$

for each $y \in [a_{n,3^m-m_k-1}, a_{n,3^m-m_k}]$. Since $\sup P_{n-1}^+ \in [\inf P_n^-, \sup P_n^+]$, Lemma 4.2.11 shows that for every $y \in [a_{n,3^m-m_k-1}, a_{n,3^m-m_k}]$ we have

$$f(y) \leq f_n(a_{n,3^m-m_k}) + \sup P_{n-1}^+(y - a_{n,3^m-m_k}) \leq f(x) + (1 - \frac{1}{n+1}) (y - x). \quad (4.56)$$

For each $y \in [a_{n,3^m-m_k}, a_{n,3^m-m_k+1}]$ by C2 (iv) Case 2 (d),

$$f_{n+1}(y) \leq f_n(a_{n,3^m-m_k}) + \inf P_{n-1}^-(y - a_{n,3^m-m_k}).$$

Since $\inf P_{n-1}^- \in [\inf P_n^-, \sup P_n^+]$, by Lemma 4.2.11 for every $y \in [a_{n,3^m-m_k}, a_{n,3^m-m_k+1}]$ we have

$$f(y) \leq f_n(a_{n,3^m-m_k}) + \inf P_{n-1}^-(y - a_{n,3^m-m_k}) \leq f(x) + (-1 + \frac{1}{n+1}) (y - x). \quad (4.57)$$
4.3 Computing the Hölder super-sub-differentials of $f$

(4.56) and (4.57) show that for every fixed $n \geq m$ on $[a_{n,3^n-m_k-1}, a_{n,3^n-m_k+1}]$ we have

$$-f(y) \geq -f(x) + t(y-x) \quad \text{for every } t \in [-1 + \frac{1}{n+1}, 1 - \frac{1}{n+1}],$$

(4.58)

This shows $[-1 + \frac{1}{n+1}, 1 - \frac{1}{n+1}] \subset \partial_{hs}(-f)(x)$ for every $n \geq m$. Hence

$$(-1,1) = \bigcup_{n=m}^{\infty} [-1 + \frac{1}{n+1}, 1 - \frac{1}{n+1}] \subset \partial_{hs}(-f)(x).$$

This concludes the proof of Theorem 4.1.1. \hfill $\Box$

**Remark 4.3.2** (1). Equation (4.33) shows that $f$ achieves a strict local minimum at $a_{m,k} \in D_2$. If $f$ has a local minimum at $x$, then $0 \in \partial_{hs}f(x)$. Since $\partial_{hs}f(x) = \emptyset$ for $x \notin D_2$, we see that $f$ achieves and can only achieve local minima at points of $D_2$. Similar arguments apply to $-f$ by equation (4.58). Hence the points where $f$ achieves strict local minima and the points where $f$ achieves strict local maxima are both dense in $\mathbb{R}$.

(2). It is shown in [19] that in $\mathbb{R}$ for a locally Lipschitz function $f$, its $\partial_a f$ may be computed from $\partial_c f$. Theorem 4.1.1 shows that in general one can not recover $\partial_p f$ from $\partial_c f$.

(3). Suppose $S_1 := Q \setminus Z, S_2 := S_1 + \sqrt{2}$. Then $(S_1 \cup S_2) \cap Z = \emptyset$ and $S_i + 1 = S_i$ for $i = 1, 2$. According to the construction, we can make $f(x+k) = f(x) + f(k), D_1 + k = D_1,$ and $D_2 + k = D_2$ for every $k \in Z$ and $x \in [0,1]$. Consider $g : \mathbb{R} \to \mathbb{R}$ defined by $g(x) := f(x) - x/2$. It follows that $g(x+k) = f(x+k) - (x+k)/2 = f(x) - x/2 = g(x)$ when $x \in [0,1]$. Thus $g$ is a 1-periodic function satisfying:

$$\partial_{hs}g(x) = \begin{cases} 
( -3/2, 1/2) & \text{if } x \in D_2, \\
\emptyset & \text{if } x \notin D_2.
\end{cases}$$

and $\partial_{hs}(-g)(x) = \begin{cases} 
( -1/2, 3/2) & \text{if } x \in D_1, \\
\emptyset & \text{if } x \notin D_1.
\end{cases}$

**Problem 4.3.3** Can one use the Baire category theorem to simplify our proof?
Chapter 5

Subdifferentials in separable Banach spaces

In this chapter, we give a systematic and unified approach to subdifferential integration on separable Banach spaces, that is, Banach spaces which contain a countable dense set. We will frequently use the following fact due to Day whose proof may be found [49, page 191]: Each separable Banach space is smoothable. It will be clear that there are strong links with Chapter 3, whose developments are in fact the prototype from which the results of this one are derived. Among the questions considered in this chapter are: when is a set-valued map a Clarke subdifferential? Is the approximate subdifferential small? What are the subdifferentiability properties of bounded continuous functions and cone-monotone functions?

Our central result illustrates the theme that “for each locally Lipschitz function $f$ on an open set $A$ there exists a residual set in some complete metric space such that every function in the residual set has at least the same subdifferential as $f$.” It turns out that the convex hull of a finite number of Clarke subdifferentials is still a Clarke subdifferential. In the space of nonexpansive functions endowed with supremum metric, the set of functions whose Clarke subdifferential and approximate subdifferential are the same, in fact identically equal to the dual unit ball, is residual. In the space of bounded continuous functions endowed with supremum metric, the set of functions whose Clarke subdifferential and approximate subdifferential are identically equal to the whole dual space is residual. The results allow
us to generalize many known results.

Although there is much literature on subdifferential integration [19, 21, 61], all of their techniques are ad hoc. The nature of the functions, so far constructed, tends to suggest that such functions are exceptions rather than the rule. Ironically, as Weierstrass first constructed a complicated nowhere differentiable function explicitly, the same thing happens in nonsmooth analysis.

5.1 Properties of \( T \)-Lipschitz functions

Let \( T \) be \( w^* \)-cusco that maps from a non-empty open subset \( A \) of a Banach space \( X \) (not necessarily separable) into its dual space \( X^* \). We consider the following (possibly empty) set of locally Lipschitz functions defined on \( A \), called the \( T \)-Lipschitz functions on \( A \);

\[
\mathcal{X}_T := \{ f \in \mathbb{R}^A : f \text{ is locally Lipschitz and } \partial_c f(x) \subseteq T(x) \text{ for all } x \in A \}.
\]

When \( X \) is smoothable, by Theorem 1.1.6 we have the following simplified definition:

\[
\mathcal{X}_T = \{ f \in \mathbb{R}^A : f \text{ is locally Lipschitz and } \nabla f(x) \in T(x) \text{ whenever } \nabla f(x) \text{ exists} \}.
\]

On \( \mathcal{X}_T \), we define a metric \( \rho \) by: \( \rho(f, g) := \min\{1, d(f, g)\} \), where \( d(f, g) := \sup_{x \in A} |f(x) - g(x)| \).

If \( T \) is identically equal to some non-empty \( w^*- \)compact, convex subset \( C \) of \( X^* \), we simply write \( \mathcal{X}_C \) in place of \( \mathcal{X}_T \).

**Definition 5.1.1** Let \( X \) be a Banach space and \( A \subset X \), let \( \mathcal{F} := \{ f_i : i \in I \} \) be a collection of functions \( X \to (-\infty, +\infty] \). \( \mathcal{F} \) is called **locally equi-Lipschitz** relative to \( A \) if each \( f_i \in \mathcal{F} \) is real-valued on \( A \) and if for every \( x \in A \) there exists a neighbourhood \( U \) of \( x \) and \( K > 0 \) such that \( |f_i(x) - f_i(y)| \leq K \|x - y\| \) for all \( x, y \in U \cap A \) and all \( i \in I \).

**Lemma 5.1.2** Let \( X \) be a Banach space and let \( A \) be a metric space, then each \( w^*-\usco \) \( T : A \to 2^{X^*} \) is locally bounded on \( A \).

**Proof.** Assume that that \( T \) is not locally bounded on \( A \). Then there exists a point \( x_0 \in A \) such that for each \( n \in \mathbf{N} \) the set \( T(B_{1/n}(x_0)) \cap nB_{X^*} \neq \emptyset \). Using this we may construct two sequences \( (x_n : n \in \mathbf{N}) \) in \( A \) and \( (x_n^* : n \in \mathbf{N}) \) in \( X^* \) so that \( (x_n : n \in \mathbf{N}) \) converges to \( x_0 \).
5.1 Properties of $T$-Lipschitz functions

and $x_n^* \in T(x_n) \setminus nB_X$. for all $n \in \mathbb{N}$. Now if we set $K := \{x_0, x_1, x_2, \ldots x_n, \ldots\}$ then $K$ is compact and so $T(K)$ is $w^*$-compact. By the uniform boundedness theorem [23] $T(K)$ is bounded, which is impossible since $\{x_n^* : n \in \mathbb{N}\} \subseteq T(K)$ is unbounded. Hence, $T$ must be locally bounded on $A$. 

Combining Proposition 2.4.4 and Lemma 5.1.2, we obtain:

**Proposition 5.1.3** Let $\mathcal{F}$ be a family of real-valued locally Lipschitz functions defined on a non-empty open subset $A$ of a Banach space $X$. Then the functions in $\mathcal{F}$ are $T$-Lipschitz for some $w^*$-cuso $T : A \to 2^{X^*}$ if and only if the family of functions $\mathcal{F}$ is locally equi-Lipschitz on $A$.

**Lemma 5.1.4** $\mathcal{X}_T$ is convex and closed under lattice operations. In other words, if $f, g \in \mathcal{X}_T$, then $\max\{f, g\}, \min\{f, g\} \in \mathcal{X}_T$ and $\lambda f + (1 - \lambda)g \in \mathcal{X}_T$ for every $0 \leq \lambda \leq 1$.

**Proof.** By Propositions 2.3.12, 2.3.3 [33], we have $\partial_c[\max\{f, g\}] \subset \text{co}\{\partial_c f, \partial_c g\} \subset T$, 

$$\partial_c[\min\{f, g\}] = \partial_c[-\max\{-f, -g\}] = -\partial_c[\max\{-f, -g\}] \subset -\text{co}[-T, -T] = T,$$

and for $0 \leq \lambda \leq 1$, $\partial_c(\lambda f + (1 - \lambda)g) \subset \lambda \partial_c f + (1 - \lambda)\partial_c g \subset \lambda T + (1 - \lambda)T = T$. 

**Lemma 5.1.5** If $T : A \to X^*$ is a norm-to-$w^*$ usco, then for each $v \in X$ the function $\sigma_T(v) := \sup_{x^* \in T(x^*)} x^*, v \in \mathcal{X}_T$ is upper semicontinuous for $y$.

**Proof.** For every $\epsilon > 0$, we define a $w^*$-neighbourhood of $Tx$ by

$$W := \{x^* \in X^* : \langle v, x^* \rangle < \sigma_T(x) + \epsilon\}.$$

By $w^*$-upper semicontinuity of $T$ at $x$, there exists $\delta > 0$ such that for $\|y - x\| < \delta$ we have $Ty \subset W$, thus $\sigma_T(y) \leq \sigma_T(x) + \epsilon$ for $\|y - x\| < \delta$. 

**Proposition 5.1.6** If $T$ is a norm-to-$w^*$ cuso, then $(\mathcal{X}_T, \rho)$ is a complete metric space.

**Proof.** Note that a sequence in $(\mathcal{X}_T, \rho)$ converges if and only if the sequence converges uniformly on $A$. Assume $\{f_n\}$ is Cauchy in $\mathcal{X}_T$. For each $x \in A$, $\{f_n(x)\}$ is Cauchy, it converges to $f(x)$. For every $\epsilon > 0$ there exists $N$ such that when $n, m \geq N$ we have 

$$|f_n(x) - f_m(x)| < \epsilon$$

for all $x \in A$. 

5.2 When is a set-valued map a Clarke subdifferential?

Letting \( m \to \infty \), we obtain \( r_1(f_n, f) \leq \epsilon \), thus \( f_n \to f \) uniformly. We now show \( f \in \mathcal{X}_T \).

Lemma 5.1.2 shows \( \{f_n\} \) are locally equi-Lipschitz, then \( f \) is locally Lipschitz on \( A \). Since \( f_n \in \mathcal{X}_T \), for every \( x, y \) with \([x, y] \subset A\), by the Lebourg mean value theorem [33] we have

\[
f_n(x) - f_n(y) \leq \sigma_{\partial_c f_n}(u_n)(x - y) \leq \sigma_{T u_n}(x - y),
\]

where \( u_n \in (x, y) \) and \( u_n \) relies on \( f_n \). As \( u_n \in [x, y] \), a compact interval, there exists a subsequence such that \( u_{n_k} \to u \in [x, y] \) in norm. Taking the limit with respect to \( n_k \), by Lemma 5.1.5 we have \( f(x) - f(y) \leq \limsup_{n_k \to \infty} \sigma_{T u_n}(x - y) \leq \sigma_{T u}(x - y) \). Then for every \( y \in A \) and \( t > 0 \) such that \([y, y + tv] \subset A\) we have \( f(y + tv) - f(y) \leq \sigma_{T u}(tv) \) where \( u \in [y, y + tv] \). That is, \((f(y + tv) - f(y))t^{-1} \leq \sigma_{T u}(v)\). Letting \( y \to x \), \( t \downarrow 0 \), by Lemma 5.1.5 we have

\[
f^0(x; v) = \limsup_{t \downarrow 0} \frac{f(y + tv) - f(y)}{t} \leq \limsup_{t \downarrow 0} \sigma_{T u}(v) \leq \sigma_{T_x}(v).
\]

Dually \( \partial_c f \subset T \). Hence \( f \in \mathcal{X}_T \). \( \square \)

5.2 When is a set-valued map a Clarke subdifferential?

**Lemma 5.2.1** If \( f \) is upper semicontinuous on a separable Banach space \( X \), then there exists a countable dense set \( D \subset X \) such that \( f(x) = \limsup_{y \to x} f(y) \) for every \( x \in X \).

**Proof.** Since \( X \) is separable, we may take a countable family \( \{U_i\} \) of norm open sets such that for every open set \( O \subset X \) we have \( O = \bigcup_{U_i \subset O} U_i \). For each \( U_n \), we let \( C_n \subset U_n \) be a countable set such that \( \sup_{y \in U_n} f(y) = \sup_{y \in C_n} f(y) \). Define \( C := \bigcup_{n=1}^{\infty} C_n \). For every \( x \), with \( O_x \) denoting arbitrary open neighbourhoods containing \( x \), we have

\[
f(x) = \inf_{O_x} \sup_{y \in O_x} f(y) = \inf_{O_x} \sup_{y \in \bigcup_{U_i \subset O_x} U_i} f(y) = \inf_{O_x} \sup_{y \in C \subset O_x} f(y) = \limsup_{y \to x} f(y). \qed \]

**Theorem 5.2.2** Let \( T \) be a norm-to-\( w^* \) cusclo. Assume \( f_j \in \mathcal{X}_F \) for \( j = 1, \ldots, +\infty \). Define \( \Omega, T_F : A \to X^* \) by

\[
\Omega(x) := \bigcup_{j=1}^{\infty} \{\partial_c f_j(x)\} \text{ for } x \in A, \text{ and } T_F := CSC(\Omega). \quad (5.1)
\]

Then in \((\mathcal{X}_T, \rho)\) the set \( \{f \in \mathcal{X}_T : \partial_c f \subset T_F \text{ on } A\} \) is residual. In particular, in \((\mathcal{X}_F, \rho)\), generically \( \partial_c f \equiv T_F \text{ on } A \).
5.2 When is a set-valued map a Clarke subdifferential?

Proof. Fix \( x \in A \) and \( v \in X \). For each \( j \), we consider \( G_k := \)
\[
\left\{ f \in \mathcal{X}_T : \frac{f(y + tv) - f(y)}{t} - f_j^0(x; v) > -\frac{1}{k} \text{ for some } 0 < t < \frac{1}{k} \text{ and some } \|y - x\| < \frac{1}{k} \right\}.
\]

(a) \( G_k \) is open in \( \mathcal{X}_T \). Let \( f_0 \in G_k \). Then for some \( 0 < t < 1/k \) and some \( \|y - x\| < 1/k \) we have
\[
\frac{f_0(y + tv) - f_0(y)}{t} - f_j^0(x; v) > -1/k.
\] (5.2)
Let \( \rho(f, f_0) < \varepsilon < 1 \) and \( f \in \mathcal{X}_T \). Consider
\[
\frac{f(y + tv) - f(y)}{t} - f_j^0(x; v)
= \frac{(f(y + tv) - f(y)) - (f_0(y + tv) - f_0(y))}{t} + \frac{f_0(y + tv) - f_0(y)}{t} - f_j^0(x; v)
\geq -\frac{|(f(y + tv) - f_0(y + tv)) + |f(y) - f_0(y)|}{t} + \frac{f_0(y + tv) - f_0(y)}{t} - f_j^0(x; v)
\geq \frac{-2\varepsilon}{t} + \left\{ \frac{f_0(y + tv) - f_0(y)}{t} - f_j^0(x; v) \right\}.
\]
The bracketed expression is greater than \(-1/k\) by equation (5.2). We may set \( \varepsilon \) (relying on \( t, y, f_0 \)) sufficiently small such that
\[
\frac{-2\varepsilon}{t} + \frac{f_0(y + tv) - f_0(y)}{t} - f_j^0(x; v) > -\frac{1}{k}.
\]
Thus, the same \( t \) and \( y \) may be used, and so \( B_\varepsilon(f_0) \subset G_k \).

(b) \( G_k \) is dense in \( \mathcal{X}_T \). With \( f \in \mathcal{X}_T \), for every \( 1/3 > \varepsilon > 0 \) we verify that open ball \( B_{3\varepsilon}(f) \) contains a point of \( G_k \). Define \( h : X \to \mathbb{R} \) by \( h(\tilde{x}) := f(x) - \varepsilon + f_j(\tilde{x}) - f_j(x) \), which is clearly in \( \mathcal{X}_T \), and set \( h_1 := \min\{f, h\}, \quad h_2 := \max\{f - 2\varepsilon, h_1\} \). Because \( f, f - 2\varepsilon, h \in \mathcal{X}_T \), Lemma 5.1.4 shows \( h_1 \in \mathcal{X}_T \), as is \( h_2 \). Since \( h_1 \leq f \) and \( f - 2\varepsilon \leq f \), we have \( f - 2\varepsilon \leq h_2 \leq f \).

It follows that \( \rho_1(f, h_2) \leq 2\varepsilon < 3\varepsilon < 1 \), so
\[
\rho(f, h_2) = \min\{1, \rho_1(f, h_2)\} = \rho_1(f, h_2) < 3\varepsilon.
\]
As \( f, f_j \) are continuous at \( x \), for \( 0 < \delta \) sufficiently small we have for \( y \in B_\delta(x) \),
\[
f(x) - \frac{\varepsilon}{2} \leq f(y) \leq f(x) + \frac{\varepsilon}{2} \text{ and, } \quad |f_j(y) - f_j(x)| < \frac{\varepsilon}{2}.
\] (5.3)
Now for \( \tilde{x} \in B_\delta(x) \) we have
\[
h(\tilde{x}) \leq f(x) - \varepsilon + \varepsilon/2 \leq f(x) - \frac{\varepsilon}{2} \leq f(\tilde{x}),
\]
and so $h_1 = \min\{f, h\} = h$ on $B_\delta(x)$. On the other hand, on $B_\delta(x)$ by equations (5.3) we have
\[ h_1(\bar{x}) = h(\bar{x}) \geq f(x) - \epsilon - \epsilon/2 \geq f(x) - \frac{3\epsilon}{2}, \quad \text{and} \quad f(\bar{x}) - 2\epsilon \leq f(x) - \frac{3\epsilon}{2}, \]
and so $h_2 = h_1 = h$ on $B_\delta(x)$. By the definition of $f_j^0(x; v)$, we may choose $0 < t < 1/k$ and $\|y - x\| < 1/k$ such that $y, y + tv \in B_\delta(x)$ and
\[ \frac{f_j(y + tv) - f_j(y)}{t} > \frac{f_j^0(x; v)}{t} - \frac{1}{k}, \quad (5.4) \]
For these $y$ and $t$, by equation (5.4) we have
\[
\frac{h_2(y + tv) - h_2(y)}{t} = \frac{h(y + tv) - h(y)}{t} = \frac{(f(x) - \epsilon - f_j(x) + f_j(y + tv)) - [(f(x) - \epsilon) - f_j(x) + f_j(y)]}{t} = \frac{f_j(y + tv) - f_j(y)}{t} > \frac{f_j^0(x; v)}{t} - \frac{1}{k},
\]
which shows $h_2 \in G_k$ while $h_2$ is arbitrarily close to $f$.

(c) Since $G_k$ is open and dense in $\mathcal{X}_T$ by (a) and (b), $G_{x,v} := \bigcap_{k=1}^\infty G_k$ is a dense $G_\delta$ in $\mathcal{X}_T$. If $f \in G_{x,v}$, then for every $k$ we can find $0 < t_k < 1/k$ and $\|y_k - x\| < 1/k$ such that
\[ \frac{f(y_k + t_k v) - f(y_k)}{t_k} - f_j^0(x; v) > -\frac{1}{k}, \]
and taking the limit we derive
\[ f^0(x; v) \geq \limsup_{\substack{y_k \to x \\ t_k \to 0}} \frac{f(y_k + t_k v) - f(y_k)}{t_k} \geq f_j^0(x; v). \]

(d) Since $f_j^0(x; v)$ is upper semicontinuous as a function of $(x, v)$ on the separable metric space $A \times X$, endowed with product metric, by Lemma 5.2.1 there exists a countable set $C := \{(x_k, v_k)\}$ being dense in $A \times X$ such that
\[ f_j^0(x; v) = \limsup_{(x_k, v_k) \in C \atop (x, v) \to (x, v)} f_j^0(x_k; v_k) \quad \text{for every} \quad (x, v) \in A \times X. \]
For each $(x_k, v_k)$, from (c) the set $G_{x_k,v_k}$ is dense $G_\delta$ in $\mathcal{X}_T$, then the set $G_{f_j} := \bigcap_{k=1}^\infty G_{x_k,v_k}$ is a dense $G_\delta$ set in $\mathcal{X}_T$. If $f \in G_{f_j}$, we have $f^0(x_k; v_k) \geq f_j^0(x_k; v_k)$ for all $k$. By the upper semicontinuity of $f^0(x; v)$ with respect to $(x, v)$, we obtain
\[ f^0(x; v) \geq \limsup_{(x_k, v_k) \in C \atop (x, v) \to (x, v)} f^0(x_k; v_k) \geq \limsup_{(x_k, v_k) \in C \atop (x, v) \to (x, v)} f_j^0(x_k; v_k) = f_j^0(x; v), \]
5.2 When is a set-valued map a Clarke subdifferential?

for all \((x, v) \in A \times X\), dually \(\partial_c f \supseteq \partial_c f_j\) on \(A\).

(e) From (d), \(G := \bigcap_{j=1}^{\infty} G_{f_j}\) is dense \(G_\delta\) in \(X_T\). If \(f \in G\), then \(\partial_c f \supseteq \partial_c f_j\) on \(A\) for all \(j\), and so \(\partial_c f(x) \supseteq \bigcup_{j=1}^{\infty} \partial_c f_j(x) = \Omega(x)\) for each \(x \in A\). Since \(CSC(\Omega)\) is the smallest cusco containing \(\Omega\), we have \(\partial_c f \supseteq CSC(\Omega) = T_x\) on \(A\) for every \(f \in G\).

The second claim follows by applying the first on the complete metric space \((X_{T_x}, \rho)\).

**Corollary 5.2.3** Let \(X\) be a separable Banach space. Assume \(\{f_j\}_{j=1}^{\infty}\) are locally equi-Lipschitz on \(A\). With \(T_x\) defined by equation (5.1), in \((X_{T_x}, \rho)\) the set

\[
\{f \in X_{T_x} : \partial_c f \equiv T_x\ \text{on } A\},
\]

is residual. In particular, \(T_x\) is a Clarke subdifferential.

**Proof.** By Theorem 5.2.2, it suffices to show \(f_j \in X_{T_x}\) for some norm-to-\(w^\ast\) cusco \(T\). Since \(\{f_j\}_{j=1}^{\infty}\) are locally equi-Lipschitz on \(A\), for every \(x \in A\) there exists a constant \(K_x\) and a neighbourhood \(U_x\) of \(x\) such that \(\partial_c f_j(y) \subseteq K_x B_{X^\ast}\) for each \(y \in U_x\) and every \(j\), then \(\Omega(y) \subseteq K_x B_{X^\ast}\) for every \(y \in U_x\). By Lemma 2.4.4 there exists a smallest cusco \(CSC(\Omega)\) containing \(\Omega\). Theorem 5.2.2 applies on the space \((X_{T_x}, \rho)\).

**Corollary 5.2.4** Assume \(\{f_j\}_{j=1}^{\infty}\) are locally equi-Lipschitz on \(A\). If the set-valued map \(T : A \to X^\ast\) defined by

\[
T(x) := \overline{\partial_c w^\ast}\{\partial_c f_j(x) : j = 1, \ldots, \infty\} \quad \text{for } x \in A,
\]  
(5.5)

is a norm-to-\(w^\ast\) cusco, then in \((X_T, \rho)\) the set \(\{f \in X_T : \partial_c f \equiv T\ \text{on } A\}\) is residual.

**Proof.** Since \(T_x(x) \supseteq T(x) \supseteq \Omega(x)\) for \(x \in A\) and \(T_x\) is the smallest cusco containing \(\Omega\), we obtain \(T_x = T\). Corollary 5.2.3 now applies.

In order to show that \(T\) defined by equation (5.5) is a \(w^\ast\)-cusco, it suffices to show it has a closed graph by Proposition 2.4.3.

**Definition 5.2.5** Let \(A\) be a non-empty convex subset of \(X^\ast\). The parallel body \(A_\delta\) is defined to be \(A_\delta := A + \delta B_{X^\ast}\).

---

The text above is a continuation of the discussion on Clarke subdifferentials, focusing on the conditions under which a set-valued map is a Clarke subdifferential. It introduces corollaries and definitions that are essential for understanding the properties of Clarke subdifferentials in the context of separable Banach spaces.
Let $A$, $B$ be non-empty $w^*$ compact convex subsets of $X^*$. The Hausdorff distance between $A$ and $B$ is defined as $D(A, B) := \inf \{ \delta : A \subset B_{\delta} \text{ and } B \subset A_{\delta} \}$.

One concrete condition to make $T$ defined by equation (5.5) a cuso is:

**Proposition 5.2.6** Assume $\partial_c f_j$ converges uniformly in the Hausdorff metric to a $w^*$-cuso $\Omega$ on $A$. Then $T : A \to X^*$ defined by equation (5.5) is a $w^*$-cuso.

**Proof.** We first show $T(x)$ is $w^*$-compact for $x \in A$. Given $\epsilon > 0$, there exists $K$ such that if $j \geq K$ we have $\partial_c f_j(x) \subset \Omega(x) + \epsilon B_{X^*}$. Then

$$T(x) \subset \overline{w^*} \{ \partial_c f_1(x), \ldots, \partial_c f_K(x), \Omega(x) + \epsilon B_{X^*} \}$$

$$= \text{co} \{ \partial_c f_1(x), \ldots, \partial_c f_K(x), \Omega(x) + \epsilon B_{X^*} \},$$

and the latter set is $w^*$-compact by Lemma 2.2.2.

Next $T$ is upper semicontinuous. For every $n$, there exists $K_n$ such that if $k \geq K_n$ we have $\Omega(x) \subset \partial_c f_k(x) + 1/n B_{X^*}$. Then for fixed $x^* \in \Omega(x)$, there exists $x^*_{K_n} \in \partial_c f_{K_n}(x)$ with $\|x^* - x^*_{K_n}\| \leq 1/n$, thus $x^*_{K_n} \to x^*$ in norm as $n \to \infty$. Hence $\Omega(x) \subset T(x)$. Let $W$ be a $w^*$-open set containing $T(x)$. We will show that $T(y) \subset W$ for $y$ near by $x$. Since $(X^*, w^*)$ is locally convex, the $w^*$-compact convex set $T(x)$ has a neighbourhood base of the form $T(x) + U$, where sets $U$ form a $w^*$-open neighbourhood base of 0. For such an $U$, we may choose $\epsilon > 0$ small such that $V + \epsilon B_{X^*} \subset U$, where $V$ is another $w^*$-open neighbourhood of 0. For this $\epsilon > 0$, we choose $K$ large such that for $j \geq K$, $\partial_c f_j(y) \subset \Omega(y) + \epsilon B_{X^*}$ for all $y \in A$, then

$$T(y) \subset \text{co} \{ \partial_c f_1(y), \ldots, \partial_c f_K(y), \Omega(y) + \epsilon B_{X^*} \}$$

$$\subset \text{co} \{ \partial_c f_1(y), \ldots, \partial_c f_K(y), \Omega(y) \} + \epsilon B_{X^*}. \quad (5.6)$$

For this $K$, as $\Omega(x) \subset T(x)$ we have $\text{co} \{ \partial_c f_1(x), \ldots, \partial_c f_K(x), \Omega(x) \} \subset T(x) \subset T(x) + V$. Since $\text{co} \{ \partial_c f_1(y), \ldots, \partial_c f_K(y), \Omega(y) \}$ is an usco by Proposition 2.4.2, for some $\delta > 0$ with $\|y - x\| < \delta$ we have $\text{co} \{ \partial_c f_1(y), \ldots, \partial_c f_K(y), \Omega(y) \} \subset T(x) + V$. Equations (5.6) shows $T(y) \subset T(x) + V + \epsilon B_{X^*} \subset T(x) + U$ for $\|y - x\| < \delta$. \hfill \qedsymbol

**Corollary 5.2.7** Assume $\{f_j\}_{j=1}^k$ are locally Lipschitz on $A$. Define

$$T(x) := \text{co} \{ \partial_c f_j(x) : j = 1, \ldots, k \} \quad \text{for } x \in A.$$
5.2 When is a set-valued map a Clarke subdifferential?

Then in \((\mathcal{X}_T, \rho)\), the set \(\{ f \in \mathcal{X}_T : \partial_c f \equiv T \text{ on } A \}\) is residual. In particular, the Clarke subdifferential is closed under the operation of taking finite convex hulls.

**Proof.** Combine Corollary 5.2.4 and Proposition 2.4.2.

Compare Corollary 5.2.7 with the “max formula”: \(\partial_c \max \{ f_j : j = 1, \ldots, k \} \subset T\) in total generality. Corollary 5.2.7 has generalized the main result of [21] in which the minimality of each \(\partial_c f_j\) was assumed.

**Example 5.2.8** Corollary 5.2.4 may fail for \(T\) defined by equation (5.5) not being upper semicontinuous. For each \(n\), consider the set-valued map \(T_n\) defined by

\[
T_n(x) := \left[ 0, \min\{ n|x|, 1 \} \right] \text{ for each } x \in \mathbb{R}.
\]

By Corollary 5.2.7, there exists \(f_n\) defined on \(\mathbb{R}\) such that \(\partial_c f_n = T_n\). Then the set-valued map \(T\) given by

\[
T(x) := \overline{\operatorname{co}} \left\{ \bigcup_{n=1}^{\infty} T_n(x) \right\} = \left\{ \begin{array}{ll} \{0\} & \text{if } x = 0, \\ [0, 1] & \text{otherwise.} \end{array} \right.
\]

is not upper semicontinuous at \(x = 0\) and so not a Clarke subdifferential mapping of any Lipschitz function on \(\mathbb{R}\). However, the the cusco generated by \(T\), \(\text{CSC}(T)(x) = [0, 1]\) for each \(x\), is always a Clarke subdifferential by Theorem 5.2.2.

**Corollary 5.2.9** In \((\mathcal{X}_C, \rho)\), the set \(\{ f \in \mathcal{X}_C : \partial_c f \equiv C \text{ on } A \}\) is residual. In particular, this holds for \(C = B_{X^*}\).

**Proof.** Since \(X\) is separable, \(C\), being \(w^*\)-compact, is \(w^*\)-metrizable. With this metric \(C\) is a compact metric space, thus there exists a countable set \(\{ x_j^* \}\), being dense in metric space \(C\), thus \(w^*\)-dense in \(C\). Define \(f_j : A \to \mathbb{R}\) by \(f_j(x) := \langle x_j^*, x \rangle\). For each \(j\), \(\partial_c f_j \equiv \{ x_j^* \}\). \(\{ f_j \}_{j=1}^{\infty}\) are 1-Lipschitz on \(A\). Now \(T := \overline{\operatorname{co}} w^* \{ x_j^* : j = 1, \ldots, \infty \} = C\), is a constant set-valued map, certainly a norm-to-\(w^*\) cusco. Corollary 5.2.4 applies.

**Corollary 5.2.10** Suppose \(f_0\) is locally Lipschitz on \(A\) and \(C \subset X^*\) is \(w^*\)-compact convex. Define \(T := \partial_c f_0 + C\). Then in \((\mathcal{X}_T, \rho)\) the set \(\{ f : \partial_c f \equiv T \text{ on } A \}\) is residual. In particular, \(T\) is a Clarke subdifferential.
5.2 When is a set-valued map a Clarke subdifferential?

Proof. By Proposition 2.4.1, $T$ is a norm-to-$w^*$ cusc. As $X$ is separable we may take a countable $w^*$-dense set $\{x_n^*\} \subset C$. For each $x_n^*$, we have $\partial_c(f_0 + \langle x_n^*, \cdot \rangle) = x_n^* + \partial_c f_0$, and $f_0(\cdot) + \langle x_n^*, \cdot \rangle \in X_T$. Moreover,

$$\partial_c f_0(x) + C \supset \omega^w \left[ \bigcup_{n=1}^{\infty} \partial_c f_0(x) + x_n^* \right] = \omega^w \left[ \partial_c f_0(x) + \{x_n^*\}_{n=1}^{\infty} \right] \supset \partial_c f_0(x) + C.$$

Since $C$ is norm bounded, $\{f_0 + \langle x_n^*, \cdot \rangle\}_{n=1}^{\infty}$ are locally equi-Lipschitz on $A$. It suffices now to apply Corollary 5.2.4.

**Corollary 5.2.11** Assume $f$ is locally Lipschitz on $A$ which is connected. Then the following are equivalent:

(i) $f$ is “strongly integrable”, that is, for each locally Lipschitz $g$ with $\partial_c g \subset \partial_c f$ we have $f - g \equiv \text{constant}$.

(ii) $f$ is “weakly integrable”, that is, for each locally Lipschitz $g$ with $\partial_c g \equiv \partial_c f$ we have $f - g \equiv \text{constant}$.

Proof. (i)$\Rightarrow$(ii) is clear. (ii)$\Rightarrow$(i): Fix $x_0 \in A$ and let $g$ be any member of $X_{\partial_c f}$. By Corollary 5.2.7 we may select, for each $0 < \varepsilon < 1$, a function $g_\varepsilon \in X_{\partial_c f}$ so that $\rho(g, g_\varepsilon) < \varepsilon$ and $\partial_c g_\varepsilon(x) = \partial_c f(x)$ for all $x \in A$. Then for any $x \in A$,

$$|(f - g)(x) - (f - g)(x_0)| \leq |(f - g_\varepsilon)(x) - (f - g_\varepsilon)(x_0)| + |(g_\varepsilon - g)(x) - (g_\varepsilon - g)(x_0)|$$

$$= |(g_\varepsilon - g)(x) - (g_\varepsilon - g)(x_0)| \quad \text{(since $f - g_\varepsilon \equiv \text{constant}$ on $A$)}$$

$$\leq |(g_\varepsilon - g)(x)| + |(g_\varepsilon - g)(x_0)| \leq 2\varepsilon.$$

As $\varepsilon$ was arbitrary, we have $(f - g)(x) = (f - g)(x_0)$, hence $f - g \equiv \text{constant}$ on $A$. \[ \square \]

Let $I$ be an open interval in $\mathbb{R}$ and let $f : I \to \mathbb{R}$. We say that $f$ is robustly lower (upper) semi-continuous if, $f(x) = \limsup_{y \to x} f(y)$ (or $f(x) = \liminf_{y \to x} f(y)$), for each Lebesgue null set $N$ of $I$. Continuous functions are obviously robustly lower semicontinuous and robustly upper semicontinuous. If $\beta_1$ and $\beta_2$ are robustly upper semicontinuous then $\max\{\beta_1, \beta_2\}$ is robustly upper semicontinuous, but $\beta_1 + \beta_2$ may not be robustly upper semicontinuous. Similarly for robustly lower semicontinuous continuous functions.
5.2 When is a set-valued map a Clarke subdifferential?

Example 5.2.12 A mapping $T : I \rightarrow 2^R$ is a cusco if and only if there are real valued functions $\alpha$ and $\beta$ on $I$ such that $\alpha \leq \beta$, $\alpha$ is lower semicontinuous, $\beta$ is upper semicontinuous and $T(x) = [\alpha(x), \beta(x)]$ for all $x \in I$. $T$ needs not be a Clarke subdifferential map unless one adds more conditions on $\alpha$ and $\beta$. One always has $T \supset \partial_c f$ where $f$ is given by $f(x) := \int_a^x \alpha(s)ds$. That is, the set $\mathcal{X}_T$ is always non-empty (but this might fail in higher dimensions). If $T$ is a minimal cusco, then $T = \partial_c f$. Hence each minimal cusco mapping on $I$ is a Clarke subdifferential map of a locally Lipschitz function. Let $N$ be an arbitrary null set, the map $T_N$ defined by $T_N(x) := \liminf_{y \in N} \alpha(y) \cup \limsup_{y \in N} \beta(y)$ is a cusco contained in $T$. As $T$ is minimal, we have $T_N = T$, thus $\alpha$ and $\beta$ are robustly lower semicontinuous and robustly upper semicontinuous respectively.

Corollary 5.2.13 Let $I$ be an open interval in $\mathbb{R}$ and $\alpha$ and $\beta$ be function on $I$ such that $\alpha \leq \beta$. Define $T := [\alpha, \beta]$. Then the following are equivalent:

(i) $\alpha$ is robustly lower semicontinuous and $\beta$ is robustly upper semicontinuous;

(ii) There exists a locally Lipschitz function $f$ on $I$ such that $\partial_c f \equiv [\alpha, \beta]$;

(iii) In $(\mathcal{X}_T, \rho)$ the set $\{f \in \mathcal{X}_T : \partial_c f \equiv [\alpha, \beta]\}$ is residual.

Proof. $(i) \Rightarrow (iii)$: Since $\alpha$ is lower semicontinuous and $\beta$ is upper semicontinuous on $I$, both of them are Lebesgue integrable. Choose any $p \in I$ and define $f_1, f_2 : I \rightarrow \mathbb{R}$ by

$$f_1(x) := \int_p^x \alpha(s)ds \quad \text{and} \quad f_2(x) := \int_p^x \beta(s)ds.$$ 

Let $N \subset I$ be a Lebesgue null set containing the nondifferentiability points of both $f_1$ and $f_2$. We have $\partial_c f_1(x) = [\liminf_{s \to z^-} \alpha(s), \limsup_{s \to z^-} \alpha(s)] = [\alpha(x), \limsup_{s \to z^-} \alpha(s)]$, and

$$\partial_c f_2(x) = [\liminf_{s \to z^-} \beta(s), \limsup_{s \to z^-} \beta(s)] = [\liminf_{s \to z^-} \beta(s), \beta(x)].$$

Since $\alpha \leq \beta$, we have $\limsup_{s \to z^-} \alpha(s) \leq \limsup_{s \to z^-} \beta(s) = \beta(x)$ and

$$\liminf_{s \to z^-} \beta(s) \geq \liminf_{s \to z^-} \alpha(s) = \alpha(x),$$

then $\text{co}\{\partial_c f_1(x), \partial_c f_2(x)\} = [\alpha(x), \beta(x)]$ for each $x \in I$, thus $T = \text{co}\{\partial_c f_1, \partial_c f_2\}$. Then Corollary 5.2.7 applies on the complete metric space $(\mathcal{X}_T, \rho)$. 

5.3 A characterization of the Clarke subdifferential

(iii) ⇒ (ii) is clear. We now show (ii) ⇒ (i). For every null set \( N \) and \( y \in I \) we have
\[
\beta(y) = \limsup_{z \to y, z \not\in N} f'(z) = \limsup_{z \to y} f'(z).
\]
Since \( [\alpha, \beta] \) is a usco, we have
\[
\limsup_{z \to y} \beta(y) \in [\alpha(x), \beta(x)],
\]
thus
\[
\beta(x) \geq \limsup_{y \to z, y \not\in N} \beta(y) = \limsup_{y \to z} \limsup_{z \to y} f'(z) \geq \limsup_{y \to z} f'(y) = \beta(x),
\]
as required. \( \Box \)

This has recovered and generalized the main result Theorem 1.2 [19].

**Example 5.2.14** The sum of two Clarke subdifferentials is usually not a Clarke subdifferential, as indicated by the sum of \( \partial | \cdot | \) and \( -\partial | \cdot | \).

The intersection of two Clarke subdifferentials is usually not a Clarke subdifferential. Define \( f_1(x) := 1 \) if \( 0 < x < \infty \) and \( 0 \) otherwise, and \( f_2(x) := 1 \) if \( -\infty < x < 0 \) and \( 0 \) otherwise. Both \( f_1 \) and \( f_2 \) are robustly lower semicontinuous, thus \( T_1 \) and \( T_2 \) defined by \( T_1(x) := [f_1(x), 2] \) and \( T_2(x) := [f_2(x), 2] \) for \( x \in \mathbb{R} \) are Clarke subdifferential maps by Corollary 5.2.13. However, \( T_1(x) \cap T_2(x) = [1, 2] \) if \( x \neq 0 \) and \( T_1(0) \cap T_2(0) = [0, 2] \) is not a Clarke subdifferential as \( T_1 \cap T_2 \) is sensitive to null set. Note that \( \max\{f_1, f_2\} \) is not robustly lower semicontinuous at \( 0 \), but lower semicontinuous at \( 0 \).

5.3 A characterization of the Clarke subdifferential

**Lemma 5.3.1** Let \( A \) be an open subset of a separable Banach space \( X \) and \( \Omega : A \to X^* \) be a densely defined locally bounded multifunction. Then there exists a countable set \( C \subset \text{Gr}(\Omega) \) such that \( \text{Gr}(\text{USC}(\Omega)) = \overline{C} \) where the closure is taken in the product topology of \( X \times X^* \) with \( X^* \) endowed with \( w^* \)-topology.

**Proof.** First every subspace of a separable metric space is separable [60, page 58]. Define \( A_m := \{ x \in A : \Omega(x) \subset mB_{X^*} \} \). As \( B_{X^*} \) is \( w^* \)-compact and metrizable in the \( w^* \)-topology, \( A_m \times mB_{X^*} \) is separable in the product topology, thus \( \text{Gr}(\Omega) \cap (A_m \times mB_{X^*}) \) is separable. Then there exists a countable set \( C_m \subset \text{Gr}(\Omega) \cap (A_m \times mB_{X^*}) \) with \( \text{Gr}(\Omega) \cap (A_m \times mB_{X^*}) \subset \overline{C_m} \). Let \( C := \bigcup_{m=1}^{\infty} C_m \). Then \( C \) is countable and
\[
\text{Gr}(\Omega) = \text{Gr}(\Omega) \cap \bigcup_{m=1}^{\infty} (A_m \times mB_{X^*}) = \bigcup_{m=1}^{\infty} \text{Gr}(\Omega) \cap (A_m \times mB_{X^*}) \subset \bigcup_{m=1}^{\infty} \overline{C_m} \subset \overline{C}.
\]
5.3 A characterization of the Clarke subdifferential

By Proposition 2.4.4, \( \text{Gr}(\text{USC}(\Omega)) = \cl\text{Gr}(\Omega) \subset \text{cl}C \subset \cl\text{Gr}(\text{USC}(\Omega)) = \text{Gr}(\text{USC}(\Omega)) \). \( \square \)

**Lemma 5.3.2** Assume \( \mathcal{F} := \{ f_\alpha : \alpha \in I \} \) is a family of locally equi-Lipschitz functions on \( A \), in which \( I \) may be countable or uncountable. Define \( \Omega : A \to 2^{X^*} \) by \( \Omega(x) := \bigcup_{\alpha \in I} \partial f_\alpha(x) \) for \( x \in A \). Then \( \text{CSC}(\Omega) \) is a Clarke subdifferential map.

**Proof.** First \( \text{CSC}(\Omega)(x) = \overset{\text{w*}}{\text{USC}(\Omega)}(x) \). By Lemma 5.3.1 there exists a countable set \( C \subset \text{Gr}(\Omega) \) such that \( \text{Gr}(\text{USC}(\Omega)) = \cl C \). Write

\[
C = \{(x_i, x_i^*) : (x_i, x_i^*) \in \text{Gr}(\Omega)\} = \{(x_i, x_i^*) : x_i^* \in \partial f_\alpha(x_i) \text{ for some } \alpha \in I\}.
\]

Define \( \hat{\Omega} : A \to 2^{X^*} \) by \( \hat{\Omega}(x) := \bigcup_{i=1}^\infty \partial f_\alpha(x) \) for \( x \in A \). Then \( \hat{\Omega}(x) \subset \Omega(x) \) for every \( x \in A \) and \( x_i^* \in \hat{\Omega}(x_i) \) for all \( i \in \mathbb{N} \). Then

\[
\text{Gr}(\text{USC}(\Omega)) = \cl C \subset \text{clGr}(\hat{\Omega}) \subset \text{Gr}(\text{USC}(\Omega)) \subset \text{Gr}(\text{USC}(\Omega)).
\]

This yields \( \text{USC}(\Omega) = \text{USC}(\hat{\Omega}) \), and so \( \text{CSC}(\Omega) = \text{CSC}(\hat{\Omega}) \). Since \( \text{CSC}(\hat{\Omega}) \) is a cuscogenerated by the subdifferentials of a countable family of locally equi-Lipschitz functions, it suffices to apply Theorem 5.2.2. \( \square \)

**Theorem 5.3.3** Let \( A \) be an open subset of a separable Banach space \( X \) and let \( T : A \to 2^{X^*} \) be a \( w^* \)-cuscog. Then the following are equivalent:

(i) \( T \) is a Clarke subdifferential map;

(ii) \( T = \text{CSC}(\Omega) \) where \( \Omega(x) := \bigcup_{\alpha \in I} \partial f_\alpha(x) \) for each \( x \in A \) and \( \{ f_\alpha : \alpha \in I \} \) are equi-locally Lipschitz on \( A \);

(iii) \( T = \text{CSC}(\Omega) \) where \( \Omega(x) := \bigcup_{i=1}^\infty \partial f_i(x) \) for each \( x \in A \) and \( \{ f_i : i \in \mathbb{N} \} \) are equi-locally Lipschitz on \( A \).

**Proof.** (i) \( \iff \) (ii) is Lemma 5.3.2. (i) \( \iff \) (iii) is Theorem 5.2.2. \( \square \)

We believe the above theorem opens a new direction in our understanding of the possible behavior of Clarke subdifferential maps. A locally equi-Lipschitz criterion is:

**Proposition 5.3.4** Let \( A \) be an open convex subset of Banach space \( X \). Let \( \mathcal{F} \) be a family of real-valued lower semicontinuous convex functions on \( A \), and suppose that for each \( x \in A \)
there is a number $M_x$ such that $|f(x)| \leq M_x$ for all $f \in \mathcal{F}$. Then $\mathcal{F}$ is locally equi-Lipschitz on $A$.

**Proof.** By the Barie category theorem there exists a non-empty open set $O \subset A$ and a constant $M$ such that $f(x) \leq M$ for all $f \in \mathcal{F}$ and $x \in O$. By convexity of $A$ and $f \in \mathcal{F}$, for every $a \in A$ there exists an open ball $B_{r_0}(a)$ of a such that $m \leq f(x) \leq M$ for each $f \in \mathcal{F}$ and $x \in B_{r_0}(a)$. Let $0 < r < r_0$. We get $|f(x) - f(y)| \leq (M - m)/(r_0 - r)\|y - x\|$ for every $x, y \in B_r(a), f \in \mathcal{F}$, which proves the stated result. \hfill $\square$

A convex function $f$ defined on an open convex subset $A$ of a Banach space $X$ is continuous on $A$ if and only if $f$ is bounded above on a neighbourhood of some point $a \in A$ [49].

**Corollary 5.3.5** [88] Let $U \subset \mathbb{R}^n$ be open, and let $\mathcal{F} := \{f_i : i \in I\}$ be a collection of convex functions $\mathbb{R}^n \to (-\infty, +\infty)$. If the set $\{f_i(x) : i \in I\}$ is bounded for each $x \in U$, then $\mathcal{F}$ is locally equi-Lipschitz relative to $U$.

One may be wondering why we do not try to use the supremum function. If $\{f_\alpha : \alpha \in I\}$ is equi-locally Lipschitz on $A$ and $\{f_\alpha(x) : \alpha \in I\}$ is bounded for each $x \in A$, then $f : X \to \mathbb{R}$ defined by $f(y) := \sup\{f_\alpha(y) : \alpha \in I\}$ is locally Lipschitz on $A$ and,

(i) If $A$ is convex, and $f_\alpha(x)$ is continuous as a function of $\alpha$ and convex as a function of $x$, then $f$ is convex.

(ii) If $I$ is sequentially compact, the map $\alpha \to f_\alpha(y)$ is upper semicontinuous for each $y \in A$, and $f_\alpha$ admits a strict derivative $D_\alpha f_\alpha(x)$ on $A$, and that $D_\alpha f_\alpha(x)$ is continuous as a function of $(\alpha, x)$, then for each $x \in A$, $f$ is regular at $x$ [33, page 87].

(iii) If $I$ is compact, $\alpha \to f_\alpha(x)$ is upper semicontinuous on $I$, and for each $y \in X$ the mapping $(x, \alpha) \to f_\alpha^+(x, y)$ is upper semicontinuous on $A \times I$, then $f \in S_c(A)$ [20, page 330].

In essence, pointwise maxima typically yield a locally Lipschitz function having a minimal Clarke subdifferential not maximal one!
5.4 The approximate subdifferential

Owing to its definition, topological properties of the approximate subdifferential are closely related to the local behavior of the Dini-Hadamard subdifferential.

**Lemma 5.4.1** Let \( X \) be a separable Banach space and \( C \subset X^* \) be a \( w^* \)-compact convex set with nonempty norm interior. Then there exists a countable \( \{ c_k \} \) from the norm interior of \( C \) such that

\[
C = w^* \text{cl}\{ \lim_{n \to \infty} c_{n_k} : \{ c_{n_k} \} \text{ is norm converging subsequences of} \{ c_k \} \}.
\]  

(5.7)

**Proof.** Since \( X \) is separable, \( C \) is metrizable in the \( w^* \)-topology. As \( C \) is a compact metric space in this metric, we may take a countable set \( \{ y^n_0 \} \) being \( w^* \)-dense in \( C \). Choose \( y^n_0 \) from the norm interior \( C \). For every \( y^n_0 \), by the convexity of \( C \), the half-open line segment \( [y^n_0, y^n_*] \) lies in the norm interior of \( C \). Define \( x^*_{n,k} = (1/k)y^n_0 + (1 - 1/k)y^n_* \). Set \( \{ c_k \} := \bigcup_{n=1}^{\infty} \bigcup_{k=1}^{\infty} x^*_{n,k} \). As \( y^n_* = (\text{norm}) \lim_{k \to \infty} x^*_{n,k} \), the proof is complete. \( \square \)

Let \( C \) be a \( w^* \)-compact convex set with non-empty norm interior. We denote the norm interior of \( C \) by \( \text{int}(C) \). In a \( C \)-Lipschitz setting we get a very broad genericity result.

**Theorem 5.4.2** In \( (X_C, \rho) \) the set \( \{ f \in X_C : \partial_a f \equiv C \text{ on } A \} \) is residual.

**Proof.** Fix \( x \in A, c \in \text{int}(C), n > 1 \). Consider \( G^n_{x,c} := \)

\[
\{ f \in X_C : \text{ there exists } \bar{x} \in A \text{ satisfying } \| \bar{x} - x \| < \frac{1}{n} \text{ such that for some } \bar{x}^* \in \partial_- f(\bar{x}) \text{ we have } \| \bar{x}^* - c \|_* < \frac{1}{n} \}.
\]

(a) \( \text{int}(G^n_{x,c}) \) is dense in \( X_C \). Choose an arbitrary \( f \in X_C \) and \( \epsilon > 0 \). Define

\[
h(\bar{x}) := f(\bar{x}) - \epsilon + \sigma_C(\bar{x} - x), \quad h_1 := \min\{ f, h \} \text{ and,} \quad h_2 := \max\{ f - 2\epsilon, h_1 \}.
\]

Then \( h_2 \in X_C \). For \( 0 < \delta < 1/n \) sufficiently small, we have

\[
h_2(\bar{x}) = h(\bar{x}) = f(\bar{x}) - \epsilon + \sigma_C(\bar{x} - x) \quad \text{on } B_\delta[x].
\]

As \( h \) is convex on \( B_\delta(x) \), \( \partial_- h_2(x) = C \), we see that \( h_2 \in G^n_{x,c} \). Moreover, \( \rho(h_2, f) < 3\epsilon \). We proceed to show that \( h_2 \in \text{int}( G^n_{x,c} ) \). Because \( c \in \text{int}(C) \), we may assume \( \tau B_{X^*} \subset C - c \)
for some $r > 0$, then for $\bar{x} \neq x$ we have $\sigma_{C-\varepsilon}(\bar{x} - x) > 0$ by the Hahn-Banach theorem. On $B_{\delta}[x]$, for $x \neq x$ we have

$$h_2(\bar{x}) - \langle c, \bar{x} \rangle = f(x) - \varepsilon + \sigma_C(\bar{x} - x) - \langle c, \bar{x} \rangle = h_2(x) - \langle c, x \rangle + \sigma_{C-\varepsilon}(\bar{x} - x),$$

and so

$$m := \inf\{h_2(\bar{x}) - \langle c, \bar{x} \rangle : \|\bar{x} - x\| = \delta\} = h_2(x) - \langle c, x \rangle + \inf\{\sigma_{C-\varepsilon}(\bar{x} - x) : \|\bar{x} - x\| = \delta\}
\geq h_2(x) - \langle c, x \rangle + r\delta > h_2(x) - \langle c, x \rangle.
$$

Let $\alpha := m - (h_2(x) - \langle c, x \rangle) > 0$. For $g \in X_C$ with $\rho(g, h_2) < \beta < \alpha/2$, for $\|\bar{x} - x\| = \delta$ we have $g(\bar{x}) - \langle c, \bar{x} \rangle = g(\bar{x}) - h_2(\bar{x}) + h_2(\bar{x}) - \langle c, \bar{x} \rangle \geq -\beta + m$, and

$$g(x) - \langle c, x \rangle = g(x) - h_2(x) + h_2(x) - \langle c, x \rangle \leq \beta + h_2(x) - \langle c, x \rangle.
$$

Then $\inf\{g(\bar{x}) - \langle c, \bar{x} \rangle : \|\bar{x} - x\| = \delta\} \geq -\beta + m > \beta + h_2(x) - \langle c, x \rangle \geq g(x) - \langle c, x \rangle$. Define $g_1 := g - \langle c, \cdot \rangle + I_{B_\delta[x]}$, which is lower semicontinuous and bounded below on $X$.

$$\inf_{X} g_1 = \inf_{B_\delta[x]} (g(\bar{x}) - \langle c, \bar{x} \rangle) \leq g(x) - \langle c, x \rangle < \inf\{g(\bar{x}) - \langle \bar{x}, c \rangle : \|\bar{x} - x\| = \delta\}.
$$

Let $0 < \nu < \min\{1/(2n), \inf\{g(\bar{x}) - \langle \bar{x}, c \rangle : \|\bar{x} - x\| = \delta\} - (g(x) - \langle x, c \rangle)\}$. Choose $x_0$ such that $g_1(x_0) = g(x_0) - \langle c, x_0 \rangle < \inf_X g_1 + \nu$. Since $X$ is separable, $X$ is smoothable. With $\lambda = 1$, Theorem 2.7.3 shows there exists a Gâteaux differentiable $\phi$ and $v \in X$ such that for all $\bar{x} \in B_\delta[x]$,

$$g(\bar{x}) - \langle c, \bar{x} \rangle + \phi(\bar{x}) \geq g(v) - \langle c, v \rangle + \phi(v),$$

$$g(v) - \langle c, v \rangle < \nu + \inf_{X} g_1,$$  

$$\|\nabla \phi(v)\| < 2\nu < \frac{1}{n}.  \tag{5.10}$$

Equation (5.9) shows $g(v) - \langle c, v \rangle < \inf\{g(\bar{x}) - \langle c, \bar{x} \rangle : \|\bar{x} - x\| = \delta\}$, thus $\|v - x\| < \delta < 1/n$. Equation (5.8) shows $0 \in \partial_{-} g(v) - \frac{c + \nabla \phi(v)}{\nu}$, thus there exist $v^* \in \partial_{-} g(v)$ with $\|v^* - c\| = \|\nabla \phi(v)\| < 1/n$ by equation (5.10). Then $g \in G^n_{x,c}$, and so $B_\beta(h_2) \subset G^n_{x,c}$, that is, $h_2 \in \text{int}(G^n_{x,c})$. [When $X$ is finite dimensional, we may choose $\phi = 0$ by the compactness of $B_\delta[x]$.]

(b) Since $\text{int}(G^n_{x,c})$ is open and dense in $X_C$, the set $G^n_{x,c} := \bigcap_{n=1}^{\infty} \text{int}(G^n_{x,c})$ is dense in $X_C$. If $f \in G_{x,c}$, then for every $n$ there exists $\tilde{x}_n^* \in \partial_{-} f(x_n)$ with $\|x_n^* - c\| < 1/n$ and $\|x_n - x\| < 1/n$. Letting $n \to \infty$, we obtain $c \in \partial_{a} f(x)$. 


5.5 The Gâteaux differentiability of C-Lipschitz functions

(e) By Lemma 5.4.1 we may take a countable dense set \( \{ c_k \} \) from \( \text{int}(C) \) satisfying equation (5.7). As \( G_{x,c_k} \) is dense \( G_\delta \) for each \( c_k \), \( G_x := \bigcap_{k=1}^{\infty} G_{x,c_k} \) is a dense \( G_\delta \) set in \( X_C \). If \( f \in G_x \), then \( c_k \in \partial_a f(x) \) for every \( k \) by (b), thus \( C \subset \partial_a f(x) \) by the \( \text{w}^* \)-closedness of \( \partial_a f(x) \). As \( f \in X_C \), we have \( \partial_a f(x) = C \) if \( f \in G_x \).

(d) Now let \( \{ x_k \} \) be a countable norm dense set in \( A \). Since \( G_{x_k} \) is dense \( G_\delta \) for every \( k \), the set \( G := \bigcap_{k=1}^{\infty} G_{x_k} \) is a dense \( G_\delta \) set in \( X_C \). If \( f \in G \), for every \( k \) we have \( \partial_a f(x_k) = C \) by (c). By the norm-to-\( \text{w}^* \) upper semicontinuity of \( \partial_a f \), we get \( C \supset \partial_a f(x) \supset \limsup_{x_k \rightarrow x} \partial_a f(x_k) \supset C \). Hence \( \partial_a f \equiv C \) for every \( f \in G \).  

Corollary 5.4.3 Let \( C \) be a \( \text{w}^* \)-compact convex set with non-empty norm interior. In \( (X_C, \rho) \) the set \( \{ f : \partial_a f = \partial_a f \equiv C \text{ on } A \} \) is residual. Hence, in \( (X_{B_X}, \rho) \) the set \( \{ f : \partial_a f = \partial_a f \equiv B_{X^*} \text{ on } A \} \) is residual. In particular, in \( (X_{B_X}, \rho) \) the set of functions with minimal Clarke or approximate subdifferentials is first category.

Corollary 5.4.4 Every \( C^1 \)-tube is an approximate subdifferential map.

Proof. By Corollary 5.6.3 [58], if \( f \) is strictly Hadamard differentiable at \( x \) and \( g \) is Lipschitz near by \( x \), then \( \partial_a(f + g)(x) = \nabla f(x) + \partial_a g(x) \). The result follows by taking \( g \) to be a Lipschitz function on \( A \) with \( \partial_a g \equiv B_{X^*} \) and \( f \) to be a Gâteaux differentiable locally Lipschitz function with \( \nabla f \) norm-to-\( \text{w}^* \) continuous on \( A \).

5.5 The Gâteaux differentiability of C-Lipschitz functions

Our result is closely related to the differentiability properties of locally Lipschitz functions on separable Banach spaces.

Lemma 5.5.1 Let \( f \) be a real-valued locally Lipschitz function defined on a nonempty open subset \( A \) of a Banach space \( X \). Then for each \( x \in A \) and \( v \in X \)

\[
f^0(x;v) = \limsup_{z \rightarrow x} f^+(z;v) = \limsup_{z \rightarrow x} f^-(z;v).
\]

Proof. Given \( v \in X \), it is clear that

\[
f^0(x;v) = \limsup_{z \rightarrow x} f^0(z;v) \geq \limsup_{z \rightarrow x} f^+(z;v) \geq \limsup_{z \rightarrow x} f^-(z;v).
\] (5.11)
5.5 The Gâteaux differentiability of C-Lipschitz functions

Given $\epsilon > 0$, in any neighbourhood of $x$ there exists $z_0 \in A$ and $\lambda_0 > 0$ such that $z_0 + \lambda_0 v \in A$ and

$$\frac{f(z_0 + \lambda_0 v) - f(z_0)}{\lambda_0} > f^0(x; v) - \epsilon.$$  

Consider $f$ restricted to the interval $[z_0, z_0 + \lambda_0 v]$. Since $f$ is locally Lipschitz, it follows from Lebesgue's mean-value theorem that there exists $0 \leq \lambda_1 \leq \lambda_0$ such that

$$f'(z_0 + \lambda_1 v; v) \geq \frac{f(z_0 + \lambda_0 v) - f(z_0)}{\lambda_0}.$$  

So $\limsup_{z \to x} f^-(z; v) \geq f^0(x; v)$ and with equation (5.11) we conclude that

$$f^0(x; v) = \limsup_{z \to x} f^-(z; v) = \limsup_{z \to x} f^+(z; v). \quad \Box$$

For a locally Lipschitz function $f$ on an open set $A \subset X$, given $x \in A$ and $v \in X$ and $p \in \mathbb{N}$ we write $f^+_p(x; v) := \sup\{|f(x + \lambda v) - f(x)|\lambda^{-1} : 0 < \lambda < 1/p\}$. Clearly

$$f^+(x; v) = \lim_{p \to \infty} f^+_p(x; v) = \inf_{p \in \mathbb{N}} f^+_p(x; v). \quad (5.12)$$

Given $v \in X$ and $p \in \mathbb{N}$, as a supremum of continuous functions, $f^+_p(\cdot; v)$ is lower semicontinuous, thus generically continuous on $A$ by Proposition 2.1.3.

**Lemma 5.5.2** [50] A locally Lipschitz function $f$ on an open set $A$ of a Banach space $X$ is, generically on $A$, partially pseudo-regular on every separable subspace.

**Proof.** Given $v \in X$ and $p \in \mathbb{N}$, by Proposition 2.1.3 there exists a dense $G_\delta$ subset $D_v$ of $A$ such that $f^+_p(\cdot; v)$ is continuous at the points of $D_v$ for every $p \in \mathbb{N}$. We show that $f^+(\cdot; v)$ is upper semicontinuous at the points of $D_v$. Given $\epsilon > 0$ and $x \in D_v$ there exists $p \in \mathbb{N}$ such that $f^+(x; v) > f^+_p(x; v) - \frac{\epsilon}{2}$. But there exists a $\delta_p > 0$ such that $|f^+_p(z; v) - f^+_p(x; v)| < \frac{\epsilon}{2}$ for all $z \in A$, $\|z - x\| < \delta_p$. By equation (5.12), $f^+(x; v) + \epsilon > f^+_p(z; v) \geq f^+(z; v)$ for all $z \in A$, $\|z - x\| < \delta_p$, thus $f^+(\cdot; v)$ is upper semicontinuous at $x$. By Lemma 5.5.1, $f^0(x; v) = \limsup_{z \to x} f^+(z; v) = f^+(x; v)$. If $Y$ is a separable subspace of $X$, then there exists a countable dense set $\{v_n\}$ in $Y$ and therefore a dense $G_\delta$ subset $D := \bigcap_{n=1}^\infty D_{v_n}$ of $A$ where for each $x \in D$ we have $f^+(x; v_n) = f^0(x; v_n)$ for all $n \in \mathbb{N}$. But both $f^+(x, \cdot)$ and $f^0(x, \cdot)$ are continuous, we conclude that for every $x \in D$,

$$f^+(x; v) = f^0(x; v) \quad \text{for all } v \in Y. \quad \Box$$
Theorem 5.5.3 Assume $C$ is $w^*$-compact convex and not singleton. In $(\mathcal{X}_C, \rho)$, the set
\[ G := \{ g : g \text{ is Dini-Hadamard subdifferentiable only on a first category subset of } A \}, \]
is residual. In particular, each $g \in G$ is Gâteaux differentiable only on a first category set.

Proof. By Corollary 5.2.9, the set $G := \{ g \in \mathcal{X}_C : \partial_c g \equiv C \text{ on } A \}$ is residual in $(\mathcal{X}_C, \rho)$. For each $g \in G$, $\partial_c (-g) = -C$, and the set $D_{-g} := \{ x \in A : \partial^+ (-g)(x) = \partial_c (-g)(x) = -C \}$ is residual in $A$ by Lemma 5.5.2. Assume $x^* \in \partial_- g(x)$ for $x \in A$. We will show that $x \in A \setminus D_{-g}$. Let us suppose that $x \in D_{-g}$. Take $y^* \in \partial^+ (-g)(x)$. For all $v \in X$ we have
\[ \langle x^*, v \rangle \leq \liminf_{t \downarrow 0} \frac{g(x + tv) - g(x)}{t}, \quad \langle y^*, v \rangle \leq \limsup_{t \downarrow 0} \frac{(-g)(x + tv) - (-g)(x)}{t}. \]
Then $\langle x^*, v \rangle \leq \langle -y^*, v \rangle$ for all $v \in X$, thus $x^* = -y^*$, and
\[ (-g)^+ (x; v) = \limsup_{t \downarrow 0} \frac{(-g)(x + tv) - (-g)(x)}{t} = \langle -x^*, v \rangle \text{ for all } v \in X, \]
which is impossible since $(-g)^+ (x; v) = (-g)^0 (x; v) = \sigma_C (v)$ is not a linear function of $v$ when $C$ is non-singleton.

Theorem 5.5.3 forms a sharp contrast to these well-known facts:

Theorem 5.5.4 (i) (Christensen) Every locally Lipschitz function on a separable Banach space is Gâteaux differentiable everywhere except for a Haar null set [30];

(ii) (Mazur) A continuous convex function on an open convex subset $A$ of a separable Banach space $X$ is Gâteaux differentiable on a dense $G_\delta$ subset of $A$ [72].

For a locally Lipschitz $f : A \to \mathbb{R}$, we define the gradient oscillation of $f$ at $x$ by
\[ \Omega^f_H (x) := \{ w^* \in \lim_{x_n \to x} \nabla f (x_n) : x_n \notin H \}, \]
where $H$ is any Haar-null set containing the set of points at which $f$ fails to be Gâteaux differentiable. Since $X$ is separable, by Proposition 2.8.3 we deduce that $\Omega^f_H (x) \neq \emptyset$. Thibault [89] showed that $\Omega^f_H (x)$ is weak$^*$ compact and $\partial_c f (x) = \mathfrak{e} \omega^* \Omega^f_H (x)$.

Corollary 5.5.5 If $X$ is an infinite dimensional smooth separable Banach space, then the set $\{ f \in \mathcal{X}_{B_X} : \Omega^f_H (x) \equiv B_{X^*} \text{ on } A \}$ is residual in $(\mathcal{X}_{B_X}, \rho)$.
5.6 Further properties of approximate subdifferentials

Proof. Assume $f$ has $\partial_x f \equiv B_{X^*}$. From the Krein-Milman converse theorem, we have $\text{ext} B_{X^*} \subset \Omega_H^f(x)$. By Proposition 2.8.2, $S_{X^*} \subset \Omega_H^f(x)$, and so $B_{X^*} \subset \Omega_H^f(x)$ by Proposition 2.8.1. Apply Corollary 5.2.9.

However, Rockafellar’s function shows Corollary 5.5.5 fails when $X$ is finite dimensional.

5.6 Further properties of approximate subdifferentials

By Lemma 5.3.1, when $f$ is a locally Lipschitz function on an open subset $A$ of a separable Banach space $X$, there exists a countable set $C \subset \text{Gr}(\partial_- f)$ such that $\text{Gr} \partial_a f = \text{cl} C$ where the closure is taken in the product topology of $X \times X^*$ with $X^*$ endowed with $w^*$-topology.

Given a locally Lipschitz function $f \in \mathcal{X}_T$, for each $(x, x^*) \in \text{Gr}(\partial_- f)$, usually the set $\{g \in \mathcal{X}_T : x^* \in \partial_- g(x)\}$ is not residual in $\mathcal{X}_T$, but the set $\{g \in \mathcal{X}_T : x^* \in \partial_a g(x)\}$ is always residual in $\mathcal{X}_T$ as Theorem 5.6.1 illustrates. As $\text{Gr}(\partial_a f)$ can be generated by a countable elements from $\text{Gr}(\partial_- f)$, we may find a residual subset $G \subset \mathcal{X}_T$ such that if $g \in G$ then $g$ has at least the same approximate subdifferential as $f$. The same holds for the Clarke subdifferential.

Theorem 5.6.1 Let $A$ be a non-empty open subset of a separable Banach space $X$. Let $T : A \to 2^{X^*}$ be a $w^*$-cuscoc on $A$. Then for each $f \in \mathcal{X}_T$, the set

$$\{g \in \mathcal{X}_T : \partial_a f(x) \subset \partial_a g(x) \text{ for all } x \in A\},$$

is residual in $(\mathcal{X}_T, \rho)$.

Proof. Since $X$ is separable, we may choose a increasing sequence of finite dimensional subspaces of $X$ such that $\text{cl} \bigcup_{n=1}^{\infty} X_n = X$. For each $(x, x^*) \in \text{Gr}(\partial_- f)$ and $n \in \mathcal{N}$ we consider the set

$$G(x, x^*, n) := \{g \in \mathcal{X}_T : \text{ there exists a point } z \in B_{1/n}(x) \text{ and } z^* \in \partial_- g(z) \text{ so that } \|x^*|x_n - z^*|x_n\| \leq 4/n\}.$$

(a) For each $(x, x^*) \in \text{Gr}(\partial_- f)$ and $n \in \mathcal{N}$, $\text{int}[G(x, x^*, n)]$ is dense in $(\mathcal{X}_T, \rho)$. Suppose $(g_0, \epsilon) \in \mathcal{X}_T \times (0, 1)$ is given. We need to verify that $B_{\epsilon}(g_0) \cap \text{int}[G(x, x^*, n)] \neq \emptyset$. Define $h_1 \in \mathcal{X}_T$ by $h_1(z) := f(z) + (g_0(x) - \epsilon/3 - f(x))$ and the function $h_2 \in \mathcal{X}_T$ by $h_2(z) :=$
5.6 Further properties of approximate subdifferentials \(106\)

\[ \min\{g_0(z), h_1(z)\}, \] that is, \(h_2 = g_0 \wedge h_1 \in \mathcal{X}_T \) by Lemma 5.1.4. Clearly, \(h_2(z) \leq g_0(z)\) for all \(z \in A\). Next, define \(h_3 \in \mathcal{X}_T\) by \(h_3(z) := \max\{h_2(z), g_0(z) - 2\varepsilon/3\}\) that is, \(h_3 = h_2 \vee (g_0 - 2\varepsilon/3) \in \mathcal{X}_T\) by Lemma 5.1.4. Clearly,

\[ g_0(z) - \frac{2\varepsilon}{3} \leq h_3(z) \leq g_0(z) \text{ for all } z \in A, \]

and so \(\rho(h_3, g_0) = \min\{1, d(h_3, g_0)\} < \varepsilon\). We claim that \(h_3 \in \text{int}[G(x,x^*, n)]\). To see this, first note that:

\[ g_0(x) - \frac{2\varepsilon}{3} < h_3(x) = h_1(x) = g_0(x) - \frac{\varepsilon}{3} < g_0(x). \]

Therefore there exists an open neighbourhood \(U\) of \(x\) so that \(h_1 = h_2 = h_3\) on \(U\). Since \(x^* \in \partial_- f(x)\), \(f\) is Lipschitz around \(x\), and \(S_{X_n}\) is compact, there exists \(0 < \delta < 1/n\) so that (i) \(B_{\delta}(x) \subset U\); (ii) \(f\) is Lipschitz on \(B_{\delta}(x)\); (iii)

\[ (f - x^*)(x + \lambda u) - (f - x^*)(x) > \frac{\lambda}{n} \geq \frac{\delta}{n} \text{ for all } 0 < \lambda < \delta \text{ and } v \in S_{X_n}. \]

We now show that \(B_r(h_3) \subset G(x,x^*, n)\) for any \(0 < r < \delta/(2n)\). To accomplish this, let \(g\) be any member of \(B_r(h_3)\). Then

\[ (g - x^*)(z) - (g - x^*)(x) > \frac{2\delta}{n} \text{ for all } z \in x + \delta B_{X_n}. \]

By Proposition 2.7.5, there exists \(z, \xi\) satisfying \(\|z - x\| < \delta\), \(\xi \in \partial_-(g - x^*)(z)\) and \(\|\xi|_{X_n}\| \leq 4/n\), that is, \(\|z^*|_{X_n} - x^*|_{X_n}\| \leq 4/n\) for some \(x^* \in \partial_- g(z)\). This shows that \(g \in G(x,x^*, n)\).

(b) Fix \((x,x^*) \in \text{Gr}(\partial_- f)\). For each \(g \in G(x,x^*) := \bigcap_{n \in \mathbb{N}} G(x,x^*, n)\), we have \(x^* \in \partial_a g(x)\).

By (a), the set \(G(x,x^*)\) is residual in \((\mathcal{X}_T, \rho)\). If \(g \in G(x,x^*)\), then for each \(n \in N\) there exists a \(x_n \in B_{1/n}(x)\) and \(x_n^* \in \partial_- g(x_n)\) so that \(\|x_n^*|_{X_n} - x^*|_{X_n}\| \leq 4/n\). Since the subspaces \(X_n : n \in \mathbb{N}\) are monotonely increasing, we see that \(\{x_n^*\}\) converges to \(x^*\) pointwise on \(\bigcup_{n=1}^\infty X_n\). Moreover, since \(g\) is locally Lipschitz around \(x\), the sequence \(\{x_n^*\}\) is norm bounded, we have that \(\{x_n^*\}\) converges to \(x^*\) pointwise on \(\overline{\bigcup_{n=1}^\infty X_n} = X\). Hence \(\{(x_n,x_n^*)\}\) converges to \((x,x^*)\) in \(A \times X^*\), with \(A\) endowed with the norm topology and \(X^*\) with the weak$^*$-topology. However, as \(\text{Gr}(\partial_a g)\) is closed in \(A \times X^*\) we see that \(x^* \in \partial_a g(x)\).

(c) By Lemma 5.3.1, we may choose a countable set \(C \subset \text{Gr}(\partial_- f)\) such that \(\text{Gr}(\partial_a f) = \text{cl} C\). Let \(G := \bigcap \{G(x,x^*) : (x,x^*) \in C\}\). By (b), \(G\) is a residual set in \((\mathcal{X}_T, \rho)\). If \(g \in G\), then for every \((x,x^*) \in C\) we have \(x^* \in \partial_a g(x)\). That is, \(C \subset \text{Gr}(\partial_a g)\). Since \(\text{Gr}(\partial_a g)\) is closed
in the product topology on $A \times X^*$ we have $\text{Gr}(\partial_a f) = \text{cl} C \subseteq \text{Gr}(\partial_ag)$. This shows that if $g \in G$ then $\partial_a f(x) \subseteq \partial_ag(x)$ for all $x \in A$.

\[ \square \]

**Corollary 5.6.2** Let $\{f_i\}$ be a sequence of locally equi-Lipschitz functions on an open set $A$ of a separable Banach space $X$. Define $\Omega(x) := \bigcup_{i=1}^{\infty} \partial_a f_i(x)$ for every $x \in A$ and $T := \text{CSC}(\Omega)$. In $(X_T, \rho)$, the set

$$\{g \in X_T : \text{USC}(\Omega)(x) \subseteq \partial_a g(x) \subseteq \partial_c g(x) \subseteq \text{CSC}(\Omega)(x) \text{ for every } x \in A\}$$

is residual. In particular, in $(X_T, \rho)$ the set $\{g \in X_T : \partial_c g(x) = \text{CSC}(\Omega)(x) \text{ for every } x \in A\}$ is residual.

**Proof.** Since each $f_i \in X_T$, we may apply Theorem 5.6.1 to obtain that the set

$$G_i := \{g \in X_T : \partial_a f_i(x) \subseteq \partial_a g(x) \text{ for all } x \in A\}$$

is residual, thus the set $G := \bigcap_{i=1}^{\infty} G_i$ is residual in $(X_T, \rho)$. If $g \in G$, for each $i$ we have $\partial_a f_i(x) \subseteq \partial_a g(x)$ for every $x \in A$, then $\Omega(x) = \bigcup_{i=1}^{\infty} \partial_a f_i(x) \subseteq \partial_a g(x)$ for every $x \in A$. Hence $\text{USC}(\Omega) \subseteq \partial_a g \subseteq T$. Since $T$ is convex-valued, the proof is complete by using $\partial_c g(x) = \overline{\text{conv}}(\partial_a g(x))$.

\[ \square \]

**Corollary 5.6.3** If $X$ is separable, then every $w^*$-compact convex subset $C \subset X^*$ is an approximate subdifferential map.

**Proof.** Since $X$ is separable, we may choose a countable set $\{c_i\}$ being $w^*$-dense in $C$. With $f_i$ defined by $f_i(x) := \langle c_i, x \rangle$ for $x \in A$, we may apply Corollary 5.6.2.

\[ \square \]

**Example 5.6.4** Theorem 5.6.1 fails if $T$ is only assume to be a $w^*$-usco. Define an usco map $T : \mathbb{R} \to 2^{\mathbb{R}}$ by $T(x) := \{0, 1\}$. Suppose Theorem 5.6.1 holds for this $T$ then there would exist a residual set in $X_T$ where $\partial_a g(x) \equiv \{0, 1\}$ for all $x \in A$. However $T$ is not an approximate subdifferential map of any Lipschitz function. Indeed, if there exists $f$ with $\partial_a f \equiv T$, then $\partial_c f \equiv \text{co}[T] = [0, 1]$. By Theorem 2.2 [19], $\partial_a f \equiv \partial_c f = [0, 1]$, a contradiction. However, $[0, 1]$ is always an approximate subdifferential map.

**Corollary 5.6.5** Suppose that $A$ is an open subset of a separable Banach space $X$ and that $f : A \to \mathbb{R}$ is locally Lipschitz with $\partial_a f = \partial_c f$. Then for every $w^*$-compact convex subset $C \subset X^*$, $\partial_a f + C$ is an approximate subdifferential map.
Proof. Choose a countable \( \text{w}^* \)-dense subset \( \{ c_i \} \) of \( C \). Define \( f_i \) by \( f_i(x) = f(x) + \langle c_i, x \rangle \).

For every \( x \in A \), we have
\[
\Omega(x) = \bigcup_{i=1}^{\infty} \partial_x f_i(x) = \bigcup_{i=1}^{\infty} \partial_x f(x) + c_i = \partial_x f(x) + \{ c_i \}_{i=1}^{\infty} \subset \partial_x f(x) + C.
\]

Since USC(\( \Omega(x) \)) \( \supset \) USC(\( \Omega(x) \)) is \( \text{w}^* \)-closed, we have USC(\( \Omega(x) \)) \( \supset \) \( \partial_x f(x) + C \).

As \( \partial_x f + C \) is a \( \text{w}^* \)-cuso, USC(\( \Omega(x) \)) \( \supset \) \( \partial_x f(x) + C \). Hence USC(\( \Omega(x) \)) = USC(\( \Omega(x) \)) = \( \partial_x f + C \).

Apply Corollary 5.6.2.

One may reformulate Theorem 5.6.1 in separable subspaces of a smoothable space:

**Theorem 5.6.6** Let \( A \) be a non-empty open subset of a smoothable Banach space \( X \) and let \( Y \) be a separable subspace of \( X \). Let \( T : A \rightarrow 2^X \) be a \( \text{w}^* \)-cuso on \( A \). Then for each \( f \in X_T \) with \( x^* \in \partial_x f(x) \), the set
\[
\{ g \in X_T : \text{there exists } y^\ast \in \partial_g(x), (y^\ast - x^*)|_Y = 0 \}
\]
is residual in \( (X_T, \rho) \).

Proof. Let \( F \) be a finite dimensional subspace \( Y \). We define
\[
G^{F}_{(x,x^*,m)} := \{ g \in X_T : \text{for some } z \in B_{1/m}(x), z^* \in \partial_g(z), \| (z^* - x^*)|_F \| < 4/m \}.
\]

Following (a) in the proof of Theorem 5.6.1, int\( G^{F}_{(x,x^*,m)} \) is dense in \( (X_T, \rho) \). If \( g \in G^{F}_{(x,x^*)} := \bigcap_{m=1}^{\infty} G^{F}_{(x,x^*,m)} \), then for every \( m \) we have some \( z^*_m \in \partial_g(z_m) \) with \( \| z_m - x \| < 1/m \) such that \( \| (z^*_m - x^*)|_F \| < 1/m \). As \( X \) is smoothable, \( X^* \) has a \( \text{w}^* \)-sequentially compact dual ball. As \( g \) is locally Lipschitz, \( \{ z^*_m \} \) is bounded, thus has a subsequence \( \{ z^*_{m_k} \} \) \( \text{w}^* \)-converging to some \( z^* \in \partial_g(x) \) Then \( (z^* - x^*)|_F = 0 \). Now take a increasing sequence of finite dimensional subspace such that \( Y = \text{cl} \bigcup_{n=1}^{\infty} F_n \). Let \( g \in G := \bigcap_{m=1}^{\infty} G^{F}_{(x,x^*)} \). For each \( F_n \), we get \( (y^*_n - x^*)|_{F_n} = 0 \) with \( y^*_n \in \partial_g(x) \). Again \( \{ y^*_n \} \) has a subsequence \( \{ y^*_n \} \) \( \text{w}^* \)-converging to some \( y^* \in \partial_g(x) \). As \( y^*_n|_{F_n} = x^*|_{F_n} \) and \( F_n : n \in \mathbb{N} \) is monotonically increasing, we have \( y^*|_{F_n} = x^*|_{F_n} \) for every \( n \), thus \( (y^* - x^*)|_Y = 0 \). \( \square \)

5.7 The subdifferentiability of permutation-invariant functions

Let \( P(n) \) denote the group of \( n \times n \) permutation matrices and \( X \) denote a finite-dimensional real inner-product space. A function \( f : X \rightarrow (-\infty, +\infty] \) is permutation invariant if \( f(Px) = f(x) \) for all \( x \in X \) and \( P \in P(n) \).
Proposition 5.7.1 [63] If \( f : X \to (-\infty, +\infty] \) is permutation invariant and lower semi-continuous, then any point \( x \in X \) and any \( P \in P(n) \) satisfy \( \partial f(Px) = P\partial f(x) \) where \( \| \cdot \| \) stands for \( -, a \). If \( f \) is Lipschitz around \( x \), the corresponding result holds for \( \partial_c \). Also \( f \) is regular at \( Px \) if and only if it is regular at \( x \).

Let \( S(n) \) denote the Euclidean space of \( n \times n \) real symmetric matrices and \( O(n) \) denote the group of \( n \times n \) real orthogonal matrices. The eigenvalue map \( \lambda : S(n) \to \mathbb{R}^n \) is defined by \( \lambda(X) := (\lambda_1(X), \lambda_2(X), \ldots, \lambda_n(X)) \) with the eigenvalues \( \lambda_k(X) \) in nonincreasing order. An eigenvalue function is an extended-real-valued function on \( S(n) \) of the form \( f \circ \lambda \) for a function \( f : \mathbb{R}^n \to (-\infty, +\infty], \) where \( f \) is permutation invariant. Lewis [63] has shown that the subgradients of the underlying function \( f \) characterize the subgradients of the matrix function \( f \circ \lambda \). More precisely, for a locally Lipschitz eigenvalue function \( f \circ \lambda \), at \( X \in S(n) \) one has

\[
\partial f \circ \lambda)(X) = \{ U^t \text{Diag}(\partial f(\lambda(X))) U : U \in O(n)^X \},
\]

where \( O(n)^X := \{ U \in O(n) : U^t \text{Diag}(\lambda(X)) U = X \} \) and \( \| \cdot \| \) stands for \( -, a, c \). It is instructive to consider the subdifferentiability of permutation invariant functions.

Theorem 5.7.2 Let \( \{ f_j \}_{j=1}^\infty \) be locally equi-Lipschitz and permutation invariant on \( \mathbb{R}^n \). Define \( \Omega \) by \( \Omega(x) := \bigcup_{j \in \mathbb{N}} \{ \partial_c f_j(x) \} \) for every \( x \in \mathbb{R}^n \). Then there exists a permutation invariant and locally Lipschitz \( f : \mathbb{R}^n \to \mathbb{R} \) such that \( \partial_c f(x) = \text{CSC}(\Omega)(x) \) for every \( x \in \mathbb{R}^n \).

To prove this, we consider the complete metric space:

\[
\{ f : f \in X_T, \ f \text{ is permutation invariant} \},
\]

with metric \( \rho_1(f, g) := \min\{1, \| f - g \|_\infty \} \) and \( T := \text{CSC}(\Omega) \). Since the class of permutation invariant functions is closed under lattice operations, the arguments of Theorem 5.2.2 still apply. However, we can not answer the following: Let \( C \) be a compact convex set and \( PC = C \) for every \( P \in P(n) \). Can one find a permutation invariant Lipschitz function \( f : \mathbb{R}^n \to \mathbb{R} \) such that \( \partial_c f \equiv C \)? The technical difficulty is: translation does not preserve the permutation-invariant property.
5.8 The subdifferentiability properties of $BC(A)$

The results we have developed have been for locally Lipschitz functions; it is natural to ask to what extent it would be possible to treat more general functions. For any topological space $S$, the space $BC(S)$, which denotes all bounded continuous functions on $S$ with norm given by $\|f\| := \sup\{|f(s)| : s \in S\}$, is a Banach space [60, page 70]. Assume $A$ is an open set in the separable Banach space $X$. For each continuous function $f$ on $A$ and $x \in A$, the set $\partial_c f(x) \subset X^*$ is a $w^*$-closed convex set (possibly empty). However, as is well-known:

**Proposition 5.8.1** Let $X$ be a Banach space and let $f : X \to (-\infty, +\infty]$ be lower semicontinuous. Then $f$ is Rockafellar-Clarke subdifferentiable at a dense subset of points in its graph.

**Proof.** Let $\epsilon > 0$ and $x_0$ with $f(x_0)$ finite be given. Since $f$ is lower semicontinuous one can choose $\epsilon > \lambda > 0$ such that $\inf\{f(x) : \|x - x_0\| \leq \lambda\} > f(x_0) - \epsilon$. Then $f_\lambda := f + \delta_{B_\lambda[x_0]}$ is lower semicontinuous and $\inf_X f_\lambda + \epsilon > f_\lambda(x_0)$. By the Ekeland variational principle, there exists $v$ such that

$$f_\lambda(v) \leq f(x_0), \quad \|v - x_0\| \leq \lambda/2,$$

(5.13)

$$f_\lambda(w) + \frac{2\epsilon}{\lambda}\|w - v\| > f_\lambda(v) \quad \text{for all } w \neq v.$$  

(5.14)

Equation (5.14) yields $0 \in \partial_c f_\lambda(v) + 2\epsilon/\lambda B_X$, by the sum rule [33, page 105]. As $\|v - x_0\| < \lambda$, $f_\lambda$ and $f$ agree in a neighbourhood of $v$, $\partial_c f_\lambda(v) = \partial_c f(v)$, thus $\partial_c f(v) \neq \emptyset$. Moreover, equation (5.13) shows $f(x_0) - \epsilon < f(v) = f_\lambda(v) \leq f_\lambda(x_0)$. Thus we obtain a point $(v, f(v))$ in the graph of $f$, which is arbitrarily close to the original point, where $\partial_c f(v) \neq \emptyset$. □

**Proposition 5.8.2 (Rockafellar)** [80] Suppose $f : \mathbb{R}^n \to (-\infty, +\infty]$ is directionally Lipschitz at $x$ with respect to at least one $y$. Then the multifunction $\partial_c f$ is closed at $x$, in the sense that: $x_k \to f(x), z^k \in \partial_c f(x_k), z^k \to z \implies z \in \partial_c f(x)$.

Proposition 5.8.2 can fail in some other cases as the following example [80] illustrates:

**Example 5.8.3** Define $S := \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_3 = x_1x_2 \text{ or } x_3 = -x_1x_2\}$. Now

$$\partial_c I_S(0, 0, 0) = \{(0, 0, x_3) : x_3 \in \mathbb{R}\} \text{ and } \partial_c I_S(x_1, 0, 0) = \{(0, x_2, x_3) : x_2, x_3 \in \mathbb{R}\},$$

for each $x_1 \neq 0$. Hence $\partial_c I_S$ is not closed at $(0, 0, 0)$ in the sense of Proposition 5.8.2.
5.8 The subdifferentiability properties of $BC(A)$

For a closed set $S \subset \mathbb{R}^n$ containing $x$, if for some neighbourhood $U$ of $x$ the set $S \cap U$ is the epigraph of a locally Lipschitz function then $\partial_x I_S$ is closed at $x$ [33, page 69]. Nevertheless, Theorem 5.8.6 below shows that the multifunction $\partial_c f$ is closed for most continuous functions even though they are nowhere directionally Lipschitz.

Suppose $x \in \text{cl} \, S$. The $G$-normal cone to $S$ at $x$ is

$$N_G(S,x) := w^* \text{cl} \bigcup \{ \lambda \partial_a d_S(x) : \lambda > 0 \}.$$  

For a function $f$ the $G$-subdifferential and singular $G$-subdifferential of $f$ at $x \in \text{dom} \, f$ are

$$\partial_y f(x) := \{ x^* : (-1, x^*) \in N_G(e^p f, (f(x), x)) \},$$

$$\partial_y^\infty f(x) = \{ x^* : (0, x^*) \in N_G(e^p f, (f(x), x)) \}.$$  

When $X$ is finite dimensional, $\partial_y f$ and $\partial_a f$ are the same.

**Proposition 5.8.4** [58]Let $f$ be lower semicontinuous on an open subset $A$ of a Banach space $X$. Then $\partial_y f(x) \subset \partial_a f(x)$. If $f$ is directionally Lipschitz at $x$, then $\partial_y f(x) = \partial_a f(x)$.

In general, $\partial_c f(x) = \overline{\text{co}} \, \overline{w^*} \, (\partial_y f(x) + \partial_y^\infty f(x))$.

The following characterization of $\partial_y f$ by Borwein and Zhu [12, 4] will help us to compute $\partial_y^\infty$:

**Proposition 5.8.5** Let $X$ be a smoothable Banach space and $f$ be a lower semicontinuous proper function on $X$. Then for any $x \in X$

$$\partial_y f(x) = \overline{w^*} \text{cl} \bigcup_{k=1}^\infty \{ w^* - \lim_{n \to \infty} x_n^* : x_n^* \in D_y^k f(x_n), x_n \to_f x \},$$  

where $D_y^k f(x)$ is the subset $D_y f(x)$ for which the support function in the definition has a Lipschitz constant no greater than $k$.

We are ready for the highlight of this section: the ‘dual space case’ theorem. It essentially says $\partial_y$ can be large even if we take the norm sequential limits of the viscosity subderivatives first then take the $w^*$-closure of these limits in equation (5.15).

**Theorem 5.8.6 (The dual space case)** If $A$ is an open subset of a separable Banach space $X$, then in $BC(A)$ the set \{ $f \in BC(A) : \partial_c f = \partial_y f = \partial_a f = X^*$ on $A$ $\}$ is residual.
Lemma 5.8.7 Assume $X$ is separable and $C \subset X^*$ is $w^*$-compact convex with non-empty norm interior. Then the set \( \{ f \in BC(A) : C \subset \partial_g f(x) \subset \partial_a f(x) \text{ for every } x \in A \} \) is residual in $BC(A)$.

Proof. Fix $x \in A$, $c \in \text{int}(C)$, $n > 1$. Consider $G^n_{x,c} := \{ f \in BC(A) : \text{there exists } \tilde{x} \in A \text{ satisfying } \| \tilde{x} - x \| < \frac{1}{n} \text{ such that for some } \tilde{x}^* \in \partial f(\tilde{x}) \text{ we have } \| \tilde{x}^* - c \| < \frac{1}{n} \}$. 

(a) $\text{int}(G^n_{x,c})$ is dense in $BC(A)$. Choose an arbitrary $f \in BC(A)$ and $\epsilon > 0$. Define 
\[ h(\tilde{x}) := f(x) - \epsilon + \sigma_C(\tilde{x} - x), \quad h_1 := \min\{ f, h \} \quad \text{and} \quad h_2 := \max\{ f - 2\epsilon, h_1 \}. \]

Clearly $h_2$ is continuous on $A$. Since $\| h_2 - f \| < 3\epsilon$ and $f$ is bounded on $A$, $h_2 \in BC(A)$. For $0 < \delta < 1/n$ sufficiently small, we have 
\[ h_2(\tilde{x}) = h(\tilde{x}) = f(x) - \epsilon + \sigma_C(\tilde{x} - x) \quad \text{on } B_\delta[x]. \]

As $h$ is convex on $B_\delta[x]$, $\partial h_2(x) = C$, we see that $h_2 \in \text{int}(G^n_{x,c})$. We proceed to show that $h_2 \in \text{int}(G^n_{x,c})$. Because $c \in \text{int}(C)$, we may assume $rB_X^* \subset C - c$ for some $r > 0$, then for $\tilde{x} \neq x$ we have $\sigma_C(\tilde{x} - x) > 0$ by the Hahn-Banach theorem. On $B_\delta(x)$, for $\tilde{x} \neq x$ we have 
\[ h_2(\tilde{x}) - \langle c, \tilde{x} \rangle = f(x) - \epsilon + \sigma_C(\tilde{x} - x) - \langle c, \tilde{x} \rangle = h_2(x) - \langle c, x \rangle + \sigma_C(\tilde{x} - x), \quad \text{and so} \]
\[ m := \inf \{ h_2(\tilde{x}) - \langle c, \tilde{x} \rangle : \| \tilde{x} - x \| = \delta \} = h_2(x) - \langle c, x \rangle + \inf \{ \sigma_C(\tilde{x} - x) : \| \tilde{x} - x \| = \delta \}
\]
\[ \geq h_2(x) - \langle c, x \rangle + r\delta > h_2(x) - \langle c, x \rangle. \]

Let $\alpha := m - (h_2(x) - \langle c, x \rangle) > 0$. For $g \in BC(A)$ with $\| g - h_2 \| < \beta < \alpha/2$, for $\| \tilde{x} - x \| = \delta$ we have $g(\tilde{x}) - \langle c, \tilde{x} \rangle = g(\tilde{x}) - h_2(\tilde{x}) + h_2(\tilde{x}) - \langle c, \tilde{x} \rangle \geq -\beta + m$, and 
\[ g(x) - \langle c, x \rangle = g(x) - h_2(x) + h_2(x) - \langle c, x \rangle \leq \beta + h_2(x) - \langle c, x \rangle. \]

Then 
\[ \inf \{ g(\tilde{x}) - \langle c, \tilde{x} \rangle : \| \tilde{x} - x \| = \delta \} \geq -\beta + m > \beta + h_2(x) - \langle c, x \rangle \geq g(x) - \langle c, x \rangle. \quad (5.16) \]

Define $g_1 := g - \langle c, \cdot \rangle + I_{B_\delta[x]}$, which is lower semicontinuous and bounded below on $X$. By equation (5.16) 
\[ \inf_{X} g_1 = \inf_{B_\delta[x]} (g(\tilde{x}) - \langle c, \tilde{x} \rangle) \leq g(x) - \langle c, x \rangle < \inf \{ g(\tilde{x}) - \langle \tilde{x}, c \rangle : \| \tilde{x} - x \| = \delta \}. \]
5.8 The subdifferentiability properties of $BC(A)$

Let $0 < \nu < \min\{1/(2n), \inf\{g(\bar{x}) - \langle \bar{x}, c \rangle : \|\bar{x} - x\| = \delta\} - (g(x) - \langle x, c \rangle)\}$. Choose $x_0$ such that

$$g_1(x_0) = g(x_0) - \langle c, x_0 \rangle < \inf_X g_1 + \nu.$$ 

Since $X$ is separable, $X$ is smoothable. With $\lambda = 1$, Theorem 2.7.3 shows there exists a Gâteaux differentiable $\phi$ and $v \in X$ such that for all $\bar{x} \in B_\delta[x]$,

$$g(\bar{x}) - \langle c, \bar{x} \rangle + \phi(\bar{x}) \geq g(v) - \langle c, v \rangle + \phi(v),$$

$$g(v) - \langle c, v \rangle < \nu + \inf_X g_1,$$  \hspace{1cm} (5.17)

$$\|\nabla \phi(v)\| < 2\nu < \frac{1}{n}.$$ \hspace{1cm} (5.18)

Equation (5.18) shows $g(v) - \langle c, v \rangle < \inf\{g(\bar{x}) - \langle \bar{x}, c \rangle : \|\bar{x} - x\| = \delta\}$, thus $\|v - x\| < \delta < 1/n$. Equation (5.17) shows $0 \in \partial_g g(v) - c + \nabla \phi(v)$, thus there exist $v^* \in \partial_g g(v)$ with $\|v^* - c\| = \|\nabla \phi(v)\| < 1/n$ by equation (5.19). Then $g \in G^a_{x,c}$, and so $B_\beta(h_2) \subset G^a_{x,c}$, that is, $h_2 \in \text{int}(G^a_{x,c})$.

(b) Since $\text{int}(G^a_{x,c})$ is open and dense in $BC(A)$, the set $G_{x,c} := \bigcap_{n=1}^\infty \text{int}(G^a_{x,c})$ is dense in $BC(A)$. If $f \in G_{x,c}$, then for every $n$ there exists $x_n^* \in \partial_\alpha f(x_n)$ with $\|x_n^* - c\| < 1/n$ and $\|x_n - x\| < 1/n$. Letting $n \to \infty$, we obtain $c \in \partial_\alpha f(x)$ by equation (1.1).

(c) By Lemma 5.4.1 we may take a countable dense set $\{c_k\}$ from $\text{int}(C)$ satisfying equation (5.7). As $G_{x,c_k}$ is dense $G_\delta$ for each $c_k$, we have $G_x := \bigcap_{k=1}^\infty G_{x,c_k}$ is a dense $G_\delta$ set in $BC(A)$. If $f \in G_x$, then $c_k \in \partial_\alpha f(x)$ for every $k$ by (b), thus $C \subset \partial_\alpha f(x)$ as $\partial_\alpha f(x)$ is $w^*$-closed.

(d) Now let $\{x_k\}$ be a countable norm dense set in $A$. Since $G_{x_k}$ is dense $G_\delta$ for every $k$, the set $G := \bigcap_{k=1}^\infty G_{x_k}$ is a dense $G_\delta$ set in $BC(A)$. If $f \in G$, for every $k$ we have $\partial_\alpha f(x_k) \supseteq C$ by (c). By $w^*$-upper semicontinuity of $\partial_\alpha f$, we get $\partial_\alpha f(x) \supseteq \limsup_{x_k \to x} \partial_\alpha f(x_k) \supseteq C$. Hence $\partial_\alpha f \supseteq C$ for every $f \in G$.

(e) Finally we compute $\partial_\alpha f$ for $f \in G$. Equation (5.17) in fact shows that $c - \nabla \phi(v) \in D_g f(v)$, i.e., the $g$-viscosity subderivatives. From (b) we see that if $f \in G_{x,c}$ then for every $n$ there exists $x_n^* \in D_g f(x_n)$ with $\|x_n^* - c\| < 1/n$ and $\|x_n - x\| < 1/n$. In other words, for every neighbourhood $U$ of $x$, for every $\epsilon > 0$ there exists $\|x_\epsilon - x\| < \epsilon$ and $x_\epsilon^* \in D_g f(x_\epsilon)$ such that $\|x_\epsilon^* - c\| < \epsilon$. From (c) we see that if $f \in G_x$ then for every $c_k$ this holds. From (d) we see that this holds for all $\{c_m\}$ at every $x_k$ if $f \in G$. 
5.8 The subdifferentiability properties of $BC(A)$

Let $f \in G$. Fix $c_m$ and $x$. For every $B_1/n(x)$ there exists $\|x_k - x\| < 1/n$ as $\{x_k\}_{k=1}^\infty$ is dense in $A$. As $f \in G_{x_k,c_m}$, there exists $\|v_n - x_k\| < 1/n - \|x_k - x\|$ and $v_n^* \in D_g f(v_n)$ such that $\|v_n^* - c_m\| < 1/n$. That is, there exists $v_n$ and $v_n^*$ such that $v_n^* \in D_g f(v_n)$ and $\|v_n^* - c_m\| < 1/n$. Letting $n \to \infty$, we obtain $c_m \in \partial g f(x)$ by Proposition 5.8.5. As $m$ is arbitrary, $\bigcup_{m=1}^\infty \{c_m\} \subset \partial g f(x)$. Because $\{c_m\}_{m=1}^\infty$ is $w^*$-dense in $C$ and $\partial g f(x)$ is $w^*$-closed, we obtain $C \subset \partial g f(x)$. 

**Proof of Theorem 5.8.6.** By Lemma 5.8.7, for each $n$ the set

$$G_n := \{f \in BC(A) : \partial a f \supset \partial g f \supset nB_{X^*}\},$$

is residual in $BC(A)$. The set $G := \bigcap_{n=1}^\infty G_n$ is a residual set in $BC(A)$. If $f \in G$, for every $n$ we have $\partial a f(x) \supset \partial g f(x) \supset nB_{X^*}$ for $x \in A$, then $\partial a f(x) \supset \partial g f(x) \supset \bigcup_{n=1}^\infty nB_{X^*} = X^*$. The proof is complete by using Proposition 5.8.4. 

By Theorem 2.9.1 [33], if $\partial c f(x) \neq \emptyset$ we have $f^\dagger(x; v) = \sup \{\langle \xi, v \rangle : \xi \in \partial c f(x)\}$. When $\partial c f(x) = X^*$, we obtain $f^\dagger(x; 0) = 0$ and $f^\dagger(x; v) = +\infty$ for nonzero $v \in X$. Hence, Theorem 5.8.6 shows:

**Corollary 5.8.8** Let $A$ be an open subset of a separable Banach space $X$. Then the set

$$\{f \in BC(A) : f^\dagger(x; v) = +\infty \text{ for } v \neq 0 \text{ and } f^\dagger(x; 0) = 0 \text{ on } A\},$$

is residual in $BC(A)$.

Let $K \subset X$ be a closed convex cone. We say that $f$ is nondecreasing with respect to $K$ if $f(x') \leq f(x'')$ when $x'' - x' \in K$. We use $BC_K(A)$ to denote bounded continuous functions on $A$ which are $K$-nondecreasing. $BC_K(A)$ is a complete metric space and is closed under lattice operations. Define

$$K^+ := \{x^* \in X^* : \langle x^*, x \rangle \geq 0 \text{ for all } x \in K\}.$$

If $f \in BC_K(A)$, from equations (1.3) and (1.1) we see that $\partial c f(x) \subset K^+$ and $\partial a f(x) \subset K^+$. 

**Theorem 5.8.9** If $K^+$ has nonempty norm interior in the dual of a separable Banach space $X$, then the set $G := \{f \in BC_K(A) : \partial g f = \partial a f = \partial c f = K^+ \text{ on } A\}$ is residual in $BC_K(A)$. If, in addition, $K$ has non-empty interior, then $\partial c (-f) \equiv -K^+$ for $f \in G$. 
5.9 The subdifferentiability properties of $C(A)$

Proof. Exactly as in Lemma 5.8.7 we can argue as follows: If $C \cap K^+$ is a $w^*$-compact convex set with non-empty norm interior, then the set

$$\{ f \in BC_K(A) : C \cap K^+ \subset \partial_g f(x) \subset \partial_a f(x) \text{ for every } x \in A \},$$

is residual in $BC_K(A)$. Since $K^+$ has nonempty norm interior, $K^+ \cap nB_X^*$ has nonempty norm interior, then each $G_n := \{ f \in BC_K(A) : (K^+ \cap nB_X^*) \subset \partial_g f \subset \partial_a f \}$ is residual in $BC_K(A)$. If $f \in G := \bigcap_{n=1}^\infty G_n$, then $K^+ \subset \partial_g f \subset \partial_a f$, and so $\partial_c f = \partial_a f = \partial_g f \equiv K^+$. When $K$ has non-empty interior, $f$ is directionally Lipschitz, thus $\partial_c (-f) = -\partial_c f$ [81]. □

When $K = 0$, Theorem 5.8.9 recovers Theorem 5.8.6.

5.9 The subdifferentiability properties of $C(A)$

In this section, our setting is an abstract finite-dimensional normed space $X$. We will show that the Clarke and approximate subdifferentials of most continuous functions on $X$ are the dual space $X^*$. We start with some general definitions.

Definition 5.9.1 A locally convex space $X$ where the locally convex topology is generated by a complete invariant metric is called a Fréchet space.

A linear topological space is said to be normable if the linear topology is generated by a norm.

Proposition 5.9.2 [49] A locally convex space $X$ is metrizable if and only if it is separated and the topology has a countable local base. The metric generating the locally convex topology may be taken to be an invariant metric.

Definition 5.9.3 Suppose that $d_i (i \in I)$ is a family of pseudo-metrics on a set $S$. The topology `induced by the $d_i$' is obtained by letting the neighbourhood of $s$ be sets that contains a set of the form $\{ t : d_i(s, t) < \alpha \text{ for each } i \in F \},$ where $\alpha > 0$ and $F$ is a finite subset of $I$.

Consider $C(X)$, continuous real functions on a topological space $X$ with the compact-open topology; that is, for each $K$ compact we define the semi-norm

$$p_K(f) := \sup\{|f(t)| : t \in K\},$$
and consider the topology generated by the family of such semi-norms for all compact set \( K \subset X \). When \( X \) is the union of a sequence of compact sets \( K_n \), the locally convex topology is separated and since there is a countable local base it is metrisable by Proposition 5.9.2. (However the topology is not normable: If there were a norm defining the topology, the unit ball would contain a neighbourhood \( U \) of 0, where \( U := \{ f \in C(X) : p_{K_n}(f) < \epsilon \} \) for some \( n \) and \( \epsilon > 0 \). But this implies that if \( p_{K_n}(f) = 0 \) then \( \| f \| = 0 \) and so \( f = 0 \). However, for each \( n \in \mathbb{N} \) there exists an \( f \in C(X) \) such that \( p_{K_n}(f) = 0 \) but \( f \neq 0 \).)

Assume \( X \) is a finite-dimensional normed space whose topological dual is denoted by \( X^* \). Let \( A \subset X \) be non-empty and open. With \( K_n \subset A \) being compact and having non-empty interior, \( A = \bigcup_{n=1}^{\infty} K_n \) and \( \rho_n := p_{K_n} \), on \( C(A) \) we define a metric \( \rho \) given by

\[
\rho(f, g) := \sum_{n=1}^{\infty} \frac{1}{2^n} \min\{\rho_n(f-g), 1\},
\]

The metric topology of \( \rho \) is the locally convex topology induced by \( \{p_{K_n}\}_{n=1}^{\infty} \). Then \( (C(A), \rho) \) is a Fréchet space.

**Theorem 5.9.4** Assume \( A \) is a non-empty open subset of a finite-dimensional normed space \( X \). Then in \( (C(A), \rho) \), the set \( \{ f \in C(A) : \partial_c f = \partial_\alpha f \equiv X^* \} \) is residual.

For \( c \in X^* \), we define \( D_f^c := \{ x \in A : f(y) - \langle c, y \rangle \) attains a local minimum at \( x \} \).

**Lemma 5.9.5** For each \( c \in X^* \) the set \( \{ f \in C(A) : D_f^c \text{ is dense in } A \} \) is residual in \( (C(A), \rho) \).

**Proof.** Fix \( x \in A \) and choose \( n_0 \) such that \( x \in \text{int}(K_{n_0}) \). For every \( n \) we define \( G_n^{x,c} := \{ f \in C(A) : \text{there exists } \hat{x} \in A \text{ satisfying } \| \hat{x} - x \| < \frac{1}{n} \text{ such that } f - \langle c, \cdot \rangle \text{ attains a local minimizer at } \hat{x} \} \).

(a) \( \text{int}(G_n^{x,c}) \) is dense. Choose an arbitrary \( f \in C(A) \) and \( 1/3 > \epsilon > 0 \). Choose \( C \subset X^* \) such that \( c \in \text{int}(C) \). Define \( h(\hat{x}) := f(x) - \epsilon + \sigma_C(\hat{x} - x) \),

\[
h_1 := \min\{ f, h \} \text{ and, } \quad h_2 := \max\{ f - 2\epsilon, h_1 \}.
\]

Then \( h_1, h_2 \in C(A) \). For \( 0 < \delta < 1/n \) sufficiently small, we have \( B_\delta[x] \subset K_{n_0} \) and

\[
h_2 = h = f(x) - \epsilon + \sigma_C(\hat{x} - x) \text{ on } B_\delta[x].
\]
5.9 The subdifferentiability properties of $C(A)$

As $0 \in C - c$, we have

$$h_2(\bar{x}) - \langle c, \bar{x} \rangle = h_2(x) - \langle c, x \rangle + \sigma_{C-c}(\bar{x} - x) \geq h_2(x) - \langle c, x \rangle \quad \text{for every } \bar{x} \in B_\delta[x],$$

thus $h_2 \in G^n_{x,c}$. Moreover, $f - 3\epsilon < h_2 < f + 3\epsilon$ on $A$, thus $\rho(h_2, f) < 3\epsilon$. We proceed to show that $h_2 \in \text{int}(G^n_{x,c})$. Because $c \in \text{int}(C)$, we may assume $rB_{X^*} \subset C - c$ for some $r > 0$, then

$$m := \inf \{h_2(\bar{x}) - \langle c, \bar{x} \rangle : \|\bar{x} - x\| = \delta \} = h_2(x) - \langle c, x \rangle + \inf \{\sigma_{C-c}(\bar{x} - x) : \|\bar{x} - x\| = \delta \} \geq h_2(x) - \langle c, x \rangle + r\delta > h_2(x) - \langle c, x \rangle.$$

Let $\alpha := m - (h_2(x) - \langle c, x \rangle) > 0$ and $0 < \beta < \min\{\alpha/2, 1\}$. For each $g \in C(A)$ with $\rho(g, h_2) < \beta/2^{n_0}$ we have $\rho_{n_0}(g - h_2) < \beta$. When $\|\bar{x} - x\| = \delta$ we have $g(\bar{x}) - \langle c, \bar{x} \rangle = g(\bar{x}) - h_2(\bar{x}) + h_2(\bar{x}) - \langle c, \bar{x} \rangle \geq -\beta + m$, and

$$g(x) - \langle c, x \rangle = g(x) - h_2(x) + h_2(x) - \langle c, x \rangle \leq \beta + h_2(x) - \langle c, x \rangle.$$

Then

$$\inf \{g(\bar{x}) - \langle c, \bar{x} \rangle : \|\bar{x} - x\| = \delta \} \geq -\beta + m > \beta + h_2(x) - \langle c, x \rangle \geq g(x) - \langle c, x \rangle. \quad (5.20)$$

Because $h_2$ is bounded on $B_\delta[x]$, $g$ is bounded on $B_\delta[x]$. Then $g_1 := g - \langle c, \cdot \rangle + I_{B_\delta[x]}$ is lower semicontinuous and bounded below on $X$. By equation (5.20)

$$\inf_{\bar{x}} g_1 = \inf_{B_\delta[x]} \{g(\bar{x}) - \langle c, \bar{x} \rangle\} \leq g(x) - \langle c, x \rangle < \inf \{g(\bar{x}) - \langle \bar{x}, c \rangle : \|\bar{x} - x\| = \delta\}. \quad (5.21)$$

As $X$ is finite dimensional, $B_\delta[x]$ is compact, thus $g_1$ attains a local minimizer at some $v \in B_\delta[x]$. Equation (5.21) shows $\|v - x\| < \delta < 1/n$. In other words, $g - \langle c, \cdot \rangle$ attains a local minimizer at $v$ with $\|v - x\| < 1/n$, thus $g \in G^n_{x,c}$. Hence $\{g \in C(A) : \rho(g, h_2) < \beta/2^{n_0}\} \subset G^n_{x,c}$.

(b) Since $\text{int}(G^n_{x,c})$ is open and dense in $C(A)$, the set $G_{x,c} := \bigcap_{n=1}^{\infty} \text{int}(G^n_{x,c})$ is dense in $C(A)$. If $f \in G_{x,c}$, then for every $n$ there exists $\|x_n - x\| < 1/n$ such that $f - \langle c, \cdot \rangle$ attains a local minimizer at $x_n$. That is, $f - \langle c, \cdot \rangle$ attains a local minimizer in every neighbourhood of $x$.

(c) As $X$ is finite dimensional, we may choose a countable dense subset $\{x_k\}_{k=1}^{\infty}$ from $A$. Since $G_{x_k,c}$ is dense $G_\delta$ for every $k$, the set $G_c := \bigcap_{k=1}^{\infty} G_{x_k,c}$ is a dense $G_\delta$ set in $C(A)$. 

Let $f \in G_c$. For every $x \in A$ and every open neighbourhood $U$ of $x$, there exists an $x_k \in U$, then $U$ is a neighbourhood of $x_k$. As $f \in G_{x_k,c}$, $f - \langle c, \cdot \rangle$ attains a local minimum at some point in $U$. Since $U$ and $x$ are arbitrary, the set of points at which $f - \langle c, \cdot \rangle$ attains a local minimum is dense in $A$. \hfill \square

**Proof of Theorem 5.9.4.** If $f \in G_c$ then $c \in \partial_a f(x)$ for every $x \in A$. By taking a countable set $\{c_k\}_{k=1}^{\infty}$ dense in $X^*$ we obtain a residual set $G := \bigcap_{k=1}^{\infty} G_{c_k}$ such that if $f \in G$ then $c_k \in \partial_a f(x)$ for every $k$. As $\partial_a f(x)$ is closed we have $\partial_a f(x) = X^*$. Hence $\partial_a f = \partial_c f = X^*$. \hfill \square

**Definition 5.9.6** A function $f : A \to \mathbb{R}$ is called a dense minimizer function on $A$ if for every open subset $U$ of $A$ there exists $x \in U$ such that $f$ attains a local minimizer at $x$; $f$ is called a dense minimizer function of the second species if there exists a countable set $\{c_n\}$ being dense in $X^*$ such that $f(\cdot) - \langle c_n, \cdot \rangle$ is a dense minimizer function on $A$ for each $c_n$.

Dense minimizer functions of the second species are exactly the generalization of nowhere monotone functions of the second species on $\mathbb{R}$ to several dimensions. The proof of Theorem 5.9.4 actually showed that the set of dense minimizer functions of the second species is a residual set in $(C(A), \rho)$.

**Remark 5.9.7** The Corollary to Theorem 2.9.5 [33] says: “Let $X$ be finite dimensional, and suppose $\text{epi} f$ is locally closed near $(x, f(x))$. Then $f$ is directionally Lipschitz at $x$ if and only $\{v : f^*(x; v) < \infty\}$ has nonempty interior”. As a typical function $f$ in $C(A)$ satisfies: $f^0(x; v) = +\infty$ if $v \neq 0$ and $f^0(x; 0) = 0$, so $f$ is not directionally Lipschitz at any $x \in A$.

### 5.10 The viscosity subdifferentiability of $BC(A)$

In this section, our goal is to show that functions with large Clarke or approximate subdifferential are viscosity subdifferentiable only on a first category set.

**Theorem 5.10.1** Let $A$ be an open subset of a separable Banach space $X$. In $BC(A)$ the set

$$\{ f \in BC(A) : f \text{ and } -f \text{ are $g$-viscosity subdifferentiable at most on a first category subset of } A \},$$
is residual.

In order to prove this theorem, we need some definitions. Given \( f : X \to (-\infty, +\infty] \), for each \( n \geq 1 \) we define \( f_n : X \to \mathbb{R} \) by \( f_n(x) := \inf\{ f(y) + n\|y - x\| : y \in X \} \) for \( x \in X \).

**Definition 5.10.2** A function \( f \) defined on \( X \) is said to be lower semi-Lipschitz at \( x \in X \) if there is a neighborhood \( U \) of \( x \) and a constant \( L \) such that \( f(y) - f(x) \geq -L\|y - x\| \) for all \( y \in U \).

Fix \( x \in A \). Since \( f \) is lower semicontinuous, given \( \epsilon > 0 \) there exists a \( \delta > 0 \) such that \( f(x) \leq \epsilon + \inf\{ f(y) : \|y - x\| \leq \delta \} \). By Ekeland’s variational principle, there exists \( v \) such that \( \|v - x\| < \delta/2 \) and \( f(y) + \frac{2}{\delta}\|y - v\| \geq f(v) \) for all \( \|y - x\| \leq \delta \). Then \( f \) is lower semi-Lipschitz at \( v \). Hence every lower semicontinuous function is densely lower semi-Lipschitz on its domain.

**Lemma 5.10.3** [93] Let \( f \) be extended real-valued and bounded below on \( X \). Then the sequence \( \{f_n\} \) has the following properties:

(a) \( f_n \) is Lipschitz on \( X \) with Lipschitz constant \( n \);

(b) \( f_n(x) \leq f_{n+1}(x) \leq f(x) \), for \( x \in X \), \( n \geq 1 \);

(c) \( \lim_{n \to \infty} f_n(x) = f(x) \) if and only if \( f \) is lower semi-continuous at \( x \);

(d) \( f_n(x) = f(x) \) for some \( n \geq 1 \) if and only if \( f \) is locally lower semi-Lipschitz at \( x \);

(e) \( f \) is locally Lipschitz around \( x \) if and only if there exists a neighborhood \( U \) of \( x \) and an integer \( n \geq 1 \) such that \( f_n(x) = f(x) \) in \( U \).

**Proof.** (a), (b) and (e) are clear, we omit their proofs.

(c) Assume \( \lim_{n \to \infty} f_n(x) = f(x) \). By definition, \( f_n(x) \leq f(y) + n\|y - x\| \) for all \( y \in X \), then \( f_n(x) \leq \liminf_{y \to x} (f(y) + n\|y - x\|) = \liminf_{y \to x} f(y) \). Then \( f(x) = \lim_{n \to \infty} f_n(x) \leq \liminf_{y \to x} f(y) \), and so \( f \) is lower semicontinuous at \( x \). Conversely, assume \( f \) is lower semicontinuous at \( x \). For each \( n \), choose \( y_n \) such that \( f(y_n) + n\|y_n - x\| \leq f_n(x) + 1/n \). Assume \( f \) is bounded below by \( m \), then \( m + n\|y_n - x\| \leq f(x) + 1/n \) shows \( y_n \to x \). Then

\[
    f(x) \leq \liminf_{n \to \infty} f(y_n) \leq \lim_{n \to \infty} f_n(x) \leq f(x),
\]
as required.

(d) Let $f_n(x) = f(x)$ for some $n \geq 1$. From the definition of $f_n(x)$, we get $f(x) \leq f(y) + n\|y - x\|$ for all $y \in X$, thus $f$ is lower semi-Lipschitz at $x \in X$ with constant $n$.

Conversely, suppose that $f$ is locally lower semi-Lipschitz at $x$. If $f_n(x) < f(x)$ for all $n \in \mathbf{N}$, for each $n \geq 1$ we choose $x_n$ such that $f(x_n) + n\|x_n - x\| < f(x)$. Since $f(x)$ is finite and $f$ is bounded below on $X$, we obtain that $x_n \to x$, and $f(x_n) - f(x) < -n\|x_n - x\|$ for all $n \in \mathbf{N}$. This is impossible since $f$ is lower semi-Lipschitz at $x$. \qed

\textbf{Lemma 5.10.4} Assume that $f$ is lower semi-continuous and proper on Banach space $X$. If $f$ is nowhere locally Lipschitz on $X$, then $f$ is lower semi-Lipschitz at most on a first category set.

\textbf{Proof.} (a) In case that $f$ is bounded below on $X$, for each $n \in \mathbf{N}$ we define $D_n := \{x \in X : f_n(x) = f(x)\}$. By Lemma 5.10.3, each $D_n$ is closed subset of $X$ since $D_n = \{x \in X : f(x) - f_n(x) \leq 0\}$ and $f - f_n$ is lower semicontinuous. If $f$ is nowhere locally Lipschitz, then $D_n$ has no interior since $f_n$ is locally Lipschitz on int $D_n$. Then $F := \bigcup_{n=1}^{\infty} D_n$ is a first category $F_\sigma$ set of $X$.

(b) In case that $f$ is not bounded below on $X$, let $G_m := \{x \in X : f(x) > -m\}$ for $m \in \mathbf{N}$. The lower semi-continuity of $f$ implies that $G_m$ is an open set for $m \in \mathbf{N}$. For each $m \in \mathbf{N}$, we define

$$g_m(x) := \begin{cases} f(x) & \text{if } x \in G_m, \\ -m & \text{if } x \in \text{cl}[G_m] \setminus G_m, \\ +\infty & \text{otherwise}. \end{cases}$$

Clearly, $g_m$ is lower semi-continuous, bounded below, and nowhere locally Lipschitz. By (a), $g_m$ is semi-Lipschitz at most on a first category $F_\sigma$-set of $X$, say $F_m$. Then $F := \bigcup_{m=1}^{\infty} F_m$ is again a first category $F_\sigma$-set of $X$. The set of points at which $f$ is semi-Lipschitz is a subset of $F$, thus first category. \qed

If $f$ is $\beta$-viscosity subdifferentiable at $x$ then $f$ is lower semi-Lipschitz at $x$. Hence:

\textbf{Proposition 5.10.5} If $f$ is lower semicontinuous and nowhere locally Lipschitz on a $\beta$-smoothable Banach space, then $f$ is $\beta$-viscosity subdifferentiable at most on a first category set.
5.11 The subdifferentiability of $C$-Lipschitz extensions

Proof of Theorem 5.10.1. If $f$ is continuous and $\partial C f \equiv X^*$ on $A$, then $f$ and $-f$ are nowhere locally Lipschitz, thus the Gâteaux-viscosity subgradient exists at most on a first category subset of $A$ by Proposition 5.10.5. To finish the proof, we apply Theorem 5.8.6. □

Note that in finite dimensional spaces for a lower semicontinuous function, the Dini, Fréchet and viscosity subdifferentials coincide (see page 9). Hence Proposition 5.10.5 shows each nowhere monotone function of the second species on $\mathbb{R}$ is Dini subdifferentiable at most on a first category subset (see Theorem 3.8.3).

5.11 The subdifferentiability of $C$-Lipschitz extensions

Let us now explore the subdifferentiability of Lipschitz extensions. Assume $f$ is a nonexpansive Lipschitz function defined on a closed subset $A$ of a separable Banach space $X$. Our main result says if for some point $x \notin A$ there exist two non-expansive Lipschitz extensions $h_1$ and $h_2$ such that $h_1(x) \neq h_2(x)$ then most non-expansive Lipschitz extensions have large Clarke and approximate subdifferentials around $x$. When $A$ is a closed proper subspace of $X$, our result recovers Theorem 2.1 given by Borwein, Giles and Vanderwerff [6].

Lemma 5.11.1 Let $A \subset X$ be a non-empty set. Suppose $f$ is $C$-Lipschitz on $A$, that is,

$$f(x) - f(y) \leq \sigma_C(x - y) \text{ for all } x, y \in A.$$ 

Then there is a $C$-Lipschitz function $f_1$ on $X$ such that

$$f_1|_A = f \quad \text{and} \quad f_1(x) - f_1(y) \leq \sigma_C(x - y) \text{ for all } x, y \in X. \quad (5.22)$$

Moreover, the largest and smallest functions satisfying (5.22) are respectively given by

$$f_+(x) := \inf \{ f(y) + \sigma_C(x - y) : y \in A \}, \quad \text{and} \quad f_-(x) := \sup \{ f(y) - \sigma_C(y - x) : y \in A \}.$$ 

Proof. Fix $x \in X$. For $x_0 \in A$ we have $f(y) \geq f(x_0) - \sigma_C(x_0 - y)$ for every $y \in A$. Then

$$f(y) + \sigma_C(x - y) \geq f(x_0) - \sigma_C(x_0 - y) + \sigma_C(x - y) \geq f(x_0) - \sigma_C(x_0 - x),$$

shows $f_+(x) > -\infty$ for each $x \in X$. 
Given $x, x' \in X$ and $\epsilon > 0$, choose $y' \in A$ such that $f_+(x') + \epsilon > f(y') + \sigma_C(x' - y')$. By definition, we also have

$$f_+(x) \leq f(y') + \sigma_C(x - y') \leq f(y') + \sigma_C(x - x') + \sigma_C(x' - y'),$$

so we obtain $f_+(x) \leq f_+(x') + \epsilon + \sigma_C(x - x')$. This holds for arbitrary $x$ and $x'$ and $\epsilon$, so $f_+(x) - f_+(x') \leq \sigma_C(x - x')$, that is, $f_+$ is $C$-Lipschitz on $X$. Let $x \in A$. By definition, $f_+(x) \leq f(x)$, and also the $C$-Lipschitz property of $f$ on $A$ implies

$$f(x) - f(y) \leq \sigma_C(x - y)$$

for each $y \in A$, thus $f(x) \leq f_+(x)$. In a word, $f_+ = f$ on $A$.

Let $\tilde{f}$ satisfy (5.22). Then $\tilde{f}(x) - f(y) = \tilde{f}(x) - \tilde{f}(y) \leq \sigma_C(x - y)$ for $y \in A$, thus $\tilde{f}(x) \leq f(y) + \sigma_C(x - y)$ for each $y \in A$ and so $\tilde{f} \leq f_+$. That is, $\tilde{f}$ minorizes $f_+$ on $X$. The proof for $f_2$ is similar. \qed

For a Banach space $X$ and $M > 0$, we will let $L^C_{X,M}$ be the space of $C$-Lipschitz functions on $X$ such that $\sup_{x \in X} |f(x)| \leq M$ with metric $\rho(f, g) := \sup\{|f(x) - g(x)| : x \in X\}$. For a closed subset $A \subseteq X$ and $f \in L^C_{A,M}$ we let

$$L^C_{f,M} := \{g \in L^C_{X,M} : g|_A = f\},$$

which is a closed subset of $L^C_{X,M}$. We will write

$$f_+ := \sup\{g : g \in L^C_{f,M}\} \quad \text{and} \quad f_- := \inf\{g : g \in L^C_{f,M}\}.$$

**Proposition 5.11.2** Suppose $A$ is a closed proper subset of $X$ and $0 \in \text{int} C$. Let $f \in L^C_{A,M}$ and $\bar{x} \in X \setminus A$. Then $f_+(\bar{x}) > f_-(\bar{x})$ if and only if $f_+(\bar{x}) > f_-(\bar{x})$.

**Proof.** Assume $f_+(\bar{x}) > f_-(\bar{x})$. As $\delta B_X(0) \subset \text{int} C$ we have

$$f_+(\bar{x}) \geq -M + \inf_{y \in A} \sigma_C(\bar{x} - y) > -M + \inf_{y \in A} \delta\|\bar{x} - y\| > -M,$$

$$f_-(\bar{x}) \leq M - \inf_{y \in A} \sigma_C(\bar{x} - y) \leq M - \inf_{y \in A} \delta\|\bar{x} - y\| < M,$$

as $\bar{x} \not\in A$. Then $f_+(\bar{x}) > (f_- \vee -M)(\bar{x})$ and $M > (f_- \vee -M)(\bar{x})$ shows $(f_+ \wedge M)(\bar{x}) > (f_- \vee -M)(\bar{x})$, and both $f_+ \wedge M$ and $f_- \vee -M$ belongs to $L^C_{f,M}$. The other direction is obvious. \qed
Theorem 5.11.3 (C-maximal subdifferential extension) Suppose $X$ is a separable Banach space, $A$ a closed set of $X$ and $C \subset X^*$ a w*-compact convex set with non-empty norm interior. Suppose $f \in L^C_{f,M}$ is such that $f-(x_n) < f+(x_n)$ for each $n$ with \{x_n\} being dense in $X \setminus A$. Then the set

$$\{ \tilde{f} \in L^C_{f,M} : \partial_{e^*} \tilde{f} = \partial_c \tilde{f} = C \text{ for each } x \in X \setminus A \}$$

is residual in $L^C_{f,M}$.

Proof. Let $x \in X \setminus A$ be a point at which $X$ is $C$-Lipschitz separated, and $c \in \text{int}(C)$. As $A$ is closed we have $d_A(x) = r > 0$. Consider $G^n_{x,c} :=$

$$\{ f \in L^C_{f,M} : \text{there exists } \bar{x} \in X \setminus A \text{ satisfying } \| \bar{x} - x \| < \frac{1}{n} < r \text{ such that for some } \bar{x}^* \in \partial_- f(\bar{x}) \text{ we have } \| \bar{x}^* - c \|_* < \frac{1}{n} \}.$$

(a) $\text{int}(G^n_{x,c})$ is dense in $L^C_{f,M}$. Given $g \in L^C_{f,M}$ and $\epsilon > 0$. As $X$ is $C$-Lipschitz separated at $x$ there exists $h \in L^C_{f,M}$ such that $|g(x) - h(x)| > \eta > 0$. Without loss of generality, we may assume $2\epsilon < \eta$ and $g(x) - h(x) > \eta$. For $\lambda > 0$ define $g_\lambda := (1-\lambda)g + \lambda h \in L^C_{f,M}$. Since $g$ and $h$ are bounded, $g_\lambda \to g$ uniformly as $\lambda \downarrow 0$. Choose $\lambda > 0$ sufficiently small such that $\rho(g, g_\lambda) < \epsilon$. For such a $\lambda$ we have $g_\lambda(x) = (1-\lambda)g(x) + \lambda h(x) < (1-\lambda)g(x) + \lambda(g(x) - 2\epsilon) < g(x)$. Now choose $g_\lambda(x) < \alpha = g(x)$ and define

$$\tilde{f}(\bar{x}) := \min \{ \max \{ g_\lambda(\bar{x}), \alpha + \sigma_C(\bar{x} - x) \}, g(\bar{x}) \}.$$

As $g_\lambda|_A = f$ and $g|_A = f$ we have $\tilde{f}|_A = f$. As $\max \{ g_\lambda(\bar{x}), \alpha + \sigma_C(\bar{x} - x) \} \geq -M$, and $g(\bar{x}) \leq M$, we have $-M \leq f \leq M$. Furthermore, as $C$-Lipschitz functions are closed under lattice operations, we have $\tilde{f} \in L^C_{f,M}$. At $x$, $g_\lambda(x) < \alpha$ and $g(x) > \alpha$, thus for sufficiently small $0 < \delta < \min\{d_A(x), 1/n\}$ when $\| \bar{x} - x \| \leq \delta$ we have $\tilde{f}(\bar{x}) = \alpha + \sigma_C(\bar{x} - x)$. Moreover, $g_\lambda \geq g - \epsilon$ shows $\rho(\tilde{f}, g) \leq \epsilon$. As $\partial_- \tilde{f}(\bar{x}) = C$, we have $\tilde{f} \in G^n_{x,c}$. We proceed to show $\tilde{f} \in \text{int}(G^n_{x,c})$. Now as $C \supset (c + r B_{X^*})$ for some $r > 0$, we have

$$m := \inf \{ \tilde{f}(\bar{x}) - \langle c, \bar{x} \rangle : \| \bar{x} - x \| = \delta \} = \tilde{f}(x) - \langle c, x \rangle + \inf \{ \sigma_C-c(\bar{x} - x) : \| \bar{x} - x \| = \delta \} \geq \tilde{f}(x) - \langle c, x \rangle + \delta r > \tilde{f}(x) - \langle c, x \rangle.$$

Let $\nu := m - (\tilde{f}(x) - \langle c, x \rangle)$. For $g \in L^C_{f,M}$ with $\rho(g, \tilde{f}) < \beta < \alpha/2$, we have

$$\inf \{ g(\bar{x}) - \langle c, \bar{x} \rangle : \| \bar{x} - x \| = \delta \} \geq -\beta + m > \beta + \tilde{f}(x) - \langle c, x \rangle \geq g(x) - \langle c, x \rangle.$$
5.11 The subdifferentiability of C-Lipschitz extensions

Define \( g_1 := g - \langle c, \cdot \rangle + I_{B_k[x]} \) which is lower semicontinuous and bounded below on \( X \). Choose \( \nu > 0 \) and \( x_0 \) such that \( 0 < \nu < \min \{ \frac{1}{2n}, \inf \{ g_1(\bar{x}) : \|\bar{x} - x\| = \delta \} - g_1(x) \} \).

\[
  g_1(x_0) < \inf_X g_1 + \nu < \inf \{ g_1(\bar{x}) : \|\bar{x} - x\| = \delta \}.
\]

As \( X \) is separable, there exists a Gâteaux differentiable \( \phi \) on \( X \) and \( v \in X \) such that

\[
  g_1(v) < \inf \{ g_1 + \nu < \inf \{ g_1(\bar{x}) : \|\bar{x} - x\| = \delta \},
\]

\[
  g_1(\bar{x}) + \phi(\bar{x}) \geq g_1(v) + \phi(v) \quad \text{for all } \bar{x} \in X, \text{ and } \|\nabla \phi(v)\| < \frac{1}{n}. \tag{5.24}
\]

Equation (5.23) shows \( \|v - x\| < \delta < 1/n \), then by equation (5.24) we have \( 0 \in \partial_g(v) - c + \nabla \phi(v) \). That is, for some \( v^* \in \partial_{-g}(v) \) we have \( \|v^* - c\| < 1/n \), thus \( g \in G^n_{x,c} \). Hence \( \tilde{f} \in \text{int}(G^n_{x,c}) \).

(b) Since \( \text{int}(G^n_{x,c}) \) is open and dense in \( L^C_{f,M} \), the set \( G_{x,c} := \bigcap_{n=1}^{\infty} \text{int}(G^n_{x,c}) \) is dense in \( L^C_{f,M} \). If \( f \in G_{x,c} \), then for every \( n \) there exists \( \|\bar{x}_n^* - c\| < 1/n \), \( \|\bar{x}_n - x\| < 1/n \) with \( \bar{x}_n^* \in \partial_{-f}(\bar{x}_n) \). Letting \( n \to \infty \) we obtain \( c \in \partial_a f(x) \).

(c) Take a countable \( w^* \)-dense set \( \{c_k\} \) from the norm interior of \( C \). Then \( G_x := \bigcap_{k=1}^{\infty} G_{x,c_k} \) is residual in \( L^C_{f,M} \). As \( c_k \in \partial_a f(x) \) for every \( k \) and \( \partial_a f(x) \) is \( w^* \)-closed, we have \( C \subset \partial_a f(x) \subset C \), then \( \partial_a f(x) = C \).

(d) Now let \( \{x_k\} \) be a countable norm-dense set in \( X \setminus A \). Then \( G := \bigcap_{k=1}^{\infty} G_{x_k} \) is residual in \( L^C_{f,M} \). If \( f \in G \), then \( \partial_a f(x_k) = C \). By the \( w^* \)-upper semicontinuity of \( \partial_a f \) we get \( \partial_a f \equiv C \) on \( X \setminus A \).

The following more general result may be proved by using Borwein, Giles and Vanderwerff’s argument [6] and Corollary 5.6.3, we omit its proof.

**Theorem 5.11.4** Suppose \( X \) is a separable Banach space, \( A \) a closed subset of \( X \) and \( C \subset X^* \) a \( w^* \)-compact convex set. Suppose \( f \in L^C_{A,M} \) is such that \( f_-(x_n) < f_+(x_n) \) for each \( n \) with \( \{x_n\} \) being dense in \( X \setminus A \). Then in \( L^C_{f,M} \) the set

\[
  \{ \tilde{f} \in L^C_{f,M} : \partial_c \tilde{f} = \partial_a \tilde{f} = C \text{ for each } x \in X \setminus A \}
\]

is residual.

**Example 5.11.5** Let \( A \) be a closed subset of \( \mathbb{R}^n \) and \( C \) a compact convex subset with \( 0 \in \text{int} \, C \), and let \( f(x) \equiv 0 \) for \( x \in A \). Then there exists \( \tilde{f} \) on \( \mathbb{R}^n \) such that \( \partial_c \tilde{f}(x) = 0 \) if
The subdifferentiability of C-Lipschitz extensions

$x \in \text{int} A$ and $\partial_c f(x) = C$ otherwise. Indeed, for $k$ sufficiently large, since $1/kB_X \subset C$ both $[-1/kd_A(.)] \land M$ and $1/kd_A(.) \land M$ belongs to $L^C_{0,M}$, then $\mathbb{R}^n \setminus A$ is C-Lipschitz separated. Theorem 5.11.3 applies.

C-Lipschitz separated spaces

Definition 5.11.6 A Banach space $X$ is said to be C-Lipschitz separated if for every closed convex subset $A$ of $X$ and $f \in L^C_{A,M}$ we have $f_{+,M} > f_{-,M}$ for all $x \in X \setminus A$.

Proposition 5.11.7 Let $0 \in C$. A Banach space $X$ is C-Lipschitz separated if and only if for every closed proper subspace $Y \subset X$ and every $f \in L^C_{Y,M}$ one has $\sup_{g \in L^C_{Y,M}} g(x) > \inf_{g \in L^C_{Y,M}} g(x)$ for every $x \in X \setminus Y$.

Proof. Let $A$ be a closed convex subset of $X$ and $\bar{x} \notin A$. By translation, we may assume $\bar{x} \neq 0$, and using the separation theorem we may choose $\phi \in X^*$ such that $\phi(\bar{x}) < 0$ and $\inf_A \phi \geq 0$. Let $Y = \ker \phi$, and let $f \in L^C_{A,M}$. Then by Lemma 5.11.1 there exists a C-Lipschitz function $f_1$ such that $f_1|_A = f$. Now let $g \in L^C_{Y,M}$ be the restriction of $(f_1 \lor -M) \land M$ to $Y$. By the hypothesis, we have $\sup_{g \in L^C_{Y,M}} g(\bar{x}) > \inf_{g \in L^C_{Y,M}} g(\bar{x})$. But now let $\tilde{f}_1(x) := (f_1(x) \lor -M) \land M$ if $\phi(x) \geq 0$ and $\tilde{f}_1(x) = g(x)$ if $\phi(x) \leq 0$. We have $\tilde{f}_1 \in L^C_{Y,M}$ because $\tilde{f}_1$ is constructed using functions that are C-Lipschitz on $\{x : \phi(x) \leq 0\}$ and $\{x : \phi(x) \geq 0\}$ and they agree on $\{x : \phi(x) = 0\}$. Since $\bar{x}$ was arbitrary and $\phi(\bar{x}) < 0$, this completes the proof of “if” part. The other direction is obvious.

Definition 5.11.8 A Banach space $X$ has the the C-maximal subdifferential extension property (C-MSEP) if for every closed proper subspace $Y$ of $X$ and each bounded C-Lipschitz function on $Y$, there exists a C-Lipschitz function $\tilde{f}$ on $X$ such that $\tilde{f}|_Y = f$ and $\partial_c \tilde{f}(x) = C$ for each $x \in X$.

It follows from Theorem 5.11.3 that if a Banach space $X$ is C-Lipschitz separated then $X$ has C-maximal extension property.

Example 5.11.9 [6]Let $X = \mathbb{R}^2$ and $f(s,0) = (s \land 2) \lor -2$ be defined on $Y = \mathbb{R} \times \{0\}$. When $C = \{x : \|x\|_1 \leq 1\}$, any C-Lipschitz function $\tilde{f}$ such that $\tilde{f}|_Y = f$ satisfies $\tilde{f}(x,y) = x$ when $\| (x,y) \|_\infty < 1$, thus in this case $f$ does not have C-MSEP.
In order to characterize $C$-Lipschitz separated Banach spaces, we need a few lemmas.

**Definition 5.11.10** Let $S \subset X$ be a closed convex set with $0 \in S$. The function $\nu_S$ defined by $\nu_S(x) := \inf \{ \lambda > 0 : x \in \lambda S \}$ is called the gauge of $S$. The set

$$S^0 := \{ x^* \in X^* : x^*(s) \leq 1 \text{ for all } s \in S \},$$

is called the polar of $S$.

Now $\nu_S$ is finite everywhere if and only if $0 \in \text{int}(S)$. When $S$ is bounded, we have $\nu_S(x) = 0$ if and only if $x = 0$, and $S$ is strictly convex if and only if $\nu_S(x) = 1, \nu_S(y) = 1, \nu_S(x+y) = 2$ imply $x = y$.

**Lemma 5.11.11** If $\nu_S(x + y) = \nu_S(x) + \nu_S(y)$, $\nu_S(x) \neq 0$, $\nu_S(y) \neq 0$, then $\nu_S(x)/\nu_S(x) + x/\nu_S(y) = 2$.

**Proof.** Assume $\frac{1}{\nu_S(x)} \geq \frac{1}{\nu_S(y)}$, then

$$2 \geq \nu_S \left( \frac{x}{\nu_S(x)} + \frac{y}{\nu_S(y)} \right) = \nu_S \left( \frac{x+y}{\nu_S(x)} \right) - \left( \frac{1}{\nu_S(x)} \right) \nu_S(y),$$

$$\geq 2 \nu_S \left( \frac{x+y}{\nu_S(x)} \right) - \nu_S \left( \frac{1}{\nu_S(x)} \right) \nu_S(y) - \left( \frac{1}{\nu_S(x)} \right) \nu_S(y),$$

$$= 1 + \frac{\nu_S(y)}{\nu_S(x)} - \frac{\nu_S(y)}{\nu_S(x)} + 1 = 2. \quad \Box$$

**Lemma 5.11.12** Assume $X$ is finite dimensional and $C \subset X^*$ is compact convex with $0 \in \text{int}C$. If $C^0$ is strictly convex, then $X$ is $C$-Lipschitz separated.

**Proof.** Assume there exists a non-empty closed convex subset $A$ such that $a = f_+(x) = f_-(x)$ for some $x \not\in A$. Without loss of generality, we may assume $x = 0$, that is, $0 \not\in A$. For every $n$ there exist $x_n, y_n \in A$ such that

$$a \leq f(y_n) + \sigma_C(-y_n) < a + \frac{1}{n}, \quad a \geq f(x_n) - \sigma_C(x_n) > a - \frac{1}{n}. \quad (5.25)$$

As $0 \in \text{int}(C)$, for some $r > 0$ we have $rB_{X^*} \subset C$, then $\sigma_C(-y_n) \geq r\|y_n\|$. By equation (5.25) we get $f(y_n) + r\|y_n\| < a + \frac{1}{n}$. As $f$ is bounded, $y_n$ is bounded in norm. The same reason shows $x_n$ is norm bounded. As $X$ is finite-dimensional, we may take
subsequences of \( \{x_n\} \) and \( \{y_n\} \) such that they converges to \( x, y \in A \) respectively. Then
\[
f(y) + \sigma_C(-y) = a = f(x) - \sigma_C(x).
\]
Hence,
\[
\sigma_C(x - y) \leq \sigma_C(x) + \sigma_C(-y) = f(x) - f(y) \leq \sigma_C(x - y).
\]
As \( \sigma_C = \nu_C^{\circ} \), we have \( \nu_C^{\circ}(x - y) = \nu_C(x) + \nu_C(-y) \). Since \( C^0 \) is compact, \( x, -y \neq 0 \), we have \( \nu_C(x) > 0 \), \( \nu_C(-y) > 0 \), then \( \nu_C^{\circ}(\frac{x}{\nu_C(x)} + \frac{-y}{\nu_C(-y)}) = 2 \) by Lemma 5.11.11. Since \( C^0 \) is strictly convex, \( \frac{x}{\nu_C(x)} + \frac{-y}{\nu_C(-y)} \), and this implies \( 0 \in A \), a contradiction. \( \square \)

**Example 5.11.13** As in Example 5.11.9, let \( X = \mathbb{R}^2 \) and \( f(s, 0) = (s \wedge 2) \vee -2 \) be defined on \( Y = \mathbb{R} \times \{0\} \). When \( C = \{ x : \|x\|_2 \leq 1 \} \), \( X \) is \( C \)-Lipschitz separated by Lemma 5.11.12, thus there exists a Lipschitz function \( \tilde{f} \) on \( \mathbb{R}^2 \) such that \( \tilde{f}|_Y = f \) and \( \partial_x \tilde{f} \equiv C \).

**Lemma 5.11.14** Assume \( 0 \in \text{int} C \) and \( C \) is \( w^* \)-compact convex in \( X^* \), the dual of a Banach space \( X \). If \( X \) is \( C \)-Lipschitz separated, then \( C^0 \) is strictly convex.

**Proof.** Suppose \( C^0 \) is not strictly convex, then there exist \( x \neq y \) with \( \nu_C^{\circ}(x) = 1 \), \( \nu_C^{\circ}(y) = 1 \) and \( \nu_C^{\circ}(x + y) = 2 \). This is the same as \( \nu_C^{\circ}(\lambda x + (1 - \lambda)y) = 1 \) for every \( 0 \leq \lambda \leq 1 \). Write \( y_0 := (x + y)/2, h_0 = (y - x)/2 \), then \( \nu_C^{\circ}(y_0 + th_0) = 1 \) for \( |t| \leq 1 \). As \( \nu_C^{\circ}(y_0) = \sigma_C(y_0) = 1 \) and \( C \) is compact, we may choose \( c \in C \) such that \( c(y_0) = 1 \). Let \( Z = \text{lin}\{y_0\} \) then either \( y_0 + h_0 \notin Z \) or \( y_0 - h_0 \notin Z \). Suppose \( y_0 + h_0 \notin Z \). As \( Z \) is closed there exists \( \Lambda \in X^* \) such that \( \Lambda(sy_0) < \Lambda(y_0 + h_0) \) for every \( s \in \mathbb{R} \), we obtain \( \Lambda(y_0) = 0 \) and \( \Lambda(h_0) > 0 \). Let \( B \) be the kernel of \( \Lambda \). Then \( X = Y + \text{lin}\{h_0\} \). Define \( f \) on \( Y \) by \( f(y) = (c(y) \vee -2) \wedge 2 \). Then \( f \) is \( C \)-Lipschitz on \( Y \). Let \( \tilde{f} \) be any \( C \)-Lipschitz extension of \( f \). For \( y \in Y \) with \( |c(y)| \leq 1 \) and \( |t| \leq 1 \) we show that \( \tilde{f}(y + th_0) = c(y) \). Indeed, when \( |c(y)| \leq 1 \) with \( y \in Y \) we have \( \Lambda(y \pm y_0) = 0 \), that is, \( y \pm y_0 \in Y \), with \( |c(y \pm y_0)| \leq 2 \) we have \( f(y \pm y_0) = c(y) \pm c(y_0) \).

When \( y \in Y, |t| \leq 1 \) we have \( \tilde{f}(y + th_0) - \tilde{f}(y - th_0) \leq \sigma_C(y_0 + th_0) = 1 \), thus
\[
\tilde{f}(y + th_0) \leq c(y) - c(y_0) + 1 = c(y).
\]

On the other hand, \( \tilde{f}(y + y_0) - \tilde{f}(y + th_0) \leq \sigma_C(y_0 - th_0) = 1 \) shows
\[
\tilde{f}(y + th_0) \geq \tilde{f}(y + y_0) - 1 = c(y) + c(y_0) - 1 = c(y).
\]

Let \( P \) be the projection map from \( X \) onto \( Y \). Then \( \tilde{f}(x) = \tilde{f}(y + th_0) = c \circ P(x) \) for \( x \in \{ z : |c \circ P(z)| < 1, |(I - P)(z)|/\|h_0\| < 1 \} \), where the latter is an open set in \( X \) as both
5.12 Sum rules and Jacobians

\( c \circ P \) and \( I - P \) are continuous. Then \( X \) is not \( C \)-Lipschitz separated. The proof for the case \( y_0 - h_0 \notin Z \) is similar. \( \square \)

Lemma 5.11.12, 5.11.14, and Theorem 5.11.3 together show that:

**Theorem 5.11.15** Let \( X \) be a finite-dimensional Banach space and \( C \subset X^* \) be compact convex with \( 0 \in \text{int} C \). Then the following are equivalent:

(i) \( C^0 \) is strictly convex;

(ii) \( X \) is \( C \)-Lipschitz separated;

(iii) \( X \) is densely \( C \)-Lipschitz separated;

(iv) \( X \) has \( C \)-MSEP.

In particular, if \( C = B_{X^*} \), then \( X \) has the \( C \)-maximal extension property if and only if \( B_X \) is strictly convex.

**Corollary 5.11.16 (Pastng Clarke subdifferentials)** Let \( f \) be a bounded \( C \)-Lipschitz function on a finite-dimensional Banach space \( X \) with \( 0 \in \text{int} C \). If \( C^0 \) is strictly convex, then for every closed convex set \( A \subset X \), the set-valued map \( T : X \to X^* \) defined by \( T(x) := \partial_c f(x) \) for \( x \in \text{int} A \) and \( T(x) := C \) otherwise, is a Clarke subdifferential.

5.12 Sum rules and Jacobians

The next “cancellation” lemma is well known.

**Lemma 5.12.1** If \( A \) and \( B \) are convex, \( B \) is closed, \( C \) is bounded in topological vector space \( X \), then \( A + C \subset B + C \) implies \( A \subset B \).

**Proof.** By convexity, \( nA + C \subset nB + C \) for \( n \in \mathbb{N} \). Now fix \( x \in A \) and \( c \in C \), for every \( n \in \mathbb{N} \) we have \( nx + c = nb_n + c_n \) for some \( c_n \in C, b_n \in B \). That is, \( x - b_n = (c_n - c)/n \). Because \( C \) is bounded, \( x - b_n \to 0 \) as \( n \to \infty \). Since \( B \) is closed, \( b_n \in B \), we have \( x \in B \). \( \square \)

When \( \{f_i\}_{i=1}^k \) are locally Lipschitz functions on an open subset \( A \) of a Banach space \( X \), we usually have \( \partial_c (\sum_{i=1}^k f_i)(x) \subseteq \sum_{i=1}^k \partial_c f_i(x) \) for each \( x \in A \). We will show that the sum
5.12 Sum rules and Jacobians

rule generically holds with equality because of maximality in appropriate complete metric
spaces. Let \( f : A \subset X \to \mathbb{R} \) be locally Lipschitz and \( T := \partial_c f \). For \( X_T \times X_T \), the product
metric \( \rho_1 \) is defined by:

\[
\rho_1((f_1, f_2), (\tilde{f}_1, \tilde{f}_2)) := \rho(f_1, \tilde{f}_1) + \rho(f_2, \tilde{f}_2) \]

for \( f_1, \tilde{f}_1 \in X_T \).

**Theorem 5.12.2** Suppose \( X \) is a separable Banach space \( X \). In \( (X_T \times X_T, \rho_1) \) there exists a
residual set \( G \) such that if \( (f_1, f_2) \in G \), then \( \partial_c(f_1+f_2) = \partial_c f_1 + \partial_c f_2 \) with \( \partial_c f_1 = \partial_c f_2 = \partial_c f \), and

\[
\partial_c[\min\{f_1, f_2\}] = \partial_c[\max\{f_1, f_2\}] = \partial_c f = \co(\partial_c f, \partial_c f).
\]

**Proof.** Fix \( x \in A, v \in X \), we consider \( G_k := \)

\[
\{(f_1, f_2) \in X_T \times X_T : \frac{(f_1 + f_2)(y + tv) - (f_1 + f_2)(y)}{t} - 2f^0(x; v) > \frac{1}{k}
\text{ for some } t, y \text{ satisfying } 0 < t < \frac{1}{k} \text{ and } \|y - x\| < \frac{1}{k}.\}
\]

(a) \( G_k \) is open in \( X_T \times X_T \). Assume \( (f_1, f_2) \in G_k \). Then for some \( t \) and \( y \) satisfying
\( 0 < t < 1/k \) and \( \|y - x\| < 1/k \) we have

\[
\frac{(f_1 + f_2)(y + tv) - (f_1 + f_2)(y)}{t} - 2f^0(x; v) > \frac{1}{k}. \quad (5.26)
\]

For this \( y \) and \( t \), when \( \rho_1((\tilde{f}_1, \tilde{f}_2), (f_1, f_2)) < \epsilon \) we have:

\[
\frac{(\tilde{f}_1 + \tilde{f}_2)(y + tv) - (\tilde{f}_1 + \tilde{f}_2)(y)}{t} - \frac{(f_1 + f_2)(y + tv) - (f_1 + f_2)(y)}{t} \geq \frac{-2\epsilon}{t} + \left\{ \frac{(f_1 + f_2)(y + tv) - (f_1 + f_2)(y)}{t} \right\},
\]

The bracketed part is greater than \( 2f^0(x; v) - 1/k \) by equation (5.26), we may choose \( \epsilon \)
small such that

\[
\frac{-2\epsilon}{t} + \frac{(f_1 + f_2)(y + tv) - (f_1 + f_2)(y)}{t} > 2f^0(x; v) - \frac{1}{k}.
\]

Hence \( G_k \) is open.

(b) \( G_k \) is dense in \( X_T \times X_T \). Let \( \epsilon > 0 \) and \( (f_1, f_2) \in X_T \times X_T \). Define

\[
h_i := \max\{f_i - 2\epsilon, \min\{f_i, f_i(x) - \epsilon + f - f(x)\}\} \quad \text{for } i = 1, 2.
\]
Then \( f_i - 2\epsilon \leq h_i \leq f_i, \rho_i((h_1, h_2), (f_1, f_2)) \leq 4\epsilon \) and \((h_1, h_2) \in \mathcal{X}_T \times \mathcal{X}_T\). Choose \( \delta > 0 \) sufficiently small such that when \( \|x - \bar{x}\| < \delta < 1/k \) we have \( h_i(\bar{x}) = f_i(x) - \epsilon + f(\bar{x}) - f(x) \). By the definition of \( f^0(x; v) \), we may choose \( t \) and \( y \) satisfying \( 0 < t < 1/k \) and \( \|y - x\| < 1/k \) such that \( y + tw, y \in B_\delta(x) \) and

\[
\frac{f(y + tv) - f(y)}{t} > f^0(x; v) - \frac{1}{2k}.
\]

For these \( y \) and \( t \) we have

\[
\frac{h_i(y + tv) - h_i(y)}{t} = \frac{(f_i(x) - \epsilon + f(y + tv) - f(x)) - (f_i(x) - \epsilon + f(y) - f(x))}{t} = \frac{f(y + tv) - f(y)}{t}.
\]

Then

\[
\frac{(h_1 + h_2)(y + tv) - (h_1 + h_2)(y)}{t} = 2\frac{f(y + tv) - f(y)}{t} > 2f^0(x; v) - \frac{1}{k}.
\]

This shows \((h_1, h_2) \in G_k\). Hence \( G_k \) is dense.

(c) Since \( G_k \) is open and dense, the set \( G_{x, v} := \bigcap_{k=1}^{\infty} G_k \) is residual in \( \mathcal{X}_T \times \mathcal{X}_T \). If \((f_1, f_2) \in G_{x, v}\) then for every \( k \) there exist \( t_k \) and \( y_k \) satisfying \( 0 < t_k < 1/k \) and \( \|y_k - x\| < 1/k \) such that

\[
\frac{(f_1 + f_2)(y_k + tkv) - (f_1 + f_2)(y_k)}{t_k} > 2f^0(x; v) - \frac{1}{k}, \quad \text{and so}
\]

\[
(f_1 + f_2)^0(x; v) \geq \limsup_{y_k \rightarrow 0, t_k \rightarrow 0} \frac{(f_1 + f_2)(y_k + tkv) - (f_1 + f_2)(y_k)}{t_k} \geq 2f^0(x; v).
\]

(d) As \( A \times X \) is separable, by Lemma 5.2.1 we may choose a countable set \( D := \{(x_k, v_k)\} \) being dense in \( A \times X \) such that for every \((x, v) \in A \times X\) we have

\[
f^0(x; v) = \limsup_{(x_k, v_k) \rightarrow (x, v)} f^0(x_k; v_k).
\]

Let \( G := \bigcap_{k=1}^{\infty} G_{x_k, v_k}\). If \( f \in G \) then \((f_1 + f_2)^0(x_k; v_k) \geq 2f^0(x_k; v_k) \) for every \((x_k, v_k) \in D\). Then for every \((x, v) \in A \times X \) we have

\[
(f_1 + f_2)^0(x; v) \geq \limsup_{(x_k, v_k) \in D} (f_1 + f_2)^0(x_k; v_k) \geq 2 \limsup_{(x_k, v_k) \in D} f^0(x_k; v_k) = 2f^0(x; v).
\]

(e) By (d), if \((f_1, f_2) \in G \) then \( \partial_c (f_1 + f_2) \supset \partial_c f \) on \( A \). But \( \partial_c (f_1 + f_2) \subset \partial_c f + \partial_c f = 2\partial_c f \). Hence \( \partial_c (f_1 + f_2) = 2\partial_c f \). Furthermore,

\[
\partial_c f + \partial_c f = \partial_c (f_1 + f_2) \subset \partial_c f_1 + \partial_c f_2 \subset \partial_c f_i + \partial_c f_i \text{ for } i = 1, 2.
\]
shows $\partial_c f \subset \partial_c f_1$ by Lemma 5.12.1. Hence $\partial_c f_1 = \partial_c f_2 = \partial_c f$, so $\partial_c (f_1 + f_2) = \partial_c f_1 + \partial_c f_2$. Write $f_1 + f_2 = \min\{f_1, f_2\} + \max\{f_1, f_2\}$. Since $\partial_c (f_1 + f_2) = 2\partial_c f$ we have

$$\partial_c f + \partial_c f \subset \partial_c (\min\{f_1, f_2\}) + \partial_c (\max\{f_1, f_2\}) \subset \partial_c (\min\{f_1, f_2\}) + \partial_c f,$$

By Lemma 5.12.1 again, we have $\partial_c f \subset \partial_c (\min\{f_1, f_2\})$. Similarly, $\partial_c f \subset \partial_c (\max\{f_1, f_2\})$. Hence $\partial_c (\min\{f_1, f_2\}) = \partial_c f = \partial_c (\max\{f_1, f_2\})$. $\square$

**Example 5.12.3** For a vector-valued locally Lipschitz function $F : \mathbb{R}^n \to \mathbb{R}^m$ given by $F(x) := [f_1(x), f_2(x), \ldots, f_m(x)]$, the Clarke Jacobian of $F$ at $x$, denoted by $\partial F(x)$, is defined by

$$\partial F(x) := \text{co}\{\lim F(x_i) : x_i \to x, x_i \not\in \Omega_F\},$$

where $\Omega_F$ denotes the set of points at which $F$ fails to be differentiable. Letting $f_1 = \ldots = f_m := f$ with $\partial_c f \equiv B_{\mathbb{R}^n}$. Then

$$\partial F(x) = \text{co}\{(x^*, \ldots, x^*) : x^* = \lim \nabla f(x_i)\} = \{(x^*, \ldots, x^*) : x^* \in B_{\mathbb{R}^n}\},$$

whereas,

$$\partial_c f_1(x) \times \ldots \times \partial_c f_m(x) \equiv B_{\mathbb{R}^n} \times \ldots \times B_{\mathbb{R}^n},$$

thus $\partial F(x) \neq \partial_c f_1(x) \times \ldots \times \partial_c f_m(x)$ for every point $x \in A$.

**Corollary 5.12.4** Let $C$ be a nonsingleton compact convex subset of $\mathbb{R}^n$. In $\mathcal{X}_C \times \mathcal{X}_C$ there exists a residual set $G$ such that for every $F \subset G$, the Clarke Jacobian $JF$ is not a minimal cusco.

**Proof.** By Corollary 5.2.9, there exist a Lipschitz function $f$ such that $\partial_c f \equiv C$. Applying Theorem 5.12.2 we obtain a residual set $G \in \mathcal{X}_C \times \mathcal{X}_C$ such that for $F := (f_1, f_2) \in G$ we have $\partial_c (f_1 + f_2) = 2C$. By Theorem 2.6.6 [33] we have $\partial_c (f_1 + f_2) = (1, 1) JF$. Since $2C$ is not a minimal cusco, $JF$ is not minimal. $\square$

**Remark 5.12.5** For simplicity, we chose to prove that generically in $\mathcal{X}_F \times \mathcal{X}_F$, $\partial_c (f_1 + f_2) = 2\partial_c f$. A similar proof applies to the case of positive linear combinations.

Compare Corollary 5.12.4 with Proposition 8.3 [20]: If $f_i$ is strictly differentiable almost everywhere on an open set $A \subset \mathbb{R}^n$ for $i = 1, \ldots, k$, then for the vector-valued function $F : A \to \mathbb{R}^k$ defined by $F := (f_1, \ldots, f_k)$, its Clarke Jacobian $JF$ is a minimal cusco on $A$. 

Now we modify the proof of Theorem 5.4.2 to show that the approximate sum rule holds with equality generically:

**Theorem 5.12.6** Suppose $X$ is a separable Banach space and $C \subset X^*$ is a w*-compact convex set with non-empty norm interior. In $(X_C \times X_C, \rho_1)$ the set

$$\{(f_1, f_2) \in X_C \times X_C : \partial_a f_1 = \partial_a f_2 \equiv C, \partial_a (f_1 + f_2) \equiv 2C \text{ on } A\}$$

is residual.

**Proof.** Fix $x \in A$, $c_1, c_2 \in \text{int}(C)$. Define $G^n_{x,c_1,c_2} :=$

$$\{(f_1, f_2) \in X_C \times X_C : \text{there exist } \bar{x}, \bar{y}, \bar{z} \in A, \|\bar{x} - x\| < \frac{1}{n}, \|\bar{y} - x\| < \frac{1}{n}, \|\bar{z} - x\| < \frac{1}{n} \text{ so that for some } \bar{x}^* \in \partial f_1(\bar{x}), \bar{y}^* \in \partial f_2(\bar{y}), \bar{z}^* \in \partial (f_1 + f_2)(\bar{z}) \text{ we have } \|\bar{x}^* - c_1\| < \frac{1}{n}, \|\bar{y}^* - c_2\| < \frac{1}{n}, \|\bar{z}^* - (c_1 + c_2)\| < \frac{1}{n}\}$$

(a) $\text{int}(G^n_{x,c_1,c_2})$ is dense in $X_C \times X_C$. Given $(f_1, f_2) \in X_C \times X_C$. For every $\epsilon > 0$, define $h_i : A \to \mathbb{R}$ by

$$h_i(x) := \max\{\min\{f_i(x) - \epsilon + \sigma_C(\bar{x} - x), f_i(x) - 2\epsilon\}, f_i(x)\}.$$

Then $\rho(f_i, h_i) \leq 2\epsilon$ and so $\rho_1((f_1, f_2), (h_1, h_2)) \leq 4\epsilon$, that is, $(h_1, h_2)$ is near-by $(f_1, f_2)$. Furthermore, for some $0 < \delta < 1/n$, we have $h_i(\bar{x}) = f_i(x) - \epsilon + \sigma_C(\bar{x} - x)$ whenever $\|\bar{x} - x\| \leq \delta$, thus $(h_1, h_2) \in G^n_{x,c_1,c_2}$. We will show that $(h_1, h_2) \in \text{int}(G^n_{x,c_1,c_2})$. As $c_i \in \text{int}(C)$, there exists $r_i > 0$ such that $c_i + r_i B_{X^*} \subset C$, then

$$m_i := \inf\{h_i(\bar{x}) - \langle c_i, x \rangle : \|\bar{x} - x\| = \delta\}$$

$$= h_i(x) - \langle c_i, x \rangle + \inf\{\sigma_{C-c_i}(\bar{x} - x) : \|\bar{x} - x\| = \delta\}$$

$$\geq h_i(x) - \langle c_i, x \rangle + r_i \delta > h_i(x) - \langle c_i, x \rangle.$$

With $\alpha_i := m_i - (h_i(x) - \langle c_i, x \rangle)$. For every $f_i \in X_C$ with $\rho(f_i, h_i) < \beta_i < \alpha_i/2$, when $\|\bar{x} - x\| = \delta$ we have $f_i(\bar{x}) - \langle c_i, x \rangle = f_i(\bar{x}) - h_i(\bar{x}) + h_i(\bar{x}) - \langle c_i, x \rangle \geq -\beta_i + m_i$, and

$$f_i(x) - \langle c_i, x \rangle = f_i(x) - h_i(x) + h_i(x) - \langle c_i, x \rangle \leq \beta_i + h_i(x) - \langle c_i, x \rangle.$$

As $2\beta_i < \alpha_i$, we have

$$\inf\{f_i(\bar{x}) - \langle c_i, x \rangle : \|\bar{x} - x\| = \delta\} > f_i(x) - \langle c_i, x \rangle. \quad (5.27)$$
5.12 Sum rules and Jacobians

Define \( g_i : A \to (-\infty, +\infty) \) by \( g_i(\bar{x}) := f_i(\bar{x}) - \langle c_i, \bar{x} \rangle + I_{B_i(\bar{x})} \). Then \( g_i \) is bounded below and lower semicontinuous, and

\[
\inf_{\bar{x}} g_i \leq f_i(x) - \langle c_i, x \rangle < \inf_{\bar{x}} \{ f_i(\bar{x}) - \langle c_i, \bar{x} \rangle : \| \bar{x} - x \| = \delta \} = \inf_{\bar{x}} \{ g_i(\bar{x}) : \| \bar{x} - x \| = \delta \}.
\]

Let \( \nu_i \in (0, \min\{1/(2n), \inf_{\bar{x}} g_i(\bar{x}) : \| \bar{x} - x \| = \delta \} - g_i(x) \}) \). Choose \( x_0 \) satisfying \( g_i(x_0) < \inf_{x} g_i + \nu_i \). Applying the Borwein-Preiss smooth variational principle, there exist Gâteaux differentiable \( \phi_i \) and \( v_i \in X \) such that

\[
f_i(\bar{x}) - \langle c_i, \bar{x} \rangle + \phi_i(\bar{x}) \geq f_i(v_i) - \langle c_i, v_i \rangle + \phi_i(v_i) \text{ when } \| \bar{x} - x \| \leq \delta, \quad (5.28)
\]

\[
g_i(v_i) < \nu_i + \inf_{\bar{x}} g_i < \inf_{\bar{x}} \{ g_i(\bar{x}) : \| \bar{x} - x \| = \delta \} ,
\]

\[
\| \nabla \phi_i(v_i) \| < 2 \nu_i < \frac{1}{n} .
\]

Then \( \| v_i - x \| < \delta < 1/n \) by equation (5.29), this together with equation (5.28) show

\[
0 \in \partial_- f_i(v_i) - c_i + \nabla \phi_i(v_i).
\]

Then for some \( v_i^* \in \partial_- f_i(v_i) \) we have \( \| v_i^* - c_i \| = \| \nabla \phi_i(v_i) \| < 1/n \) by equation (5.30).

We proceed to show that \( (f_1 + f_2) \) has the required subdifferential property. Indeed, by equation (5.27) we have

\[
\inf\{(f_1 + f_2)(\bar{x}) - \langle c_1 + c_2, \bar{x} \rangle : \| \bar{x} - x \| = \delta \} \geq \sum_{i=1}^{2} \inf\{f_i(\bar{x}) - \langle c_i, \bar{x} \rangle : \| \bar{x} - x \| = \delta \} > (f_1 + f_2)(x) - \langle c_1 + c_2, x \rangle. \quad (5.31)
\]

Define \( g : A \to (-\infty, +\infty) \) by \( g(\bar{x}) := (f_1 + f_2)(\bar{x}) - \langle c_1 + c_2, \bar{x} \rangle + I_{B_2(\bar{x})} \), which is lower semicontinuous and bounded below. Then by equation (5.31), \( \inf_{\bar{x}} g(\bar{x}) : \| \bar{x} - x \| = \delta \} > g(x) \). Letting \( \nu \in (0, \min\{1/(2n), \inf_{\bar{x}} g(\bar{x}) : \| \bar{x} - x \| = \delta \} - g(x) \}) \). Choose \( x_0 \) satisfying \( g(x_0) < \inf_{x} g + \nu \), with \( \lambda = 1 \) and applying Borwein-Preiss principle we have a Gâteaux differentiable \( \phi \) and \( v \in X \) such that when \( \| \bar{x} - x \| \leq \delta \) we have

\[
(f_1 + f_2)(\bar{x}) - \langle c_1 + c_2, \bar{x} \rangle + \phi(\bar{x}) \geq (f_1 + f_2)(v) - \langle c_1 + c_2, v \rangle + \phi(v), \quad (5.32)
\]

\[
g(v) < \nu + \inf_{\bar{x}} g < \inf_{\bar{x}} \{ g(\bar{x}) : \| \bar{x} - x \| = \delta \} ,
\]

\[
\| \nabla \phi(v) \| < \frac{1}{n}. \quad (5.34)
\]
Then \( \|v - x\| < \delta < 1/n \) by equation (5.33), and this together with equation (5.32) shows

\[
0 \in \partial_- (f_1 + f_2)(v) - (c_1 + c_2) + \nabla \phi(v).
\]

That is, for some \( v^* \in \partial_- (f_1 + f_2)(v) \) we have \( \|(c_1 + c_2) - v^*\| = \|\nabla \phi(v)\| < 1/n \) by equation (5.34). This shows \( (f_1, f_2) \in G^n_{x, c_1, c_2} \). Since this holds for every \( (f_1, f_2) \) satisfying \( \rho(f_1, h_1) < \beta_1, \rho(f_2, h_2) < \beta_2 \), it follows that \( (h_1, h_2) \in \text{int}(G^n_{x, c_1, c_2}) \).

(b) The set \( G_{x, c_1, c_2} := \bigcap_{n=1}^\infty \text{int}(G^n_{x, c_1, c_2}) \) is dense \( G_\delta \) in \( \mathcal{X}_C \times \mathcal{X}_C \). If \( (f_1, f_2) \in G_{x, c_1, c_2} \), then for every \( n \) there exists \( x_n, y_n, z_n \in A \) satisfying \( \|x_n - x\| < 1/n, \|y_n - x\| < 1/n, \|z_n - x\| < 1/n \) such that for some \( x_n^* \in \partial_- f_1(x_n), y_n^* \in \partial_- f_2(y_n), z_n^* \in \partial_- (f_1 + f_2)(z_n) \) we have

\[
\|x_n^* - c_1\| < 1/n, \|y_n^* - c_2\| < 1/n, \text{ and } \|z_n^* - (c_1 + c_2)\| < 1/n.
\]

Letting \( n \to \infty \), we see that \( c_1 \in \partial_a f_1(x) \) and \( c_1 + c_2 \in \partial_a (f_1 + f_2)(x) \).

(c) By Lemma 5.4.1 there exists \( \{c_k\}_{k=1}^\infty \subset \text{int}(C) \) such that

\[
C = w^* \text{cl}\{ \lim_{n_k \to \infty} c_{n_k} : \{c_{n_k}\} \text{ is norm converging subsequences of } \{c_k\} \}.
\]

The set \( G_x := \bigcap_{k=1}^\infty \bigcap_{l=1}^\infty G_{x, c_k, c_l} \) is a dense \( G_\delta \) set in \( \mathcal{X}_C \times \mathcal{X}_C \). If \( (f_1, f_2) \in G_x \) then

\[
c_k \in \partial_a f_1(x), c_l \in \partial_a f_2(x), c_k + c_l \in \partial_a (f_1 + f_2)(x) \text{ for all } k, l,
\]

and so \( C \subset \partial_a f_1(x) \) and \( C + C \subset \partial_a (f_1 + f_2)(x) \). But

\[
\partial_a f_1(x) \subset C, \text{ and } \partial_a (f_1 + f_2)(x) \subset \partial_a f_1(x) + \partial_a f_2(x) \subset C + C = 2C.
\]

Hence \( \partial_a (f_1 + f_2)(x) = 2C \) and \( \partial_a f_1(x) = C \).

(d) Let \( \{x_k\}_{k=1}^\infty \) be a countable norm dense subset of \( A \). The set \( G := \bigcap_{k=1}^\infty G_{x_k} \) is dense \( G_\delta \) in \( \mathcal{X}_C \times \mathcal{X}_C \). If \( (f_1, f_2) \in G \), then \( \partial_a f_1(x_k) = C \), and \( \partial_a (f_1 + f_2)(x_k) = 2C \) for all \( k \). Since \( \partial_a f \) is norm-to-\( w^* \) upper semicontinuous, we have \( \partial_a f_1 \equiv C \) and \( \partial_a (f_1 + f_2) \equiv 2C \). ∎
Chapter 6

Subdifferentials in arbitrary
Banach spaces

In general Banach spaces we obtain weaker, but still highly useful results. In order to obtain appropriate residual sets, we need to work not only on the domain but also on the range of a given Lipschitz function \( f \). Our working environment is still \( \mathcal{X}_T \) defined in Section 5.1

6.1 A dense approximation lemma

**Lemma 6.1.1** Let \( E \) be a Lebesgue measurable subset of \( \mathbb{R} \). Let \( f, g \in \mathcal{X}_T \) then the function
\[
h(x) := \lambda_E((f - g)(x)) + g(x) \in \mathcal{X}_T \quad \text{and} \quad \|h - g\|_\infty \leq \mu(E) \quad \text{where} \quad \lambda_E(y) := \int_{(f-g)(x_0)}^y \chi_E(s) \, ds
\]
for \( y \in \mathbb{R} \). If \( E \) is an open dense set in \( \mathbb{R} \), then for every connected open set \( U \) there exists an open subset \( \hat{U} \subset U \) such that \( h - f \) is constant on \( \hat{U} \).

**Proof.** Suppose \( f, g \in \mathcal{X}_T \), we need to show \( h \in \mathcal{X}_T \). To this end, let us fix \( x \in A \) and choose \( \delta > 0 \) so that \( B_{2\delta}(x) \subseteq A \). For each \( v \in S_X \) we define the function \( K_v : B_\delta(x) \times (0, \delta) \to [0,1] \) by,
\[
K_v(z, \lambda) := \begin{cases} 0 & \text{if} \ (f - g)(z + \lambda v) = (f - g)(z), \\ \int_{(f-g)(z)}^{(f-g)(z+\lambda v)} \chi_E(t) \, dt & \text{otherwise.} \\ \frac{(f-g)(z+\lambda v) - (f-g)(z)}{(f-g)(z)} & \end{cases}
\]

135
Then we have,
\[
\frac{h(z + \lambda v) - h(z)}{\lambda} = (1 - K_v(z, \lambda)) \frac{g(z + \lambda v) - g(z)}{\lambda} + K_v(z, \lambda) \frac{f(z + \lambda v) - f(z)}{\lambda}
\]
\[
\leq \max \left\{ \frac{g(z + \lambda v) - g(z)}{\lambda}, \frac{f(z + \lambda v) - f(z)}{\lambda} \right\},
\]
for all \((z, \lambda) \in B_\delta(x) \times (0, \delta)\). Therefore \(h^0(x; v) \leq \max\{g^0(x; v), f^0(x; v)\}\) for all \(v \in S_X\) and so \(\partial_c h(x) \subseteq \text{co}\{\partial_c g(x), \partial_c f(x)\} \subseteq T(x)\). Since the point \(x\) was arbitrary we have that \(h \in \mathcal{X}_T\). The proof that \(\rho(g, h) \leq \mu(E)\) is obvious.

Now assume \(E\) is dense open in \(I\). Let \(U\) be a connected open subset of \(X\). If \(\text{diam}(f - g)(U) = 0\), then \(f - g\) is constant on \(U\). If \(\text{diam}(f - g)(U) > 0\), as \((f - g)(U)\) is a nondegenerate interval, we see that \([(f - g)(U)] \cap E\) contains an open interval \((r, s)\). Choose \(z \in U\) such that \((f - g)(z) = (r + s)/2\). Then there exists \(\delta > 0\) such that for \(\|x - z\| \leq \delta\) we have \((f - g)(x) \in (r, s)\). Now for \(\|x - z\| \leq \delta\) we have:
\[
h(x) - h(z) = \int_{(f - g)(z)}^{(f - g)(x)} \chi_E(t) dt + g(x) - g(z)
\]
\[
= (f - g)(x) - (f - g)(z) + g(x) - g(z) = f(x) - f(z),
\]
then \(h - f\) is constant on \(B_\delta(z)\).

\[
\square
\]

## 6.2 Approximate subdifferentials in smooth Banach spaces

**Theorem 6.2.1** Let \(A\) be a nonempty open subset of a smoothable Banach space \(X\). Let \(T : A \to 2^{X^*}\) be a \(w^*-\)cuso on \(A\). If \(f \in \mathcal{X}_T\), then the set
\[
\{g \in \mathcal{X}_T : \partial_a g(x) \cap (-\partial_a (-f)(x)) \neq \emptyset \text{ for all } x \in A\}
\]
is residual in \((\mathcal{X}_T, \rho)\).

In particular, the set \(\{g \in \mathcal{X}_T : \partial_a f(x) \subset \partial_a g(x) \text{ for all } x \in A\}\) is residual in \((\mathcal{X}_T, \rho)\) when \(\partial_a f\) is a minimal \(w^*-\)usco on \(A\).

**Proof.** For each \(m \in \mathbb{N}\), let \(A_m := \text{int}\{t \in A : T(t) \subseteq mB_{X^*}\}\). Then by Lemma 5.1.2, \(A = \bigcup_{m \in \mathbb{N}} A_m\) and each \(g \in \mathcal{X}_T\) is \(m\)-Lipschitz on each convex subset of \(A_m\). Let \(J := \{J_n : n \in \mathbb{N}\}\) be an enumeration of all the open intervals in \(I\) with rational end-points.

For each \((m, n, p, \varepsilon) \in \mathbb{N}^3 \times (0, \infty)\) we consider the set,
\[
O_{(m, n, p, \varepsilon)} := \{g \in \mathcal{X}_T : \text{ for each connected open set } U \text{ with } U + \frac{1}{p}B_X \subseteq A_m\}
\]
and $J_n \subseteq (f - g)(U)$ there exists a $z_0 \in U$ and $0 < r_0 < 1/p$
so that $\inf_{B_{r_0}[z_0]}(g - f)(z) + \varepsilon r_0 > (g - f)(z_0)$.

(a) For each $(m, n, p, \varepsilon) \in \mathbb{N}^3 \times (0, \infty)$, $\text{int} \ O_{(m, n, p, \varepsilon)}$ is dense in $(\mathcal{X}_T, \rho)$.
Suppose $(g_0, \delta) \in \mathcal{X}_T \times (0, 1)$, we need to verify that $B_\delta(g_0) \cap \text{int} \ O_{(m, n, p, \varepsilon)} \neq \emptyset$. To this end, suppose $J_n := (r_n, s_n)$ and $0 < \delta' := \min\{(s_n - r_n)/5, \delta\}$. Now let us choose a dense open subset $E$ of $\mathbb{R}$ such that $\mu(E) < \delta'$ and define $h : A \to \mathbb{R}$ by,
$$h(x) := \lambda_E((f - g_0)(x)) + g_0(x).$$
By Lemma 6.1.1 we have $h \in \mathcal{X}_T$ and $\rho(g_0, h) < \delta' \leq \delta$. We claim that $h \in \text{int} \ O_{(m, n, p, \varepsilon)} \cap B_\delta(g_0)$. To this end, choose $0 < r < 2m/p$ and $t \in \mathbb{R}$ so that $[t - r, t + r] \subseteq (r_n + 2\delta', s_n - 2\delta') \cap E$ and set $0 < d < \min\{(\varepsilon r)/4m, \delta'\}$. We will show that $B_d(h) \subseteq O_{(m, n, p, \varepsilon)}$. Let $g \in B_d(h)$ and let $U$ be any connected open subset of $A_m$ with $U + 1/pB_X \subseteq A_m$ and $J_n \subseteq (f - g)(U)$. Then $[t - r, t + r] \subseteq (r_n + 2\delta', s_n - 2\delta') \subseteq (f - g_0)(U)$ since $(f - g_0)(U)$ is connected (hence convex) and
$$\| (f - g_0) - (f - g) \|_\infty = \| g_0 - g \|_\infty \leq \| g_0 - h \|_\infty + \| g - h \|_\infty < \delta' + d \leq 2\delta'.$$
Choose $z_0 \in U$ so that $(f - g_0)(z_0) = t$ then for any $z \in B_{r_0}[z_0]$ with $r_0 = r/2m < 1/p$,
$$ (f - g_0)(z) \in [t - r, t + r] \subseteq E \text{ and so } h(z) - h(z_0) = f(z) - f(z_0).$$
Therefore by our choice of $d$,
$$ (g - f)(z) - (g - f)(z_0) = (g - h)(z) - (g - h)(z_0) > -2d > -\varepsilon r_0,$$
for all $z \in B_{r_0}[z_0]$. This shows that $g \in O_{(m, n, p, \varepsilon)}$.

(b) The set $G := \bigcap \{O(n_1, n_2, n_3, 1/n_4) : (n_1, n_2, n_3, n_4) \in \mathbb{N}^4\}$ is residual in $(\mathcal{X}_T, \rho)$, and for each $g \in G$ we have $\partial g(x) \cap (-\partial(-f)(x)) \neq \emptyset$ for every $x \in A$.

Indeed, if this is not the case then there exists a $g \in G$ and $x_0 \in A$ such that $\partial g(x_0) \cap (-\partial(-f)(x_0)) = \emptyset$, then for some $w^*$-neighborhood $V$ of zero,
$$ (\partial g(x_0) + V) \cap (-\partial(-f)(x_0) + V) = \emptyset. \quad (6.1)$$
Choose $n_4 \in \mathbb{N}$ and another $w^*$-neighbourhood of $0$ such that $2W + 2/n_4B_X \subset V$. If there exists a neighbourhood $B_\delta(x_0)$ such that $(f - g)(B_\delta(x_0)) = \{a\}$, then $g - f \equiv a$ on $B_\delta(x_0)$ which is impossible since
$$ \partial g(x_0) \cap (-\partial(-f)(x_0)) = \partial f(x_0) \cap (-\partial(-f)(x_0)) \neq \emptyset.$$
6.2 Approximate subdifferentials in smooth Banach spaces

Therefore, for every \( n \in \mathbb{N} \) there exists \( J_n \subset (f - g)(B_{1/n}(x_0)) \). We may now select \( n_1, n_3 \in \mathbb{N} \) so that \( B_{2/n_3}(x_0) \subseteq A_{n_1} \), and for \( x \in B_{2/n_3}(x_0) \) we have

\[
-\partial_a(-f)(x) \supset -\partial_a(-f)(x_0) + W; \quad \partial_ag(x) \subset \partial_ag(x_0) + W.
\]

For \( B_{1/n_3}(x_0) \) there is some \( n_2 \in \mathbb{N} \) so that \( J_{n_2} \subseteq (f - g)(B_{1/n_3}(x_0)) \). Since \( g \in O(n_1, n_2, n_3, 1/n_4) \), there exist \( z_0 \in B_{1/n_3}(x_0) \) and \( 0 < r_0 < 1/n_3 \) such that

\[
\inf_{B_{r_0}[z_0]} (g - f) + \frac{r_0}{n_4} > (g - f)(z_0).
\]

By the Borwein-Preiss variational principle, there exists a Gâteaux smooth \( \phi \) (see page 5) on \( X \) and and \( \tilde{z}_0 \in X \) such that \( \|z_0 - \tilde{z}_0\| < r_0 \), \( \|\nabla \phi(\tilde{z}_0)\| \leq 2/n_4 \), and

\[
(g - f)(z) + \phi(z) \geq (g - f)(\tilde{z}_0) + \phi(\tilde{z}_0) \quad \text{for all } z \in B_{r_0}[\tilde{z}_0].
\]

As \( \nabla \phi \) is norm-to-\( w^* \) continuous, \( \partial_a \phi(\tilde{z}_0) \subset \nabla \phi(\tilde{z}_0) + W \), we have

\[
0 \in \partial_a(g)(\tilde{z}_0) + \partial_a(-f)(\tilde{z}_0) + \frac{2}{n_4}B_{X^*} + W.
\]

Therefore, there exists \( \tilde{z}_0^* \) such that \( \tilde{z}_0^* \in -\partial_a(-f)(\tilde{z}_0) \subset -\partial_a(-f)(x_0) + W \), and

\[
\tilde{z}_0^* \in \partial_a(g(\tilde{z}_0)) + \frac{2}{n_4}B_{X^*} + W \subset \partial_ag(x_0) + W + W + \frac{2}{n_4}B_{X^*} \subset \partial_ag(x_0) + V.
\]

which is impossible by equation (6.1). Therefore, for each \( g \in G \),

\[
(-\partial_a(-f)(x)) \cap \partial_ag(x) \neq \emptyset \quad \text{for all } x \in A.
\]

When \( \partial_a f \) is a minimal \( w^* \)-usco, \( \partial_a f \) is single-valued on a dense \( G_\delta \) set \( D \subset A \) by [76]. For \( x \in D \), \( \partial_a f(x) = -\partial_a(-f)(x) = \{\nabla f(x)\} \). The result now follows by Lemma 2.4.5. \( \square \)

While minimality of an usco is restrictive, it does hold for \( \partial_a f \) when the Gâteaux derivative of \( f \) is norm-to-\( w^* \) continuous or \( f \) is a concave continuous function on \( A \). Even this allows for nice applications:

**Corollary 6.2.2** Let \( A \) be a non-empty open subset of a smoothable Banach space \( X \). Suppose \( C \subset X^* \) is \( w^* \)-compact convex and \( w^* \)-separable, then in \((X_C, \rho)\) the set

\[
\{g \in X_C : \partial_ag \equiv C \text{ on } A\}
\]

is residual.
6.3 Clarke subdifferentials in arbitrary Banach spaces

When $X$ is a separable Banach space, every $w^*$-compact set is $w^*$-separable, thus Corollary 6.2.2 recovers Corollary 5.6.3. By Goldstine’s theorem, for any normed linear space $X$, $B_X$ is $w^*$-dense in $B_{X^{**}}$ [38]. If $X$ is separable, then $B_{X^{**}}$ is $w^*$-separable. If $X^*$ has a smooth renorm, Corollary 6.2.2 applies for $B_{X^{**}}$. Note that if $X^*$ has a Fréchet renorm, then $X$ is reflexive! If $X^*$ has a Gâteaux differentiable norm, then $X$ does not contain an isomorphic copy of $l_1$ [37, page 52]. Compare Theorem 6.2.1 to:

**Theorem 6.2.3** (Preiss, Phelps and Namioka) [72, 76] A continuous convex function on an open convex subset of a smoothable Banach space is Gâteaux differentiable on a dense $G_δ$ subset of its domain.

6.3 Clarke subdifferentials in arbitrary Banach spaces

**Theorem 6.3.1** Let $A$ be a non-empty open subset of a Banach space $X$ and $T : A \to 2^{X^*}$ be a $w^*$-cusco on $A$. Then for each $f \in X_T$, the set

$$\{g \in X_T : \partial_c g(x) \cap \partial_c f(x) \neq \emptyset \text{ for all } x \in A\}$$

is residual in $(X_T, \rho)$. In particular, the set $\{g \in X_T : \partial_c g(x) \supseteq \partial_c f(x) \text{ for all } x \in A\}$ is residual in $(X_T, \rho)$ when $\partial_c f$ is a minimal $w^*$-cusco on $A$.

**Proof.** For each $m \in \mathbb{N}$, let $A_m := \text{int}\{t \in A : T(t) \subseteq mB_{X^*}\}$. Then by Lemma 5.1.2, $A = \bigcup_{m \in \mathbb{N}} A_m$ and each $g \in X_T$ is $m$-Lipschitz on each convex subset of $A_m$. Let $J := \{J_n : n \in \mathbb{N}\}$ be an enumeration of all the open intervals in $\mathbb{R}$ with rational end-points. For each $(m, n, \varepsilon) \in \mathbb{N}^2 \times (0, +\infty)$ consider the set,

$$O_{(m, n, \varepsilon)} := \{g \in X_T : \text{ for each connected open set } U \text{ with } U + \varepsilon B_X \subseteq A_m \text{ and } J_n \subseteq (f - g)(U) \text{ there exist } z_0 \in U \text{ and } 0 < \lambda_0 < \varepsilon \text{ so that } (g - f)(z) - (g - f)(z_0) \geq -\lambda_0 \varepsilon \text{ for all } z \in B_{\lambda_0}[z_0]\}.$$

(a) $\text{int} O_{(m, n, \varepsilon)}$ is dense in $(X_T, \rho)$ for each $(m, n, \varepsilon) \in \mathbb{N}^2 \times (0, +\infty)$.

Suppose $(g_0, \delta) \in X_T \times (0, 1)$. We need to verify that $B_\delta(g_0) \cap \text{int} O_{(m, n, \varepsilon)} \neq \emptyset$. To this end, suppose $J_n := (r_n, s_n)$ and $\delta' := \min\{(s_n - r_n)/5, \delta\}$. Now let us choose a dense open subset
6.3 Clarke subdifferentials in arbitrary Banach spaces

Let $E$ be a Banach space and let $f : E \to \mathbb{R}$ be a function. Define $g : E \to \mathbb{R}$ by $g(x) := \lambda E((f - g_0)(x)) + g_0(x)$.

Then by Lemma 6.1.1, $h \in \mathcal{X}_T$ and $\rho(g_0, h) < \delta' \leq \delta$. We claim that $h \in \text{int}(O_{(m,n,\varepsilon)} \cap B_\delta(g_0))$. To this end, choose $0 < r < 2m\varepsilon$ and $t \in \mathbb{R}$ so that $[t - r, t + r] \subseteq (r_n + 2\delta', s_n - 2\delta') \cap E$ and set $d := \min \{ (\varepsilon r)/(4m), \delta' \}$. We show that $B_d(h) \subseteq O_{(m,n,\varepsilon)}$. Let $g \in B_d(h)$ and let $U$ be any connected open subset of $A_m$ with $U + \varepsilon B_X \subseteq A_m$ and $J_n \subseteq (f - g)(U)$. Since $(f - g_0)(U)$ is connected (hence convex) and

$$\| (f - g) - (f - g_0) \|_\infty = \| g_0 - g \|_\infty \leq \| g_0 - h \|_\infty + \| g - h \|_\infty \leq d + \delta' \leq 2\delta',$$

we have $[t - r, t + r] \subseteq (r_n + 2\delta', s_n - 2\delta') \subseteq (f - g_0)(U)$. Choose $z_0 \in U$ so that $(f - g_0)(z_0) = t$, then for $\lambda_0 = r/(2m) < \varepsilon$ we have $(f - g_0)(z) \in [t - r, t + r] \subseteq E$ for $z \in B_{\lambda_0}[z_0]$, thus,

$$(g - f)(z) - (g - f)(z_0) = (g - h)(z) - (g - h)(z_0) \geq -2d \geq -\lambda_0 \varepsilon.$$

This shows that $g \in O_{(m,n,\varepsilon)}$.

(b) The set $G := \bigcap \{ O_{(n_1,n_2,1/n_3)} : (n_1,n_2,n_3) \in \mathbb{N}^3 \}$ is residual in $(\mathcal{X}_T, \rho)$ and for each $g \in G$ we have $\partial_c g(x) \cap \partial_c f(x) \neq \emptyset$ for all $x \in A$. Indeed, if this is not the case then there exists a $g \in G$ and $x_0 \in A$ such that $\partial_c g(x_0) \cap \partial_c f(x_0) = \emptyset$. By the strong separation theorem, there exists a $y \in S_X$, $\alpha \in \mathbb{R}$ and $\varepsilon > 0$ such that, $-f^0(x_0, -y) = \min \{ x^*(y) : x^* \in \partial_c f(x_0) \} > \alpha + \varepsilon > \alpha - \varepsilon > \max \{ x^*(y) : x^* \in \partial_c g(x_0) \} = g^0(x_0, y)$.

Now $x_0 \in A_{n_1}$ for some $n_1 \in \mathbb{N}$, and by the definitions of $-f^0(x_0, -y)$ and $g^0(x_0, y)$ there exists an $n_3 \in \mathbb{N}$ such that $1/n_3 < \varepsilon$, $B_{1/n_3}(x_0) \subseteq A_{n_1}$ and

$$\inf \left\{ \frac{f(z + \lambda y) - f(z)}{\lambda} : 0 < \lambda \leq \frac{1}{n_3}, z \in B_{1/n_3}(x_0) \right\} \geq \alpha + \frac{1}{n_3} > \alpha - \frac{1}{n_3} \geq \sup \left\{ \frac{g(z + \lambda y) - g(z)}{\lambda} : 0 < \lambda \leq \frac{1}{n_3}, z \in B_{1/n_3}(x_0) \right\}. \quad (*)$$

Let $U := B_{1/n_3}(x_0)$, then $U + 1/n_3 B_X \subseteq A_{n_1}$. Now $(f - g)(U)$ is convex, so either $(f - g)(U) = \{ a \}$ for some $a \in \mathbb{R}$ or $J_n \subseteq (f - g)(U)$ for some $n \in \mathbb{N}$. In the first case, we get that $f(x) - g(x) + a$ on $U$ which is impossible since $\partial_c f(x_0) \neq \partial_c g(x_0)$. Therefore, there exists some $n_2 \in \mathbb{N}$ so that $J_{n_2} \subseteq (f - g)(U)$. Since $g \in G$, we have that $g \in O_{(n_1,n_2,1/n_3)}$ and so there exists $z_0 \in U$ and $0 < \lambda_0 < 1/n_3$ so that,

$$\frac{(g - f)(z) - (g - f)(z_0)}{\lambda_0} \geq -\frac{1}{n_3} \text{ for every } z \in B_{\lambda_0}[z_0],$$
6.3 Clarke subdifferentials in arbitrary Banach spaces

which contradicts (*). Therefore, for each \( g \in G \) we have \( \partial_c f(x) \cap \partial_c g(x) \neq \emptyset \) for each \( x \in A \). The case when \( \partial_c f \) is a minimal \( w^* \)-cusco follows from Lemma 2.4.5. \( \square \)

In part (b) of the proof of Theorem 6.3.1, instead of using the strong separation theorem, one can use Ekeland’s variational principle. To use Theorem 6.3.1, we need Lipschitz functions with minimal Clarke subdifferential mappings. In [8] the authors have shown that each essentially smooth locally Lipschitz function on \( A \) gives rise to a minimal Clarke subdifferential.

**Theorem 6.3.2** Let \( A \) be a non-empty open subset of a Banach space \( X \). Let \( \{f_i\}_{i=1}^{\infty} \) be a sequence of locally equi-Lipschitz real-valued functions. Define \( T : A \to 2^{X^*} \) by \( T(x) := \text{CSC} [\bigcup_{i=1}^{\infty} \partial_c f_i ](x) \). Suppose each \( \partial_c f_i \) is a minimal \( w^* \)-cusco, then in \( (X_T, \rho) \) the set \( \{ g \in X_T : \partial_c g \equiv T(x) \text{ for each } x \in A \} \) is residual.

**Proof.** By Theorem 6.3.1, the set \( G_i := \{ g \in X_T : \partial_c g(x) \supset \partial_c f_i(x) \text{ for every } x \in A \} \) is residual in \( (X_T, \rho) \). If \( g \in G := \bigcap_{i=1}^{\infty} G_i \), then \( \partial_c g(x) \supset \bigcup_{i=1}^{\infty} \partial_c f_i(x) \). Since \( T \) is the minimal \( w^* \)-cusco containing \( \bigcup_{i=1}^{\infty} \partial_c f_i \), we have \( T(x) \subset \partial_c g(x) \) for every \( x \in A \). But \( g \in X_T \) implies \( \partial_c g(x) \subset T(x) \) for every \( x \in A \). \( \square \)

**Corollary 6.3.3** Let \( A \) be a nonempty open convex subset of a Banach space \( X \) and \( \{T_i : i = 1 \ldots n\} \) be a finite family of maximal cyclically monotone operators from \( A \) into non-empty subsets of \( X^* \). Then there exists a real-valued locally Lipschitz function \( g \) defined on \( A \) such that \( \partial_c g(x) = \text{co}\{T_1(x), T_2(x), \ldots, T_n(x)\} \) for every \( x \in A \).

**Proof.** Since each maximal cyclically monotone operator is a minimal cusco and the Clarke subdifferential of a continuous convex function on \( A \) by Proposition 3.26 [72], for each \( T_i \), we may find a locally Lipschitz \( f_i : A \to R \) such that \( \partial_c f_i = T_i \). With \( T : A \to 2^{X^*} \) defined by \( T(x) := \text{co}\{T_1(x), \ldots, T_n(x)\} \), we apply Theorem 6.3.1. \( \square \)

This generalizes Corollary 2 [21] to nonseparable spaces.

**Example 6.3.4** Some Clarke subdifferentials can not be expressed as the cusco generated by a countable family of minimal cuscors. Indeed, let \( f : \mathbb{R} \to \mathbb{R} \) be a differentiable nowhere monotone Lipschitz function, then \( \{ x : f \text{ is strictly differentiable at } x \text{ and } f'(x) = 0 \} \) is residual in \( \mathbb{R} \) and so the only minimal cusco lying inside \( \partial_c f \) is \( T \equiv \{0\} \) (see page 58).
6.3 Clarke subdifferentials in arbitrary Banach spaces

Often local minimality of \( \partial_c f \) suffices in applications. For a Banach space \( X \), we call a set \( S \subset X \) an \( \varepsilon \)-net if (a) \( \|x - y\| \geq \varepsilon \) for any two distinct points \( x, y \) of \( S \) and (b) \( S \) is maximal with respect to (a). Zorn's lemma yields that \( \varepsilon \)-nets exist for every \( \varepsilon > 0 \).

**Corollary 6.3.5** Let \( C \) be any \( w^\ast \)-compact convex subset in \( X^\ast \) with non-empty norm interior. In \( (\mathcal{X}_C, \rho) \) the set \( \{ f \in \mathcal{X}_C : \partial_c f \equiv C \} \) is residual. In particular, this holds when \( C = B_{X^\ast} \).

**Proof.** Without loss of generality, we may assume \( \delta B_{X^\ast} \subset C \) for some \( \delta > 0 \). Since \( C \) is \( w^\ast \)-compact, \( C \) has a norm bound \( M > 0 \). Choose a \( 2/n \)-net in \( A \) denoted by \( \{ x_n^\alpha : \alpha \in \Gamma_n \} \). For each \( x_n^\alpha \), we define \( f_n^\alpha(x) := \sigma_C(x - x_n^\alpha) \). Since \( 0 \in C \) and \( C \) is norm bounded, \( f_n^\alpha \geq 0 \) and is finite-valued everywhere. Define \( f_n(x) := \inf \{ f_n^\alpha(x) : \alpha \in \Gamma_n \} \). Then \( f_n \in \mathcal{X}_C \). When \( \|x - x_n^\alpha\| \leq \delta/(Mn) \) we have \( \sigma_C(x - x_n^\alpha) \leq M\|x - x_n^\alpha\| \leq \frac{\delta}{n} \), and

\[
\sigma_C(x - x_n^\beta) = \sup_{c \in C} (\langle x - x_n^\alpha, c \rangle + \langle x_n^\beta - x_n^\alpha, c \rangle) \geq \sup_{c \in C} (\langle x_n^\alpha - x_n^\beta, c \rangle - \sigma_C(x - x_n^\alpha)) \\
\geq \sigma_C(x_n^\alpha - x_n^\beta) - \sigma_C(x - x_n^\alpha) \geq 2\delta/n - M\|x - x_n^\alpha\| \geq \frac{\delta}{n},
\]

thus \( f_n(x) = \sigma_C(x - x_n^\alpha) \) on \( B_{\delta/(Mn)}(x_n^\alpha) \) and this holds for every \( \alpha \in \Gamma_n \). For each \( f_n \), the set \( G_n := \{ f : \partial_c f(x) \cap \partial_c f_n(x) \neq \emptyset \text{ for every } x \in A \} \) is residual \( \mathcal{X}_C \). On each \( B_{\delta/(Mn)}(x_n^\alpha) \), because \( \partial_c f_n \) is minimal, we have \( C \supset \partial_c f(x_n^\alpha) \supset \partial_c f_n(x_n^\alpha) = C \), that is, \( \partial_c f(x_n^\alpha) = C \). The set \( G := \bigcap_{n=1}^{\infty} G_n \) is residual in \( \mathcal{X}_C \). If \( f \in G \), then \( \partial_c f(x_n^\alpha) = C \) holds for every \( n \in \mathbf{N} \) and \( \alpha \in \Gamma_n \). Since \( \bigcup_{n=1}^{\infty} \{ x_n^\alpha : \alpha \in \Gamma_n \} \) is dense in \( A \) and \( \partial_c f \) is norm-to-\( w^\ast \) cusc, we obtain \( C \subset \partial_c f(x) \) for every \( x \in A \). Hence \( \partial_c f \equiv C \) on \( A \). \( \square \)

**Corollary 6.3.6** Let \( A \) be a non-empty open subset of a Banach space \( X \). Then

\[
\{ g \in \mathcal{X}_{B_{X^\ast}} : \text{for each } v \in S_X, \{ x \in A : g'(x; v) = -g'(x; -v) \} \text{ is first category} \},
\]
is residual in \( (\mathcal{X}_{B_{X^\ast}}, \rho) \).

**Proof.** Let \( G := \{ g \in \mathcal{X}_{B_{X^\ast}} : \partial_0 g(x) = B_{X^\ast} \text{ for all } x \in A \} \). We claim that for each \( g \in G \), \( D_y := \{ x \in A : g'(x; y) = -g'(x; -y) \} \) is first category in \( A \) for each \( y \in S_X \). To see this, let us fix \( y \in S_X \). Then by, by Lemma 5.5.2 there exists a dense \( G_\delta \) subset \( P_y \) of \( A \) where \( g^0(x; y) = g^+(x; y) \) and \( -g^0(x; -y) = -g^+(x; -y) \) for each \( x \in P_y \). We will now show that \( D_y \subset A \setminus P_y \). Indeed, if \( x_0 \in P_y \cap D_y \) then,

\[
1 = g^0(x_0; y) = g^0(x_0; y) = -g^0(x_0; -y) = -g^0(x_0; -y) = -1,
\]
which is absurd. Therefore, $D_y \subseteq A \setminus P_y$ and so first category in $A$. □

In contrast, every essentially smooth function $f$ is, generically on $A$, strictly differentiable for each direction $v \in S_X$ by Proposition 1.4.2.

**Theorem 6.3.7** Let $X$ be a Banach space with a smooth norm. Then in $(\mathcal{X}_{B_X}, \rho)$ the set 
\[ \{ g \in \mathcal{X}_{B_X} : \partial_a g = \partial_c g \equiv B_{X^*} \} \]

is residual.

**Proof.** If $X$ is separable, this follows from Theorem 5.4.2. If $X$ is nonseparable, by
Corollary 6.3.5, the set $G := \{ g \in \mathcal{X}_{B_X} : \partial_a g \equiv B_{X^*} \}$ is residual. Let $g \in G$. As
\[ \partial_c g(x) = \overline{w^*} \partial_a g(x) \] and \[ \partial_a g(x) \] is $w^*$-closed, the converse of the Krein-Milman theorem
shows $\text{ext } B_{X^*} \subset \partial_a g(x)$. Since $X$ has a smooth norm, $S_{X^*} \subset \partial_a g(x)$ by Proposition 2.8.2,
so $B_{X^*} = w^* \text{cl } S_{X^*} \subset \partial_a g(x)$ by Proposition 2.8.1. Hence $\partial_a g = \partial_c g \equiv B_{X^*}$.

However, we do not know whether Theorem 6.3.7 holds or not when $X$ is not smooth or separable. Following Borwein and Fitzpatrick [14], we define the sequential $\beta$-derivative and the topological $\beta$-derivative respectively by
\[ D^\beta_\sigma f(x) := \{ w^* - \lim \nabla \beta f(x_n) : x_n \to x \}, \text{ and } D^\beta_\sigma f(x) := \bigcap_{n=1}^{\infty} w^* \text{cl } \{ \nabla \beta f(y) : \| y - x \| < \frac{1}{n} \}. \]

By Theorem 2.6 [14]: If $B_{X^*}$ is weak$^*$ sequentially compact, then $D^\beta_\sigma f(x) = w^* \text{cl } D^\beta_\sigma f(x)$; If $X$ is WCG, then $D^\beta_\sigma f(x) = D^\beta_\sigma f(x)$. Preiss [74] showed that every locally Lipschitz function $f : A \to \mathbb{R}$ satisfies: \[ \partial_c f(x) = \overline{w^*} D^\beta_\sigma f(x) \text{ if } X \text{ is } \beta\text{-smooth}; \] and \[ \partial_c f(x) = \overline{w^*} D^\sigma f(x) \] if $X$ is Asplund. Let $B_{X^*}$ be the dual ball of a $\beta$-smooth renorm. Suppose that $f$ is a locally Lipschitz function on $A$ such that $\partial_c f(x) = B_{X^*}$ for every $x \in A$. By the converse of the Krein-Milman theorem, $\text{ext } B_{X^*} \subset D^\beta_\sigma f(x)$. By Proposition 2.8.2, $S_{X^*} \subset D^\beta_\sigma f(x)$. By Proposition 2.8.1, when $X$ is infinite dimensional we have $B_{X^*} \subset D^\beta_\sigma f(x)$. On each Asplund space with a Gâteaux smooth norm, similar arguments show that $B_{X^*} \subset D^\sigma f(x)$. Hence

**Corollary 6.3.8** Let $X$ be an infinite dimensional Banach space.

(i) If $X$ is endowed with a $\beta$-smooth norm, then the set \{ $f \in \mathcal{X}_{B_X} : D^\beta_\sigma f \equiv B_{X^*}$ on $A$ \} is residual in $(\mathcal{X}_{B_X}, \rho)$. When $X$ is also WCG, one may replace $D^\beta_\sigma f$ by $D^\beta_\sigma f$.

(ii) If $X$ is a smooth Asplund space, then the set \{ $f \in \mathcal{X}_{B_X} : D^\sigma_\sigma f \equiv B_{X^*}$ on $A$ \} is residual in $(\mathcal{X}_{B_X}, \rho)$. When $X$ is also WCG, one may replace $D^\sigma_\sigma f$ by $D^\sigma_\sigma f$. 


Since both $D^0 f$ and $D^1 f$ are measures of the $\beta$-gradient oscillation of $f$ at $x$, Corollary 6.3.8 says that on infinite dimensional $\beta$-smooth spaces, generically non-expansive Lipschitz functions have maximally oscillatory gradients.

Finally, the reader may compare the results in this section with:

**Theorem 6.3.9 (Kenderov)** In a Banach space $X$, for every continuous convex function $f$ on an open convex subset $A$ of $X$ there exists a dense $G_\delta$ subset of $A$ at each point $x$ of which $\partial_c f(x)$ lies in the face of a sphere of $X^*$.

For the proof, see [49, pages 135-137]. If $X$ can be equivalently renormed so that the dual $X^*$ is strictly convex, Theorem 6.3.9 shows every continuous convex function $f$ on an open convex subset $A$ of $X$ is Gâteaux differentiable on a dense $G_\delta$ subset of $A$.

### 6.4 $\sigma$-minimal Clarke subdifferentials plus the dual ball

We say that a Clarke subdifferential map $T$ is $\sigma$-minimal if there exists a countable family of minimal Clarke subdifferential maps $\{\partial_c f_i : i \in \mathbb{N}\}$ such that $T = \text{CSC}(\bigcup_{i \in \mathbb{N}} \partial_c f_i)$.

**Lemma 6.4.1** Let $f$ be a locally Lipschitz function defined on a non-empty open subset $A$ of a Banach space $X$. Assume $\partial_c f$ is a minimal $w^*$-cuso. Then for each $y \in X$ and each dense set $D \subset A$ we have $f^0(x; y) = \limsup_{z \to x} f^-(z; y)$.

**Proof.** Given a dense set $D$, suppose for some $y$, $f^0(x; y) > \alpha > \limsup_{z \to x} f^-(z; y)$. Then there exists an open neighbourhood $U$ of $x$ such that $f^-(z; y) < \alpha$ for all $z \in D \cap U$. On the other hand, $f^0(x; y) > \alpha$ shows $\partial_c f(U) \nsubseteq \{x^* \in X^* : \langle x^*, y \rangle \leq \alpha\}$. By Lemma 2.6.5 there exists a non-empty open $V \subset U$ such that $\partial_c f(V) \cap \{x^* \in X^* : \langle x^*, y \rangle \leq \alpha\} = \emptyset$. Then for all $z \in V \cap D$, $f^-(z; y) \geq -f^0(z; -y) \geq \alpha$, a contradiction. \hfill \Box

**Lemma 6.4.2** Let $T$ be a $w^*$-cuso defined on a non-empty open subset $A$ of a Banach space $X$. If $\partial_c f$ is a minimal $w^*$-cuso and $\partial_c f(x) + B_{X^*} \subseteq T(x)$ for all $x \in A$, then

\[ \{g \in X_T : \partial_c g(x) \supseteq \partial_c f(x) + B_{X^*} \text{ for all } x \in A\} \text{ is residual in } (X_T, \rho). \]

**Proof.** For each $n \in \mathbb{N}$ we choose a $2/n$-net in $A$ denoted by $C_n := \{x^n_\alpha : \alpha \in \Gamma_n\}$ in $A$, and define $f_n : A \to \mathbb{R}$ by $f_n(x) := f(x) + d_{C_n}(x)$. Let $G_n$ be any residual set in $(X_T, \rho)$
6.5 When is $\mathcal{X}_T$ non-empty?

such that for each $g \in G_n$, $\partial g(x) \cap \partial f_n(x) \neq \emptyset$ for all $x \in A$. Set $G := \cap_{n=1}^{\infty} G_n$. We will show that for each $g \in G$, $\partial g(x) \supseteq B_{X^*}$ for all $x \in A$. Let $g \in G$. On $B_{1/n}(x_n^0)$, $f_n(x) = f(x) + \|x - x_n^0\|$, thus $\partial f_n$ is a minimal cusc on $B_{1/n}(x_n^0)$ by Lemma 2.6.8, and so $\partial f_n \subseteq \partial g$ on $B_{1/n}(x_n^0)$. In particular, for each $y \in X$ we have

$$f^-(x_n^0; y) + \|y\| = f_n^-(x_n^0; y) \leq f_n^0(x_n^0; y) \leq g^0(x_n^0; y).$$

As $C := \{x_n^0 : n \in \mathbb{N}, \alpha \in \Gamma_n\}$ is dense in $A$, by Lemma 6.4.1 we obtain

$$f^0(x; y) + \|y\| = \lim_{z \to y} \sup_{z \in C} f^-(z; y) + \|y\| \leq \lim_{z \to y} \sup_{z \in C} g^0(z; y) \leq g^0(x; y).$$

Hence, $\partial g(x) \supseteq \partial f(x) + B_{X^*}$ for all $x \in A$.  

Theorem 6.4.3 Assume $\{f_j : j \in \mathbb{N}\}$ are locally equi-Lipschitz on an open subset $A$ of a Banach space $X$ with each $\partial f_j$ being minimal $w^*$-cusc. Define $\Omega : A \to 2^{X^*}$ by $\Omega(x) := \bigcup_{j \in \mathbb{N}} \partial f_j(x)$ for $x \in A$. Then $\text{CSC}(\Omega) + B_{X^*}$ is a Clarke subdifferential. That is, the sum of a $\sigma$-minimal Clarke subdifferential and the dual ball is a Clarke subdifferential.

Proof. The multifunction $T : A \to 2^{X^*}$ defined by $T := \text{CSC}(\Omega) + B_{X^*}$ is a $w^*$-cusc. For each $f_j$, by Lemma 6.4.2 there exists a residual set $G_j$ in $(X_T, \rho)$ such that $\partial f_j(x) + B_{X^*} \subseteq \partial g(x)$ for all $x \in A$ when $g \in G_j$. Then $G := \cap_{j \in \mathbb{N}} G_j$ is residual in $(X_T, \rho)$. If $g \in G$ we have $\partial f_j(x) + B_{X^*} \subseteq \partial g(x)$ for all $x \in A$ and $j \in \mathbb{N}$, then

$$\text{CSC}(\Omega) + B_{X^*} \subseteq \text{CSC}[\Omega + B_{X^*}] \subseteq \partial g \subseteq \text{CSC}(\Omega) + B_{X^*}.$$

Hence $\text{CSC}(\Omega) + B_{X^*} = \partial g$ for $g \in G$.  

In particular, the sum of a maximal cyclically monotone operator and $B_{X^*}$ is a Clarke subdifferential. Theorem 6.4.3 still holds if one replace $B_{X^*}$ by any $w^*$-compact convex subset with non-empty norm interior.

6.5 When is $\mathcal{X}_T$ non-empty?

We have not thus far dwelt much upon the question of when $\mathcal{X}_T$ is non-empty. Below we show that this issue is in fact finitely determined, that is, determined by the behavior of $T$ on finite dimensional subspaces.
6.5 When is $\mathcal{X}_T$ non-empty?

Let $A$ be a non-empty open subset of a Banach space $X$ and let $T : A \to 2^{X^\ast}$ be a $w^\ast$-cuscó defined on $A$. Then for each subspace $Y$ of $X$ with $Y \cap A \neq \emptyset$ we define $T_Y : Y \cap A \to 2^{Y^\ast}$ by: $T_Y(x) := \{ y^\ast \in Y^\ast : y^\ast = x^\ast|_Y \text{ and } x^\ast \in T(x) \}$.

**Theorem 6.5.1** Let $A$ be a non-empty open connected subset of a Banach space $X$ and let $T : A \to 2^{X^\ast}$ be a $w^\ast$-cuscó on $A$. Then $\mathcal{X}_T \neq \emptyset$ if and only if there exists an upwardly directed set $(D, \subseteq)$ of finite dimensional subspaces of $X$ such that, (i) $A \subseteq \bigcup_{Y \in D} Y$ and (ii) $\mathcal{X}_{T_Y} \neq \emptyset$ for each $Y \in D$ with $Y \cap A \neq \emptyset$.

**Proof.** For fixed $x_0 \in A$ we define $\mathcal{X}_T^0 := \{ f : f(x_0) = 0, f \in \mathcal{X}_T \}$, then $\mathcal{X}_T$ is non-empty if and only if $\mathcal{X}_T^0$ is non-empty.

Assume $f \in \mathcal{X}_T^0$. We choose $D$ to be the collection of all finite dimensional subspaces of $X$ containing $x_0$. If $Y \in D$, then $f|_{Y \cap A} \in \mathcal{X}_T^0_Y$. Indeed, for every $x \in A \cap Y$ and $y \in Y$ we have $(f|_{Y \cap A})^0(x; y) \leq f^0(x; y)$. For every $x^\ast |_Y \in \partial_c(f|_{Y \cap A})(x)$, applying the Hahn-Banach theorem to get $x^\ast \in \partial_c f(x) \subseteq T(x)$ such that $x^\ast(y) = x^\ast|_Y(y)$ for all $y \in Y$. Hence $\partial_c(f|_{Y \cap A})(x) \subseteq T_Y(x)$.

Conversely, assume $g_Y \in \mathcal{X}_T^0_Y$. For each such function we consider the following extension $\tilde{g}_Y : A \to \mathbb{R}$ defined by

$$
\tilde{g}_Y(x) := \begin{cases}
g_Y(x) & \text{if } x \in A \cap Y, \\
0 & \text{otherwise}.
\end{cases}
$$

Thus, $\{ \tilde{g}_Y : Y \in D \}$ is a net in $(\mathbb{R})^A$, which is compact by Tychonoff’s theorem [60]. It follows that $\{ \tilde{g}_Y : Y \in D \}$ has a subnet $\{ \tilde{g}_{\phi(i)} : i \in E \}$ converging to some $g$ defined on $A$, where $\phi$ is a function on $E$ with values in $D$ such that for every $Y \in D$ there exists $n \in E$ such that if $p \geq n$, then $\phi(p) \geq Y$. Since for each $i \in E$, $\tilde{g}_{\phi(i)}(x_0) = 0$, we have $g(x_0) = 0$. We first show that $g$ is real-valued. Given $y \in A$, there exists a polygonal path $P_{x_0, y} \subseteq A$ consisting of finitely many line segments and connecting $x_0$ and $y$. As $D$ is directed upwardly, by (i) there exists a finite dimensional subspace $Y_{x_0, y} \in D$ such that $P_{x_0, y} \subseteq Y_{x_0, y}$. Choose $n \in E$ such that if $p \geq n$ then $\phi(p) \geq Y_{x_0, y}$. Write $P_{x_0, y} = \bigcup_{i=1}^k \{ [x_i, x_{i+1}] \}$ for some $k$. On each $[x_i, x_{i+1}]$, being compact, $T$ is norm bounded by $M_i$, by the Lebourg mean-value theorem there exists $\xi_i \in (x_i, x_{i+1})$ such that

$$
\tilde{g}_{\phi(p)}(x_{i+1}) - \tilde{g}_{\phi(p)}(x_i) = g_{\phi(p)}(x_{i+1}) - g_{\phi(p)}(x_i) \\
\leq \sigma_{T_{\phi(p)}(\xi_i)}(x_{i+1} - x_i) = \sigma_T(\xi_i)(x_{i+1} - x_i) \leq M_i \| x_{i+1} - x_i \|.
$$
6.5 When is $\mathcal{X}_T$ non-empty?

Switching $x_i$ and $x_{i+1}$ we obtain $|\tilde{g}_{\phi(p)}(x_{i+1}) - \tilde{g}_{\phi(p)}(x_i)| \leq M_i \|x_{i+1} - x_i\|$. As $g_{\phi(p)}(x_0) = 0$ we have

$$|\tilde{g}_{\phi(p)}(y)| = |g_{\phi(p)}(y)| \leq \sum_{i=1}^k |g_{\phi(p)}(x_i) - g_{\phi(p)}(x_{i+1})| \leq \sum_{i=1}^k M_i \|x_{i+1} - x_i\|.$$ 

Hence for $p \geq n$, $\tilde{g}_{\phi(p)}(y)$ lies in a bounded interval, thus $g(y)$ is real-valued.

Next, $g$ is locally Lipschitz. Indeed, for $x \in A$, by Lemma 5.1.2 we may choose a convex open neighbourhood $U$ of $x$ such that $U \subseteq A$ and $T(U)$ is norm bounded by $M$. For $y, z \in U$, by (i) there exists a finite dimensional subspace $Y \subseteq D$ with $y, z \in Y$. Choose $n \in E$ such that if $p \geq n$ then $\phi(p) \geq Y$. On $[y, z]$ we apply the Lebour mean-value theorem to get

$$\tilde{g}_{\phi(p)}(y) - \tilde{g}_{\phi(p)}(z) = g_{\phi(p)}(y) - g_{\phi(p)}(z) \leq \sigma_{T_{\phi(p)}(\xi)}(y - z) = \sigma_{T_{\phi(p)}(\xi)}(y - z) \leq M \|y - z\|.$$ 

Switching $y$ and $z$ we get $|\tilde{g}_{\phi(p)}(y) - \tilde{g}_{\phi(p)}(z)| \leq M \|y - z\|$. Taking limit, we get $|g(y) - g(z)| \leq M \|y - z\|$, thus $g$ is Lipschitz on $U$.

Thirdly, we show $\partial_x g \subset T$ on $A$. Suppose not, there exist $x \in A, v \in X$ such that

$$g^0(x; v) > \alpha > \sigma_T(v) \quad \text{with } \|v\| = 1.$$ 

Since $T$ is a $w^*$-usco, there exists $\delta > 0$ such that

$$\sigma_T(z)(v) < \alpha \quad \text{for } z \in A \text{ satisfying } \|z - x\| < \delta. \tag{6.2}$$ 

By definition of $g^0(x; v)$ there exist $z$ and $t > 0$ with $\|z - x\| < \delta/2$ and $t < \delta/2$ such that $[g(z + tv) - g(z)]/t > \alpha$. For such $z$ and $z + tv$ there exists a finite dimensional subspace $Y$ from $D$ such that $z, z + tv \in Y_{z, z + tv}$, in particular $v \in Y_{z, z + tv}$. As $\{\tilde{g}_{\phi(i)}, i \in E\}$ converges to $g$, there exists $n \in E$ such that if $p \geq n$ then $\phi(p) \geq Y_{z, z + tv}$ and $[g_{\phi(p)}(z + tv) - g_{\phi(p)}(z)]/t > \alpha$. On such an $\phi(p)$, again applying Lebourg’s mean value theorem we obtain

$$\sigma_T(\xi)(v) = \sigma_{T_{\phi(p)}(\xi)}(v) \geq \frac{g_{\phi(p)}(z + tv) - g_{\phi(p)}(z)}{t} > \alpha,$$

where $\xi \in [z, z + tv]$. This contradicts equation (6.2). Hence $g \in \mathcal{X}_T^0$.  \hfill $\Box$

The problem of determining when $\mathcal{X}_T \neq \emptyset$ now reduces to the semi-classical problem of determining when $\mathcal{X}_{TV} \neq \emptyset$. The following theorem, which follows from Theorem 7.3.1 in Chapter 7, gives a first step in this direction.
Theorem 6.5.2  Let $A$ be a non-empty open connected subset of a finite dimensional normed linear space $X$ and let $T : A \to 2^{X^*}$ be a locally bounded $w^*$-cuso on $A$. Then $X_T \neq \emptyset$ if and only if there exists a Borel set $E \subseteq A$ with $\mu (A \setminus E) = 0$ and a $w^*$-measurable selection $\sigma : (E, \mathcal{B}_E) \to X^*$ of $T$ so that for each closed polygonal path $P$ of $A$ we have 
\[ \lim_{\varepsilon \to 0^+} \int_{P(\varepsilon)} \sigma(z) dz = 0, \] 
where $P(\varepsilon)$ is any closed $E$-admissible $\varepsilon$-path of $P$ in $A$.

6.6 The existence of $s$-minimal subdifferentials

By Zorn’s lemma, each usco [cuso] map contains a minimal usco map [minimal cusco map] [72, page 103]. What happens if we restrict the usco maps to be the Clarke subdifferential maps or the approximate subdifferential maps?

Theorem 6.6.1 ($s$-minimal approximate subdifferential) Let $A$ be a non-empty open and connected subset of a smoothable Banach space $X$ and let $f : A \to \mathbb{R}$ be locally Lipschitz on $A$. Then there exists a locally Lipschitz $g : A \to \mathbb{R}$ such that $\partial_a g \subseteq \partial_a f$ and $\partial_a g$ is $s$-minimal, i.e., for each $\hat{g}$ satisfying $\partial_a \hat{g} \subseteq \partial_a g$ we have $\partial_a \hat{g} = \partial_a g$ on $A$.

Proof. Define 
\[ P := \{ G : A \to 2^{X^*} : \text{there exists a locally Lipschitz } g \text{ on } A \text{ with } \partial_a g = G, G \subseteq \partial_a f \}. \]

Define $G_1 \preceq G_2$ if $G_2 (x) \subseteq G_1 (x)$ for each $x \in A$. Let $C$ be a chain in $P$. We will show that $C$ has an upper bound. Let $(D, \succeq)$ be the directed set obtained by considering $C$ with the induced order from $P$. For each $\alpha \in D$ we choose a locally Lipschitz $g_\alpha$ such that $g_\alpha (x_0) = f (x_0)$ and $\partial_a g_\alpha = \alpha$. Then \{ $g_\alpha, \alpha \in D$ \} is a net in $(\mathbb{R}_e)^A$, which is compact by Tychonoff’s Theorem. Thus \{ $g_\alpha, \alpha \in D$ \} has a subnet, \{ $g_{\phi(i)}, i \in E$ \}, converging pointwise to some $g \in (\mathbb{R}_e)^A$, where $\phi$ is a function on $E$ with values in $D$ such that for each $\alpha \in D$ there is $n \in E$ with the property that if $p \geq n$ then $\phi (p) \geq \alpha$. Since $D$ is linearly ordered and $\phi (E)$ is cofinal in $D$, we have $\bigcap_{i \in E} \phi (i) = \bigcap_{\alpha \in D} \alpha = \bigcap_{\alpha \in C} \alpha$, and so $\partial_a g$ is an upper bound for $C$ if we show that $\partial_a g (x) \subseteq \bigcap_{i \in E} \phi (i) (x)$ for each $x \in A$.

First, $g$ is finite-valued. Since $A$ is connected, for each $y \in A$ there exists a polygonal path in $A$ connecting $x_0$ and $y$. As the polygonal path, consisting of finitely many line segments, is compact, $\partial_c f$ is norm bounded on it. Applying the Lebourg mean-value theorem on each segment, we obtain $|g_{\phi(i)} (y)| \leq M$ for each $i \in E$ for some $M > 0$, so $g (y)$ is finite.
Moreover, $g$ is locally Lipschitz. To see this, for $x \in A$, we take a convex open neighborhood $U \subseteq A$ of $x$ such that $\partial_c f(U)$ is norm bounded by $M$. For all $y, z \in U$,

$$g_{\phi(i)}(y) - g_{\phi(i)}(z) \leq \sigma_{\partial g_{\phi(i)}}(\xi)(y - z) \leq \sigma_{\partial f}(\xi)(y - z) \leq M \|y - z\|,$$

where $\xi \in [y, z]$, then $g(y) - g(z) \leq M \|y - z\|$ as $g_{\phi(i)}$ converges to $g$ pointwise. Switching $y$ and $z$, we obtain $|g(y) - g(z)| \leq M \|y - z\|$. We proceed to show that $\partial_ag \subseteq \phi(i)$ for each fixed $i \in E$. To this end, we first show that $\partial_-g(x) \subseteq \phi(i)(x)$ for each $x \in A$. When $\partial_-g(x) = \emptyset$, this inclusion holds trivially. Assume $x^* \in \partial_-g(x)$, then for every $v \in X$ we have $\langle x^*, v \rangle \leq g^{-}(x, v)$. Fix a convex $w^*$-neighborhood $W$ of 0, and choose $\varepsilon > 0$ and a finite dimensional subspace $Y \subseteq X$ such that $4\varepsilon B_{X^*} + Y^\perp \subseteq W$. As $Y$ is finite dimensional, we have

$$\liminf_{y \in Y} \frac{g(y) - g(x) - \langle x^*, y - x \rangle}{\|y - x\|} \geq 0.$$

For the given $\varepsilon$, there exists an $\delta > 0$ such that when $\|y - x\| < \delta, y \in Y$ we have

$$g(y) - g(x) - \langle x^*, y - x \rangle \geq -\varepsilon \|y - x\| \geq -\varepsilon \cdot \delta,$$

and when $\|z - x\| < \delta$ we have $\phi(i)(z) \subseteq \phi(i)(x) + W$. As $\{g_{\phi(k)}, k \in E\}$ converges to $g$ pointwise on $A$ and $\{g_{\phi(k)} : k \in E\}$ are locally equi-Lipschitz, $\{g_{\phi(k)}, k \in E\}$ converges to $g$ uniformly on each compact subset of $A$, in particular on the compact subset $x + \delta B_Y$. Choose $n \in E$ such that if $p \geq n$ then $\phi(p) \geq \phi(i)$ and $\sup\{|g_{\phi(p)}(y) - g(y)| : y \in x + \delta B_Y\} < \frac{\delta}{2}$. Then for $y \in x + \delta B_Y$ we have

$$g_{\phi(p)}(y) - g_{\phi(p)}(x) - \langle x^*, y - x \rangle =$$

$$g(y) - g(x) - \langle x^*, y - x \rangle + (g_{\phi(p)}(y) - g(y)) - (g_{\phi(p)}(x) - g(x)) \geq -2\varepsilon \cdot \delta.$$

Applying Proposition 2.7.5, we have $z \in B_{\delta}(x)$ and $z^* \in \partial_- (g_{\phi(p)} - x^*)(z) \cap (4\varepsilon B_{X^*} + Y^\perp)$. Then,

$$x^* \in \partial_- g_{\phi(p)}(z) - z^* \subseteq \partial_a g_{\phi(p)}(z) + 4\varepsilon B_{X^*} + Y^\perp \subseteq \phi(i)(z) + W \subseteq \phi(i)(x) + 2W.$$

As $W$ is arbitrary, $x^* \in \bigcap_{\partial_a} (\phi(i)(x) + 2W) = \mathcal{w}^*\text{cl } \phi(i)(x) = \phi(i)(x)$. Hence $\partial_-g(x) \subseteq \phi(i)(x)$. Since $\partial_ag = \text{USC}(\partial_-g)$, we obtain $\partial_ag \subseteq \phi(i)$. Therefore, $\partial_ag$ is an upper bound of $C$. By Zorn's lemma, there exists a maximal element $G$ of $P$. If for some locally Lipschitz function $\tilde{g}$ we have $\partial_ag \subseteq G$, then $\partial_ag = G$ by the maximality of $G$. \qed

One may similarly prove the following:
Theorem 6.6.2 (s-minimal Clarke subdifferential) Assume that $A$ is a non-empty open and connected subset of Banach space $X$ and that $f : A \to \mathbb{R}$ is locally Lipschitz. Then there exists a locally Lipschitz $g : A \to \mathbb{R}$ such that $\partial_c g \subseteq \partial_c f$ and $\partial_c g$ is s-minimal, i.e., for each $\hat{g}$ satisfying $\partial_c \hat{g} \subseteq \partial_c g$ we have $\partial_c \hat{g} = \partial_c g$ on $A$.

Note that an approximate (Clarke) subdifferential mapping may contain several s-minimal approximate (Clarke) subdifferential mappings. On the real line, Theorem 6.6.2 is clear since every minimal cuuco on the line is a Clarke subdifferential map of some locally Lipschitz function (see page 96). On $\mathbb{R}$, since the Clarke subdifferential determines the approximate subdifferential by Theorem 2.2 [19], $\partial_c g$ is a minimal cuuco if and only if $\partial_c g$ is Clarke s-minimal, and if and only if $\partial_a g$ is approximate s-minimal. However, we observe that:

Example 6.6.3 A distance function $d_C$ on $\mathbb{R}^2$ such that (i) $\partial_c d_C$ is s-minimal but not a minimal cuuco; (ii) $\partial_a d_C$ is s-minimal but not a minimal usco.

Let $\mathbb{R}^2$ be endowed with $l^1$ norm and $f$ be 1-Lipschitz on $\mathbb{R}$ with $\partial_a f = \partial_c f \equiv [0,1]$. Define $C = \{(x, y) : y \geq f(x)\}$. Then the associated distance function is $d_C(x, y) = f(x) - y$ if $(x, y) \notin C$, otherwise 0. This follows from the following computation:

$$|f(x) - y| \leq |y - f(t)| + |f(t) - f(x)| \leq |y - f(t)| + |x - t| \quad \text{for all } t.$$ 

Let $g$ stands for $a$ or $c$. Suppose $g$ is a real-valued locally Lipschitz function on $\mathbb{R}^2$ such that $\partial_g g \subseteq \partial_c d_C$. On int $C$, as $\partial_c d_C(x, y) = \{0\}$ we have $g \equiv c_1$ on $C$ for some constant $c_1$. When $(x, y) \notin C$ we have $\partial_c d_C(x, y) = [0,1] \times \{-1\}$. Then for $(x, y) \notin C$ we have $\partial_c g(x, y) + y = [0,1] \times \{0\}$ and so $g(x, y) + y = h(x)$. By continuity we have $h(x) = g(x, f(x)) + f(x) = c_1 + f(x)$. Hence $g(x, y) = f(x) - y + c_1 = d_C(x, y) + c_1$ for $(x, y) \notin C$. Then $g = d_C + c_1$ on $\mathbb{R}^2$, in particular, $d_C$ is Clarke and approximate s-minimal. On the other hand, $\partial_a d_C = [0,1] \times \{-1\}$ on $\mathbb{R}^2 \setminus C$, which is neither a minimal cuuco or usco.

6.7 The Mordukhovich-Shao sequential subdifferential

The nonconvex subdifferential we considered so far is introduced by Ioffe on the basis of topological limits of Dini-Hadamard subdifferentials and $\varepsilon$-subdifferentials. Another kind of non-convex subdifferential construction is coined by Mordukhovich and Shao [70] on the
basis of sequential limits of Fréchet $\epsilon$-normals and subdifferentials. In finite dimensions, they are the same, but in infinite dimensional spaces they may be different (see Example 6.7.5).

Assume $X$ is an Asplund space. Since every Gâteaux differentiability space has a $w^*$-sequentially compact dual ball, $X$ has a $w^*$-sequentially compact dual ball [38, page 230]. Let $f : X \to (-\infty, +\infty]$ be lower semicontinuous and $x \in \text{dom } f$. The Mordukhovich and Shao's sequential subdifferential is the sequential Fréchet subderivative:

$$\partial_{ms}f(x) := \{w^* - \lim_{n \to \infty} x^*_n : x^*_n \in \partial^F f(x_n), \|x_n - x\| < 1/n, |f(x_n) - f(x)| < 1/n\},$$

whereas $\partial_a f(x)$ may be defined by taking the topological upper limit:

$$\partial_a f(x) = \bigcap_{k=1}^{\infty} w^* \text{cl} \left\{ \partial_{1/k}^{-} f(y) : \|y - x\| < 1/k, |f(y) - f(x)| < 1/k \right\}.$$

When $f$ is locally Lipschitz at $x$, $\partial_{ms} f(x)$ is non-empty and bounded in the norm topology. However, $\partial_{ms} f$ may not be $w^*$-closed.

**Example 6.7.1** Outside Asplund spaces $\limsup_{y \to x} \partial^F f(y)$ may be empty. If a Banach space $X$ is not an Asplund space, then there exists an equivalent norm $\|\cdot\|$ on $X$ which is nowhere Fréchet differentiable [72]. For a continuous concave function $f$ there is $x^* \in \partial_F f(x)$ if and only if $x^* = \nabla_F f(x)$. If $f(x) := -\|x\|$, then $f$ is nowhere Fréchet differentiable.

The main relationships between $\partial_a f$ and $\partial_{ms} f$ are:

**Proposition 6.7.2** [70] Let $X$ be an Asplund space. If $f$ is lower semi-continuous around $x$ then $\partial_{ms} f(x) \subset \partial_a f(x)$; if $f$ is Lipschitz continuous around $x$, then $w^* \text{cl} \partial_{ms} f(x) = \partial_a f(x)$.

If, in addition, $X$ is a WCG space then $\partial_{ms} f(x)$ is $w^*$-closed and $\partial_{ms} f(x) = \partial_a f(x)$.

**Proof.** As $\partial^F f(y) \subset \partial_{1/\epsilon}^{-} f(y)$ for every $y$ and $\epsilon > 0$, we have $\partial_{ms} f(x) \subset \partial_a f(x)$.

Assume $f$ is Lipschitz continuous around $x$. Based on the definition, we have

$$\partial_a f(x) = \bigcap_{k=1}^{\infty} w^* \text{cl} S_k \quad \text{where} \quad S_k := \bigcup \{\partial_{1/k}^{-} f(y) : \|y - x\| \leq 1/k \}.$$  

Obviously $S_{k+1} \subset S_k$ for all $k = 1, 2, \ldots$. Moreover, all the sets $S_k$ are bounded in $X^*$ due to the Lipschitz continuity of $f$ around $x$. By Proposition 2.4.4, we conclude that $\partial_a f(x) = w^* \text{cl} \partial_a^o f(x)$ where $\partial_a^o f(x) := \{\lim_{k \to \infty} x^*_k : x^*_k \in S_k \text{ for all } k\}$. We show that
\( \partial^* f(x) \subset \text{w}^* \text{cl} \partial_{ms} f(x) \). Let \( V^* \) be any \( \text{w}^* \)-closed neighbourhood of 0 in \( X^* \). For each \( V^* \) under consideration there exists a finite dimensional subspace \( L \subset X \) and \( r > 0 \) such that \( L^+ + 3r B_{X^*} \subset V^* \). Let \( x^* \in \partial^*_a f(x) \), i.e., there exists \( x_k \to x \), \( \epsilon_k \downarrow 0 \), and \( x_k^* \xrightarrow{w^*} x^* \) with \( x_k^* \in \partial_{\epsilon_k} f(x_k) \) for all \( k \). We may take \( k \) big enough to get \( 0 < \epsilon_k \leq r \) and \( 1/k \leq r \). By the definition of the Dini \( \epsilon \)-subdifferential, one conclude that for any \( r > 0 \) and finite dimensional subspace \( L \subset X \) the function

\[
  f_k(y) := f(y) - \langle x_k^*, y - x_k \rangle + 2r \| y - x_k \| + I_L(y - x_k),
\]

attains a local minimum at \( x_k \) for each \( k \). Therefore, \( 0 \in \partial^F f_k(x_k) \). As \( X \) is Asplund we can apply the Fuzzy Sum Rule to get \( \| x_k - x_k^* \| \leq 1/k \) and

\[
  x_k^* \in \partial^F f(x_k) + L^+ + 3r B_{X^*} \subset \partial^F f(x_k) + V^* \text{ for all } k.
\]

That is, \( x_k^* = x_k^* + v_k^* \) for \( x_k^* \in \partial^F f(x_k) \) and \( v_k^* \in V^* \). Since \( f \) is Lipschitz continuous around \( x \), one has the uniform boundedness of \( \partial^F f(y) \) around \( x \), thus there exists a subsequence \( x_k^* \) converges in \( \text{w}^* \)-topology to \( x^* \in \partial_{ms} f(x) \). Then \( x^* - x_k^* \in V^* \). This shows \( \partial^*_a f(x) \subset \partial_{ms} f(x) + V^* \). Taking the intersection over \( V^* \) shows \( \partial^*_a f(x) \subset \text{w}^* \text{cl} \partial_{ms} f(x) \). Finally, when \( X \) is WCG, \( \partial_{ms} f(x) \) is \( \text{w}^* \)-closed by Proposition 2.4.4, thus \( \partial_a f(x) = \partial_{ms} f(x) \).

Every WCG Asplund space admits an equivalent Fréchet differentiable renorm [37, page 286]. In WCG Asplund spaces \( \partial_{ms} f \) is in fact the \( \text{w}^* \)-sequential limits of \( F \)-viscosity subderivatives (see page 9). Separable Banach spaces with separable duals and reflexive Banach spaces are WCG Asplund spaces.

**Theorem 6.7.3** Let \( X \) be a WCG Asplund space with a Fréchet differentiable norm, then in \( (X_{B_X^*}, \rho) \) the set \( \{ g \in X_{B_{X^*}} : \partial_{ms} g = \partial_a g = \partial_{r} g \equiv B_{X^*} \} \) is residual.

**Proof.** Combine Theorem 6.3.7 and Proposition 6.7.2.

**Theorem 6.7.4** Let \( A \) be a non-empty open subset of a WCG Asplund space \( X \). Suppose \( C \subset X^* \) is \( \text{w}^* \)-compact convex and \( \text{w}^* \)-separable, then in \( (X_C, \rho) \) the set

\[
  \{ g \in X_C : \partial_{ms} g = \partial_a g = \partial_{r} g \equiv C \text{ on } A \}
\]

is residual.

**Proof.** Apply Corollary 6.2.2 and Proposition 6.7.2.

The following example due to Borwein and Fitzpatrick [14] illustrates the need of WCG hypothesis.
Example 6.7.5 A continuous concave function $f$ on the continuous function space $C(\Omega)$ (non-separable) such that $\partial_a f(x) \neq \partial^F_a f(x) = D^F_a f(x) \neq \partial_{ms} f(x) = D^F_{ms} f(x)$ for some $x \in C(\Omega)$, where $\Omega$ denotes the set of ordinals $\leq \omega_1$, the first uncountable ordinal.

First, $\Omega$ is compact with the order topology. Sequences fail to describe the order topology on $\Omega$. Now $C(\Omega)$ has an equivalent Fréchet differentiable renorm [37, page 313], thus is Asplund. Its dual is the space of all regular, finite Borel measures $\mu$ on the Borel sets of $\Omega$, $\text{rca}(\Omega)$. Define $f(x) := -\|x\|_{\infty}$ for $x \in C(\Omega)$. Let $\mu_\omega$ be the point mass at $\omega \in \Omega$, that is, the continuous linear functional defined by $(\mu_\omega, x) = x(\omega)$. The norm $\| \cdot \|_{\infty}$ of $C(\Omega)$ is Fréchet differentiable at $x \in C(\Omega)$ if and only if there is $\omega$, an isolated point of $\Omega$, such that $|x(\omega)| > |x(t)|$ for $t \neq \omega$ [37]. In that case the derivative is $\mu_\omega$. By considering for $\omega$ any non-limit ordinal

$$y_\omega(t) := \begin{cases} 
1 + \varepsilon & \text{if } t = \omega, \\
1 & \text{otherwise},
\end{cases}$$

at $x = 1$, by Proposition 6.7.2 we get

$$\partial_a f(x) = D^F_a f(x) = \{ -\mu_\omega : \omega \in \Omega \} \neq \partial_{ms} f(x) = D^F_{ms} f(x) = \{ -\mu_\omega : \omega < \omega_1 \},$$

because $\omega_1$ is not the limit of a sequence of countable ordinals, while $\omega < \omega_1$ is a limit of sequences of non-limit ordinals. On the other hand, at $x = 1$ we have

$$\partial_c f(x) = \mathcal{w}D^F_{\omega} f(x) = \{ \mu \in \text{rca}(\Omega) : \mu \geq 0, \mu(\Omega) = 1 \}$$

and

$$\partial_c f(x) \neq \partial_a f(x).$$

Problem 6.7.6 Is it possible to remove the minimality assumption in Theorems 6.2.1 and 6.3.1?
Chapter 7

Green’s theorem in Banach spaces

In the final chapter we prove an infinite dimensional calculus theorem by which we generalize one of Borwein, Moors and Shao’s results. In some sense, our calculus theorem provides an explicit way to recover an essentially smooth locally Lipschitz function from any selection of its Clarke subdifferential map. In order to accomplish this, we need to consider line integrals in general Banach spaces.

7.1 Line integrals

Assume that $A$ is a non-empty open connected subset of a Banach space $X$ and that $T : A \to 2^{X^*}$ is a maximal cyclically monotone mapping. In order to recover a convex function $f$ having $T$ as its subdifferential, Rockafellar did as follows: Given $(x_0, x_0^*) \in \text{Gr}(T)$, for $x \in X$ one defines

$$f(x) := \sup \{\langle x - x_m, x_m^* \rangle + \cdots + \langle x_1 - x_0, x_0^* \rangle : (x_i, x_i^*) \in \text{Gr}(T), 1 \leq i \leq m, m \in \mathbb{N}\}.$$ 

Then $f$ is a proper lower semicontinuous convex function on $X$ and $\partial f = T$ [72]. Examining the definition of $f$ carefully, it reminds us of the ‘Riemann integration’ on the polygonal curve $[x_0, x_1], [x_1, x_2], \ldots, [x_{m-1}, x_m], [x_m, x]$. This motivates the following definition:

**Definition 7.1.1** Let $(\mathcal{M}, \sigma)$ be a measurable space and let $X$ be a Banach space. We say that a function $f : (\mathcal{M}, \sigma) \to X^*$ is weak$^*$ measurable if $f^{-1}(U) \in \sigma$ for each weak$^*$ open subset $U$ of $X^*$. Note: if $X$ is a separable normed linear space then this is equivalent to
demanding that for each \( x \in X \), the mapping \( \hat{x} \circ f : A \to \mathbb{R} \) defined by \( (\hat{x} \circ f)(t) := f(t)(x) \) is measurable.

For a non-empty open subset \( A \) of a Banach space \( X \), the line integral along a line segment \( [a, b] \subseteq A \) of a weak* measurable function \( f : (A, B_A) \to X^* \) is the Lebesgue integral:

\[
\int_{[a, b]} f(z) \, dz := \int_0^1 \langle f(tb + (1-t)a), b-a \rangle \, dt.
\]

A polygonal path \( P \) in \( A \) is an ordered collection of line segments \( \{[a_i, a_{i+1}] : 1 \leq i \leq n-1 \} \) for some integer \( n \). If \( a_1 = a_n \), we call \( P \) a closed polygonal path. Moreover, we write \( -P := \{[a_{n-i+1}, a_{n-i}] : 1 \leq i \leq n-1 \} \). The line integral of \( f \) on \( P \) is defined as:

\[
\int_P f(z) \, dz := \sum_{i=1}^{n-1} \int_{[a_i, a_{i+1}]} f(z) \, dz.
\]  

(7.1)

Each open connected set is polygonally path connected [1, page 180]. Given a polygonal path \( P := \{[a_i, a_{i+1}] : 1 \leq i \leq n-1 \} \) in \( A \), for any \( \varepsilon > 0 \) we will call a collection of line segments \( P(\varepsilon) \) an \( \varepsilon \)-path of \( P \) provided that \( P(\varepsilon) \) may be written as: \( P(\varepsilon) = \{[a_i', b_i'] : 1 \leq i \leq n-1 \} \) and

\[
\sum_{i=1}^{n-1} \|a_i - a_i'\| + \sum_{i=1}^{n-1} \|a_{i+1} - b_i'\| < \varepsilon.
\]

Such an \( \varepsilon \)-path is closed if \( a_1 = a_n \). For a Borel subset \( E \) of \( A \) we say that \( P(\varepsilon) \) is an \( E \)-admissible \( \varepsilon \)-path of a polygonal path \( P \) if \( P(\varepsilon) \) is an \( \varepsilon \)-path of \( P \) and

\[
\mu(\{t \in [0, 1] : tb_i' + (1-t)a_i' \notin E \}) = 0 \text{ for } 1 \leq i \leq n-1.
\]

Line integrals on an \( \varepsilon \)-path are defined similarly as (7.1).

**Proposition 7.1.2** Let \( A \) be a non-empty open connected subset of a Banach space \( X \) and \( B \subset A \) be a Borel set with \( A \setminus B \) Haar null. Suppose \( \sigma : B \to X^* \) is w*-measurable and locally bounded such that for each closed polygonal path \( P \) of \( A \) we have \( \lim_{\varepsilon \downarrow 0} \int_{P(\varepsilon)} \sigma(z) \, dz = 0 \) (where \( P(\varepsilon) \) denotes any closed \( B \)-admissible \( \varepsilon \)-path of \( P \)), then for any fixed \( a \in A \), the function \( f : A \to \mathbb{R} \) defined by \( f(x) := \lim_{\varepsilon \downarrow 0} \int_{P(\varepsilon)} \sigma(z) \, dz \) (where here \( P(\varepsilon) \) is any \( B \)-admissible \( \varepsilon \)-path of a polygonal path from \( a \) to \( x \) in \( A \)) is locally Lipschitz. When \( [x_1, x_2] \subset A \) is \( B \)-admissible, we have \( f(x_2) - f(x_1) = \int_{[x_1, x_2]} \sigma(z) \, dz \).
7.2 Line integrals

Proof. Let $P$ be an arbitrary polygonal path from $a$ to $x$ in $A$. Then $\int_{P(\varepsilon)} \sigma(z)dz$ is finite as $\sigma$ is locally bounded and $P(\varepsilon)$ is compact. For any $\varepsilon_1, \varepsilon_2 > 0$, let $P_i(\varepsilon_i)$ be $B$-admissible $\varepsilon_i$-paths of $P$, $i = 1, 2$. Then $P_1(\varepsilon_1) \cup (-P_2(\varepsilon_2))$ is a closed $B$-admissible $\varepsilon_1 + \varepsilon_2$-path of $P \cup -P$. By assumption we deduce

$$|\int_{P_1(\varepsilon_1)} \sigma(z)dz - \int_{P_2(\varepsilon_2)} \sigma(z)dz| = |\int_{P_1(\varepsilon_1) \cup (-P_2(\varepsilon_2))} \sigma(z)dz| \to 0,$$

as $\varepsilon_1, \varepsilon_2 \to 0$, therefore $f$ is well-defined for $P$. Moreover, $f$ does not rely on $P$, that is, for arbitrary two polygonal paths $P_1, P_2$ from $a$ to $x$ in $A$, we have $\lim_{\varepsilon \downarrow 0} \int_{P(\varepsilon)} \sigma(z)dz = \lim_{\varepsilon \downarrow 0} \int_{P(\varepsilon)} \sigma(z)dz$. Indeed, as $-P_1 \cup P_2$ is a closed polygonal path, by assumption we have

$$- \lim_{\varepsilon \downarrow 0} \int_{P_1(\varepsilon)} \sigma(z)dz + \lim_{\varepsilon \downarrow 0} \int_{P_2(\varepsilon)} \sigma(z)dz = \lim_{\varepsilon \downarrow 0} \int_{P_1(\varepsilon)} \sigma(z)dz + \lim_{\varepsilon \downarrow 0} \int_{P_2(\varepsilon)} \sigma(z)dz = \lim_{\varepsilon \downarrow 0} \int_{(-P_1(\varepsilon)) \cup P_2(\varepsilon)} \sigma(z)dz = 0.$$

Next, we verify that $f$ is locally Lipschitz. Let $x_0 \in A$ and $U$ be convex neighborhood of $x_0$ contained in $A$ such that $\sigma(B \cap U) \subseteq LB_\varepsilon$ for some $L > 0$. Consider any two points $x, y \in U$ and fix $\delta > 0$. By the definition of $f$ we may choose $\delta > \varepsilon_0 > 0$ sufficiently small such that if $P(\varepsilon)$ is any $B$-admissible $\varepsilon$-path ($0 < \varepsilon < \varepsilon_0$) from $a$ to $x$ and $P'(\varepsilon)$ is any $B$-admissible $\varepsilon$-path from $a$ to $y$, then one has

$$|f(x) - \int_{P(\varepsilon)} \sigma(z)dz| < \delta, \quad |f(y) - \int_{P'(\varepsilon)} \sigma(z)dz| < \delta.$$

Now suppose $P(\varepsilon) = \{[a_i, b_i] : 1 \leq i \leq n\}$ is a $B$-admissible $\varepsilon$-path from $a$ to $x$ with $0 < \varepsilon < \varepsilon_0$. We may extend $P(\varepsilon)$ to a $B$-admissible $\varepsilon$-path $P'(\varepsilon)$ from $a$ to $y$, where $P'(\varepsilon) = \{[a_i, b_i] : 1 \leq i \leq n + 1\}$. It follows that

$$|f(x) - f(y)|
\leq |f(y) - \int_{P'(\varepsilon)} \sigma(z)dz| + |\int_{P'(\varepsilon)} \sigma(z)dz - \int_{P(\varepsilon)} \sigma(z)dz| + |\int_{P(\varepsilon)} \sigma(z)dz - f(x)|
\leq \delta + |\int_{[a_{n+1}, b_{n+1}]} \sigma(z)dz| + \delta \leq L\|a_{n+1} - b_{n+1}\| + 2\delta
\leq L(\|a_{n+1} - x\| + \|x - y\| + \|b_{n+1} - y\|) + 2\delta \leq L\|x - y\| + (2L + 2)\delta.$$

As $\delta$ is arbitrary, this gives $|f(x) - f(y)| \leq L\|x - y\|$ as required.

Finally, if $[x_1, x_2]$ is $B$-admissible and $P(\varepsilon)$ is an $\varepsilon$-path of a polygonal path $P$ connecting $a$ and $x_1$, then $P(\varepsilon) \cup [x_1, x_2]$ is an $\varepsilon$-path from $a$ to $x_2$. Hence $f(x_2) = \lim_{\varepsilon \downarrow 0} \int_{P(\varepsilon) \cup [x_1, x_2]} \sigma(z)dz = \lim_{\varepsilon \downarrow 0} \int_{P(\varepsilon)} \sigma(z)dz + \int_{[x_1, x_2]} \sigma(z)dz = f(x_1) + \int_{[x_1, x_2]} \sigma(z)dz$. \qed
7.2 Integration of subdifferentials of essentially smooth functions

When \( f \) is essentially smooth on \( A \), with line integral we may use any selection \( \sigma \) from \( \partial_c f \) to recover \( f \). To see this, let \( a \in A \) be fixed. Let \( x_1 \) be a point in \( A \) such that \([a,x_1] \subset A\). Let \( a' \rightarrow a \), \( x_1' = a' + (x_1 - a) \). According to Theorem 2.5.4, for almost every \( a' \) near \( a \), \([-f^0(\cdot,-(x_1-a)),f^0(\cdot,x_1-a)]\) is single-valued almost everywhere on the line segment \([a',x_1']\) in the one-dimensional measure. On such a line segment \([a',x_1']\),

\[
-f^0(\cdot,-(x_1-a)) = \sigma(\cdot)(x_1 - a) = f'(\cdot, x_1 - a) = f^0(\cdot, x_1 - a) \text{ almost everywhere, and so }
\]

\[
f(x_1') - f(a') = \int_0^1 \langle \sigma(a' + t(x_1 - a)), x_1 - a \rangle dt.
\]

However \( f(x_1') \rightarrow f(x_1) \), \( f(a') \rightarrow f(a) \) as \( a' \rightarrow a \). Thus

\[
f(x_1) = f(a) + \lim_{a' \rightarrow a} \int_0^1 \langle \sigma(a' + t(x_1 - a)), x_1 - a \rangle dt: \sigma(\cdot)(x_1 - a) = f^0(\cdot, x_1 - a) = -f^0(\cdot, -(x_1 - a)) \text{ almost everywhere on } [a',x_1'] \}. \quad (7.2)
\]

Define \( E_1 := \{ x \in A : f^0(x; x_1-a) = -f^0(x; -(x_1-a)) \} \). The latter is exactly:

\[
\lim_{\varepsilon \downarrow 0} \int_{P_1(\varepsilon)} \sigma(z)dz \text{ where } P_1(\varepsilon) \text{ is any } E_1\text{-admissible } \varepsilon\text{-path of } [a,x_1] \text{ in } A. \text{ Then }
\]

\[
f(x_1) = f(a) + \lim_{\varepsilon \downarrow 0} \int_{P_1(\varepsilon)} \sigma(z)dz.
\]

For an arbitrary \( x \in A \), we may choose a polygonal path \( P \) in \( A \) to connect \( a \) and \( x \), say \([x_0,x_1],[x_1,x_2],\ldots,[x_{k-1},x_k]\) with \( x_0 = a \) and \( x_k = x \). Define

\[
E_i := \{ x \in A : f^0(x; x_i-x_{i-1}) = -f^0(x; -(x_i-x_{i-1})) \} \text{ for } 1, \ldots, k.
\]

We have

\[
f(x) = f(a) + \sum_{i=1}^{k} \lim_{\varepsilon \downarrow 0} \int_{P_i(\varepsilon)} \sigma(z)dz = f(a) + \lim_{\varepsilon \downarrow 0} \int_{P(\varepsilon)} \sigma(z)dz,
\]

where \( P_i(\varepsilon) \) are arbitrary \( E_i \)-admissible \( \varepsilon \)-paths of \([x_{i-1},x_i]\) in \( A \).

7.3 Green's theorem and almost everywhere differentiability

We first show the finite dimensional analogue of fundamental theorem of calculus.
**Theorem 7.3.1** Assume $A \subset \mathbb{R}^n$ is open connected and $B \subset A$ is a Borel set with $A \setminus B$ Lebesgue null. Suppose there exists a locally bounded and Borel measurable $\sigma : B \to \mathbb{R}^n$ such that for each closed polygonal path $P$ of $A$ we have $\lim_{\varepsilon \to 0} \int_{P(\varepsilon)} \sigma(x)dx = 0$ (where $P(\varepsilon)$ is any closed $B$-admissible $\varepsilon$-path of $P$ in $A$), then there exists a locally Lipschitz function $f$ on $A$ such that its Fréchet gradient equals $\sigma$ almost everywhere.

**Proof.** For $a \in A$, we define $f(x) := \lim_{\varepsilon \to 0} \int_{P(\varepsilon)} \sigma(z)dz$ for all $x \in A$ where $P(\varepsilon)$ is a $B$-admissible $\varepsilon$-path of some polygonal path $P$ from $a$ to $x$ in $A$. By Proposition 7.1.2, $f$ is locally Lipschitz.

(1) Since $f$ is continuous, 

$$
\overline{D_v}f(x) = \limsup_{t \to 0} \frac{f(x + tv) - f(x)}{t} = \lim_{k \to \infty} \sup_{0 <|t| < 1/k} \frac{f(x + tv) - f(x)}{t},
$$

is Borel measurable, as is $\underline{D_v}f(x) = \liminf_{t \to 0} (f(x + tv) - f(x))t^{-1}$. Since $B$ is Borel measurable,

$$
\overline{A}_v := \{x \in B : \overline{D_v}f(x) = \langle \sigma(x), v \rangle\} \quad \text{and} \quad \underline{A}_v := \{x \in B : \underline{D_v}f(x) = \langle \sigma(x), v \rangle\},
$$

are Borel measurable. Fix $v \in X$ we show $A \setminus \overline{A}_v$ is null. For any closed hyperplane $H$ of $\mathbb{R}^n$ not containing $v$, we consider the isomorphism $T : H \times \mathbb{R} \to X$ defined by $T(h, r) := h + rv$. If we set $N_H := \{h : \mu(\{r \in H : T(h, r) \in A \setminus B\}) > 0\}$ then $N_H$ is a null subset of $H$ by Fubini’s theorem. For $x = h + rv \in A$ with $h \in H \setminus N_H$, by Proposition 7.1.2 we may write

$$
f(x + tv) - f(x) = \int_0^1 \langle \sigma(x + stv), tv \rangle ds = \int_0^1 \langle \sigma(x + sv), v \rangle ds.
$$

Then $\phi(t) := f(x + tv) + \sigma(y, v)$ is differentiable and $\overline{D_v}f(y) = \langle \sigma(y), v \rangle$ a.e on $\{t : x + tv \in A\}$. Hence $\mu((A \setminus \overline{A}_v) \cap L) = 0$ for each line $L$ parallel to $v$ passing $x = h + rv \in A$ with $h \in H \setminus N_H$. Fubini’s theorem implies $A \setminus \overline{A}_v$ is null. Then $A \setminus (\overline{A}_v \cap \underline{A}_v)$ is Lebesgue null, and $D_v f(x) = \langle \sigma(x), v \rangle$ on $\overline{A}_v \cap \underline{A}_v$.

(2) Now choose $\{v_k\}$ to be countable, dense subset of $B_{R^n}$. Set 

$$
A_k := \{x \in B : D_{v_k}f(x) \text{ exists and } D_{v_k}f(x) = \langle \sigma(x), v_k \rangle\},
$$

for $k = 1, 2, \ldots$, and define $D := \bigcap_{k=1}^\infty A_k$. Then $A \setminus D$ is Lebesgue null.
7.3 Green's theorem and almost everywhere differentiability

(3) We claim \( f \) is differentiable at each \( x \in D \) and \( \nabla f(x) = \sigma(x) \). Fix \( x \in D \). Choose \( v \in S_{R^n}, \ t \in \mathbb{R}, \ t \neq 0 \), and write

\[
Q(x, v, t) := \frac{f(x + tv) - f(x)}{t} - \langle \sigma(x), v \rangle.
\]

Then if \( v' \in S_{R^n} \), we have

\[
|Q(x, v, t) - Q(x, v', t)| \leq |\frac{f(x + tv) - f(x + tv')}{t}| + |\langle \sigma(x), v - v' \rangle| \leq 2L \|v - v'\|.
\]

Now fix \( \varepsilon > 0 \). By compactness we may choose \( N \) so large that if \( v \in S_{R^n} \), then \( \|v - v_k\| \leq \varepsilon/(4L) \) for some \( k \in \{1, \ldots, N\} \). Now \( \lim_{t \to 0} Q(x, v_k, t) = 0 \) for \( k = 1, \ldots, N \), thus there exists \( \delta > 0 \) so that \( |Q(x, v_k, t)| < \varepsilon/2 \) for all \( 0 < \|t\| < \delta, \ k = 1, \ldots, N \). Consequently, for each \( v \in S_{R^n} \), there exists \( k \) such that \( |Q(x, v, t)| \leq |Q(x, v_k, t)| + |Q(x, v, t) - Q(x, v_k, t)| < \varepsilon \) if \( 0 < \|t\| < \delta \). Note the same \( \delta \) works for all \( v \in S_{R^n} \). \( \square \)

We now generalize Theorem 7.3.1 to separable Banach spaces.

**Lemma 7.3.2** Let \( Y \) be a finite dimensional subspace of \( X \). If \( N \subseteq X \) is Borel measurable, and \( (N + x) \cap Y \) has Lebesgue measure zero in \( Y \) for all \( x \in X \), then \( A \) is Haar null.

**Proof.** Let \( \mu \) be the Lebesgue measure \( \lambda \) on \( Y \) restricted to the unit cube of \( Y \). Then \( \mu \) is a test measure for \( N \). \( \square \)

**Theorem 7.3.3** Let \( A \) be a non-empty open connected subset of a separable Banach space \( X \). Assume that there exists a Borel set \( B \subset A \) with \( A \setminus B \) Haar-null such that \( \sigma : B \to X^* \) is locally norm bounded and Borel measurable and that for each closed polygonal path \( P \) in \( A \) we have \( \lim_{t \to 0} \int_{P(t)} \sigma(x) dx = 0 \) where \( P(\varepsilon) \) is any \( B \)-admissible \( \varepsilon \)-path of \( P \) in \( A \). Then there exists a locally Lipschitz function \( f \) on \( A \) such that the Gâteaux gradient of \( f \) equals \( \sigma \) except for a Haar-null set. In particular, \( \partial f \equiv \text{CSC}(\sigma|_{D_f}) \subset \text{CSC}(\sigma) \) where \( D_f \) denote the set of points at which \( f \) is Gâteaux differentiable.

**Proof.** By Proposition 7.1.2, the function \( f : A \to \mathbb{R} \) given by \( f(x) := \lim_{\varepsilon \to 0^+} \int_{P(\varepsilon)} \sigma(z) dz \), where \( P(\varepsilon) \) is a \( B \)-admissible \( \varepsilon \)-path in \( A \) of some polygonal path \( P \) from \( a \) to \( x \), is locally Lipschitz. Let \( X_1 \subset X_2 \subset \cdots X_n \subset \cdots \) be an increasing sequence of finite dimensional subspaces of \( X \) so that \( X = \text{cl} \bigcup_{n=1}^{\infty} X_n \). For each \( n \), let \( D_n := \{x \in B : f'(x;y) = \langle \sigma(x), y \rangle \text{ for all } y \in X_n \} \). Then \( D_n \) is Borel measurable. We claim that \( A \setminus D_n \) is Haar null.
7.3 Green's theorem and almost everywhere differentiability

By Lemma 7.3.2, it suffices to show that for each fixed \( z \in X \), \((A \setminus D_n) + z) \cap X_n\) is a set of Lebesgue measure 0 in \( X_n \). But this set is exactly:

\[
A_z := \{ x \in X_n : g_z'(x; a) \neq \langle \sigma(x + z), a \rangle \text{ for some } a \in X_n \},
\]

where \( g_z : X_n \to \mathbb{R} \) is defined by \( g_z(x) := f(z + x) \). By Theorem 7.3.1, \( A_z \) is Lebesgue null in \( X_n \). Put \( D'_f := \bigcap_{n=1}^{\infty} D_n \). Since \( A \setminus D_n \) is Haar null for each \( n \), \( A \setminus D'_f \) is Haar null by Proposition 2.5.2. If \( x \in D'_f \), \( f'(x; a) = \langle \sigma(x), a \rangle \) for \( a \in \bigcup_{n=1}^{\infty} X_n \). For \( v \in X \), given \( \varepsilon > 0 \) we may find \( v_k \in \bigcup_{n=1}^{\infty} X_n \) with \( \| v - v_k \| < \varepsilon / L \). Since \( f \) is locally Lipschitz around \( x \), for sufficiently small \( t > 0 \) we have

\[
\left| \frac{f(x + tv) - f(x)}{t} - \langle \sigma(x), v \rangle \right| \leq 2L \| v_k - v \| + \left| \frac{f(x + tv_k) - f(x)}{t} - \langle \sigma(x), v_k \rangle \right| \leq 3\varepsilon,
\]

thus \( f'(x; v) = \langle \sigma(x), v \rangle \) for every \( v \in X \). Now by [89] for any Haar null set \( N \) we have

\[
\partial f(x) = \overline{w}^* \{ \lim_{n \to \infty} \sigma(x_n) : \lim_{n \to \infty} x_n = x, x_n \in D'_f \setminus N \}.
\]

and so \( \partial_c f = \text{CSC}(\sigma|_{D'_f \setminus N}) \) since the dual ball \( B_{X^*} \) is \( w^* \)-sequentially compact. \[ \square \]

Assume that \( \Omega : A \to 2^{X^*} \) is a minimal \( w^* \)-cusco and that \( \sigma : B \to X^* \) satisfies \( \sigma(x) \in \Omega(x) \) for each \( x \in B \) and the condition of Theorem 7.3.3. Then \( \partial_c f = \Omega \), which is one of the main results in [9]. The following is a separable subspace version of Theorem 7.3.3.

**Theorem 7.3.4** Let \( Y \) be a closed separable subspace of a Banach space \( X \). Let \( f \) be the real-valued locally Lipschitz function defined in Proposition 7.1.2. Then there exists a Borel subset \( D \subset A \) such that \( A \setminus D \) is a Haar-null set and for each \( x \in D \) we have \( f'(x; y) = \langle \sigma(x), y \rangle \) for \( y \in Y \).

**Proof.** Let \( Y_1 \subset Y_2 \subset \ldots \subset Y_n \subset \ldots \) be an increasing sequence of finite dimensional subspaces of \( Y \) whose union is dense in \( Y \). Consider

\[
D_n := \{ x \in B : f'(x, y) = \langle \sigma(x), y \rangle \text{ for all } y \in Y_n \}.
\]

Then \( D_n \) is Borel measurable. Applying the same arguments as in Theorem 7.3.3, we have \( A \setminus D_n \) is Haar null. Put \( D := \bigcap_n D_n \). Then \( A \setminus D \) is Haar null. If \( x \in D \), \( f'(x; y) = \langle \sigma(x), y \rangle \) for each \( y \in Y \). \[ \square \]
Lemma 7.3.5 Let \( f : A \to \mathbb{R} \) be locally Lipschitz on an open subset \( A \) of a Banach space. Let \( \sigma : A \to X^* \) be densely defined. Then \( \partial f \subset \text{CSC} (\sigma) \) if and only if for \( x \in A \) and \( h \in X \)

\[
f^0(x; h) \leq \limsup \{ \langle \sigma(y), h \rangle : y \in \text{dom}(\sigma), y \to x \}. \tag{7.3}
\]

Proof. Fix \( \varepsilon > 0 \). If \( \partial f \subset \text{CSC} (\sigma) \), then \( \partial f(x) \subset \overline{w^*}\sigma[B_\varepsilon(x)] \). Hence for each \( h \in X \), we have \( f^0(x; h) \leq \sup \{ \langle \sigma(y), h \rangle : y \in B_\varepsilon(x) \} \) and (7.3) holds. Conversely if (7.3) holds, then for each \( \varepsilon > 0 \) we have \( f^0(x; h) \leq \sup \{ \langle \sigma(y), h \rangle : y \in B_\varepsilon(x) \} \) for all \( h \in X \), and so \( \partial f(x) \subset \overline{w^*}\sigma[B_\varepsilon(x)] \) and \( \partial f \subset \text{CSC} (\sigma) \). \( \square \)

The following generalizes Theorem 2.1 [9].

Theorem 7.3.6 Assume that \( A \) is a nonempty connected open subset of a Banach space \( X \). Suppose there exists a locally bounded and Borel measurable map \( \sigma : A \to X^* \) such that \( \int_P \sigma(z) dz = 0 \) for every closed polygonal path \( P \) in \( A \). Then there exists a locally Lipschitz function \( f \) on \( A \) with \( \partial f \subset \text{CSC} (\sigma|_{A \setminus N}) \) for any Haar null set \( N \subset A \). For every closed separable subspace \( Y \subset X \), there exists a Haar null set \( N_Y \subset X \) such that \( f^0(x; y) = \limsup_{z \to x} \langle \sigma(z), y \rangle \) for every \( y \in Y \).

Proof. For fixed \( a \in A \), the function \( f : A \to \mathbb{R} \) given by \( f(x) := \int_\Gamma \sigma(z) dz \), where \( \Gamma \) is any polygonal path in \( A \) from \( a \) to \( x \in A \), is locally Lipschitz by Proposition 7.1.2. We will show that for each \( \varepsilon > 0 \) and each \( \delta > 0 \)

\[
f^0(x; y) - \varepsilon < \sup \{ \langle \sigma(z), y \rangle : \| z - x \| \leq \delta \text{ and } z \in A \setminus N \}.
\]

Suppose \( \varepsilon > 0 \) and \( \delta > 0 \) are given. By possibly making \( \delta \) smaller, we may assume that \( B_\delta(x) \subset A \). Let \( H \) be any closed hyperplane in \( X \), that does not contain \( y \), and consider the isomorphism \( T : H \times \mathbb{R} \to X \) given by \( T(h, r) := h + ry \). If we set

\[
N_H := \{ h \in H : \mu(\{ r \in \mathbb{R} : T(h, r) \in N \}) > 0 \},
\]

then by Theorem 2.5.4 we have that \( N_H \) is a Haar null set of \( H \). By definition of \( f^0(x; y) \), there exists a point \( z = h_x + r_x y \in A \) and \( \lambda \in (0, 1) \) such that \( \| z + \lambda y - x \| < \delta, \| z - x \| < \delta \) and

\[
\frac{f(z + \lambda y) - f(z)}{\lambda} > f^0(x; y) - \varepsilon.
\]
7.4 Another characterization of the Clarke subdifferential

Since $f$ is locally Lipschitz and $H \setminus N_H$ is dense in $H$, we may assume $h_z \in H \setminus N_H$. The definition of $f$ allows us to write

$$f(z + \lambda y) - f(z) = \int_0^1 \langle \sigma(z + t\lambda y), \lambda y \rangle dt = \int_0^1 \langle \sigma(z + s y), y \rangle ds.$$  

By the Lebesgue mean-value theorem on $[z, z + \lambda y]$, there exists a set $M$ of positive measure in the interval $(z, z + \lambda y)$ where $f'(z'; y) = \langle \sigma(z'), y \rangle$ and

$$f'(z'; y) \geq \frac{f(z + \lambda y) - f(z)}{\lambda} > f^0(x, y) - \varepsilon.$$  

Since $\mu(\{r \in \mathbb{R} : T(h_z, r) \in N\}) = 0$, we may choose $z_0 \in M \setminus N$. Then $\|z_0 - x\| < \delta$ and

$$f^0(x; y) - \varepsilon < \langle \sigma(z_0), y \rangle \leq \sup\{\langle \sigma(z), y \rangle : \|z - x\| < \delta \text{ and } z \in A \setminus N\}.$$  

Letting $\varepsilon, \delta \to 0$ shows $f^0(x; y) \leq \lim \sup\{\langle \sigma(z), y \rangle : z \in A \setminus N, z \to x\}$. Lemma 7.3.5 shows $\partial f \subset \text{CSC}(\sigma|_{A \setminus N})$. By Theorem 7.3.4, $f'(x; y) = \langle \sigma(x), y \rangle$ for $x \in X \setminus N_Y$. By Theorem 3.1 in [20], $f^0(x; y) = \lim \sup_{z \to x} f^-(z; y)$ for every Haar null set $N$. Let $N = N_Y$, we have $f^0(x; y) = \lim \sup_{z \to x} \langle \sigma(z), y \rangle$ for $y \in Y$. \hfill \Box

If for every $y \in X$, the set $\{x \in A : \lim \sup_{z \to x} \langle \sigma(z), y \rangle \neq \lim \inf_{z \to x} \langle \sigma(z), y \rangle\}$ is Haar null, then $f$ is essentially smooth on $A$. In separable Banach space, this is the same as requiring $\sigma$ to be norm-to-$w^*$ continuous on $A$ except for a Haar null set.

7.4 Another characterization of the Clarke subdifferential

Let $A$ be a non-empty open connected subset of a separable Banach space $X$. A w*-cuso $T : A \to 2^X^*$ (locally bounded on $A$ by Lemma 5.1.2) is called cyclically normal on $A$ if for a Haar null set $E \subset A$,

(i) there exists a selection $\sigma$ from $T$ such that for any closed polygonal path $P$ in $A$ one has $\lim_{\varepsilon \to 0} \int_{P(\varepsilon)} \sigma(z)dz = 0$, where $P(\varepsilon)$ is any closed $(A \setminus E)$-admissible $\varepsilon$-path of $P$ in $A$;

(ii) $T = \text{CSC}(\sigma|_{A \setminus N})$ where $N$ is any Haar null set containing $E$.

Theorem 7.4.1 Let $A$ be a non-empty open connected subset of a separable Banach space $X$. Then a $w^*$-cuso $T : A \to 2^X^*$ is a Clarke subdifferential of some locally Lipschitz functions if and only if $T$ is cyclically normal.
**Proof.** Suppose \( T = \partial_c f \) for a locally Lipschitz function \( f \). As \( f \) is Gâteaux differentiable everywhere on \( A \) except for a Haar null set \( E \), we choose \( \sigma = \nabla f \) on \( A \setminus E \). By [89], CSC(\( \nabla f|_{A \setminus N} \)) = \( \partial_c f \) where \( N \) is any Haar null set containing \( E \). For each closed polygonal path \( P \) consisting of \([x_0, x_1], [x_1, x_2], \ldots, [x_{k-1}, x_k], [x_k, x_0] \) in \( A \) we have (see Equation 7.2)

\[
\lim_{\varepsilon \downarrow 0} \int_{P(\varepsilon)} \sigma(z)dz = f(x_1) - f(x_0) + f(x_2) - f(x_1) + \ldots + f(x_0) - f(x_k) = 0,
\]

where \( P(\varepsilon) \) is any \((A \setminus E)\)-admissible \( \varepsilon \)-path of \( P \). Hence \( \partial_c f \) is cyclically normal.

Conversely, if \( T \) is cyclically normal we may choose a selection \( \sigma \) from \( T \) defined on \( A \setminus E \) satisfying (i) and (ii). Define \( f : A \to \mathbb{R} \) by \( f(x) := \lim_{\varepsilon \downarrow 0} \int_{P(\varepsilon)} \sigma(z)dz \), where \( P(\varepsilon) \) is any \((A \setminus E)\)-admissible \( \varepsilon \)-path of a given polygonal path connecting \( a \) and \( x \). By Theorem 7.3.3, \( \partial_c f = \text{CSC}(\sigma|_{D_f \setminus N}) = T \) where \( N \) is any Haar null set containing \( E \). \( \square \)

One advantage of our \( \varepsilon \)-path definition over the one given by Borwein, Moors and Shao [9] is that we may only require local boundedness of \( T \) rather than global boundedness in Theorem 7.4.1. When the \( \text{w}^* \)-cuso \( T \) is single-valued almost everywhere, the selection \( \sigma \) is essentially uniquely defined as observed in [9, 77]. Theorem 7.4.1 generalizes Corollary 3.1 [9] and Theorem 10 [77] where \( T \) is assumed to be a minimal \( \text{w}^* \)-cuso. Compare Theorem 7.4.1 to Theorems 5.3.3 and 1.3.3.

### 7.5 Examples

**Example 7.5.1** The norm \( \| x \| := \sum_{n=1}^{\infty} |x_n| \) in the nonseparable Banach space \( l^1(\Gamma) \) (\( \Gamma \) uncountable) is not Gâteaux differentiable at any point. Let \( f(x) := \| x \| \). Then \( \partial f \) is a maximal cyclically monotone operator and a minimal norm-to-\( \text{w}^* \) cuso. Let \( \sigma \) be any selection of \( \partial f \). We have \( f(x) = \int_{[0,x]} \sigma(z)dz = \int_0^1 \langle \sigma(tx), x \rangle dt \). Since \( f \) is nowhere Gâteaux differentiable, we can only say \( \partial f = \text{CSC}(\sigma|_{\overline{1(\Gamma) \setminus N}}) \) where \( N \) is any Haar null set by Theorem 7.3.6. Nevertheless, for each separable Banach subspace \( Y \) of \( l^1(\Gamma) \) there exists a Haar null set \( N \) of \( l^1(\Gamma) \) such that \( \dot{f}(x; y) = \langle \sigma(x), y \rangle \) for every \( y \in Y \) when \( x \in l^1(\Gamma) \setminus N \).

Theorem 7.3.3 fails for Fréchet differentiability in infinite dimensions.

**Example 7.5.2** On separable \( l^1 \) there exists an equivalent norm \( \| \cdot \| \) which is nowhere Fréchet differentiable but which is Gâteaux differentiable at each nonzero point [72]. Again,
we let $f(x) := \|x\|$. Let $\sigma$ be any selection of $\partial f$ defined on $l^1$. Then $f(x) = \int_0^1 (\sigma(tx), x) dt$. The Gâteaux gradient of $f$ equals $\sigma$ on $l^1$ except for a Haar null set.

**Example 7.5.3** Let $c_0^+$ be the non-negative cone in the space of null sequences $c_0$ with supremum norm. Then $d_{c_0}^+$ is not Fréchet differentiable on $c_0^+$.

First $c_0^+$ is not Haar null. Observe that a closed subset $K \subset c_0$ is compact if and only if $K \subset \{ y \in c_0 : |y_n| \leq x_n, n = 1, 2, \ldots \}$ for a certain $x \in c_0$. Thus for each compact subset $K$ there exists $x \in c_0$ such that $K \subset c_0^+ - x$. If $c_0^+$ is Haar null, then $\mu(K) = 0$ for every compact set $K$, which is impossible by Ulam’s theorem.

Next $d_{c_0}^+$ is nowhere Fréchet differentiable on $c_0^+$. Let $x \in c \setminus c_0^+$ with $d_{c_0}^+$ being Gâteaux differentiable at $x$. Since $d_{c_0}^+(x) = \|x-y\|$ where $y := x^+ = x \vee 0 \in c_0^+$, we have $\|\nabla d_{c_0}^+(x)\| = 1$. If $d_{c_0}^+$ is Fréchet differentiable at some $x \in c_0^+$, then $\nabla d_{c_0}^+(x) = 0$. As $d_{c_0}^+$ is convex, by Proposition 2.8 [72] the subdifferential map $\partial d_{c_0}^+$ is norm-to-norm continuous at $x$. As $c_0$ is separable, $d_{c_0}^+$ is Gâteaux differentiable almost everywhere on $c_0 \setminus c_0^+$, we may select a sequence $x_n \notin c_0^+$ with $\nabla d_{c_0}^+(x_n)$ exists and $x_n \to x$, then $\lim_{n \to \infty} \|\nabla d_{c_0}^+(x_n)\| = \|\nabla d_{c_0}^+(x)\| = 1$, a contradiction.

Let $f := d_{c_0}^+$ and $\sigma$ be any selection of $\partial_c f$ defined on $c_0$. Then $f(x) = \int_{[0, \infty]} \sigma(z) dz = \int_0^1 (\sigma(tx), x) dt$. The Gâteaux gradient of $f$ equals $\sigma$ on $c_0$ except for a Haar null set. \hfill $\Box$

**Lemma 7.5.4** Let $C$ be a norm closed convex set in a reflexive space $X$ with a Fréchet differentiable norm. Then $d_C$ is Fréchet differentiable for every $x \in X \setminus C$.

**Proof.** Let $x \in X \setminus C$. As $X$ is reflexive, there exists $p(x) \in C$ such that $\|x-p(x)\| = d_C(x)$. To see this, consider the closed ball $B_{d_C(x)+\delta}[x]$ with $\delta > 0$. Now $B_{d_C(x)+\delta}[x] \cap C$ is a bounded closed convex set. Since $X$ is reflexive, $B_{d_C(x)+\delta}[x] \cap C$ is weakly compact. Since the norm is weakly lower semi-continuous, there exists $p(x) \in C$. Lemma 2.7.4 applies. \hfill $\Box$

**Example 7.5.5** For every finite (or locally finite) Radon measure $\mu$ on an infinite dimensional reflexive Banach space $X$, there exists a convex function $f : X \to \mathbb{R}$ such that the set of points at which $f$ is Fréchet differentiable is a null set for $\mu$, but $f$ is Fréchet differentiable everywhere except a Haar null set.

Following [74] we can find a sequence $C_i$ of compact convex sets containing the origin
such that \( \mu(X \setminus \bigcup_{i=1}^{\infty} C_i) = 0 \). By Lemma 7.5.4 each \( d_{C_i} \) is Fréchet differentiable on \( X \setminus C_i \). Define \( f(x) := \sum_{i=1}^{\infty} 2^{-i} d_{C_i}(x) \). We claim:

(1) \( f \) is Fréchet differentiable on \( X \setminus \bigcup_{i=1}^{\infty} C_i \). If \( x \in X \setminus \bigcup_{i=1}^{\infty} C_i \), then \( d_{C_i} \) is Fréchet differentiable for every \( i \). For every \( \epsilon > 0 \) we choose a large \( k \) such that

\[
\sum_{i=k}^{\infty} 2^{-i} d_{C_i}(y) - d_{C_i}(x) - \nabla d_{C_i}(x)(y-x) \leq \sum_{i=k}^{\infty} 2^{-i+1} < \epsilon/2.
\]

For this \( k \), we may choose small \( \delta > 0 \) such that for \( 0 < ||y - x|| < \delta \),

\[
\sum_{i=1}^{k-1} 2^{-i} d_{C_i}(y) - d_{C_i}(x) - \nabla d_{C_i}(x)(y-x) \leq \epsilon/2.
\]

From this, we conclude that \( \nabla_F f(x) = \sum_{i=1}^{\infty} 2^{-i} \nabla d_{C_i}(x) \).

(2) \( f \) is not Fréchet differentiable on \( \bigcup_{i=1}^{\infty} C_i \). Assume \( x \in C_k \setminus C_{k-1} \). For \( 1 \leq i \leq k-1 \), \( x \notin C_i \), Lemma 7.5.4 shows that \( d_{C_i} \) is Fréchet differentiable at \( x \), so does \( \sum_{i=1}^{k-1} d_{C_i} \) at \( x \). It suffices to show \( \sum_{i=k}^{\infty} 2^{-i} d_{C_i} \) is not Fréchet differentiable at \( x \). When \( x \in C_k \), \( d_{C_k} \) is not Fréchet differentiable. Choose a sequence \( y_m \) with \( ||y_m|| \to 0 \) and \( \liminf_{y_m \to 0} d_{C_k}(x + y_m) / ||y_m|| > 0 \). Since \( d_{C_i}(x) = 0 \) for \( i \geq k \), we have

\[
\liminf_{y_m \to 0} \sum_{i=k}^{\infty} 2^{-i} d_{C_i}(x + y_m) - d_{C_i}(x) \geq \liminf_{y_m \to 0} 2^{-k} d_{C_k}(x + y_m) / ||y_m|| > 0.
\]

Since \( \bigcup_{i=1}^{\infty} C_i \) is Haar null by Proposition 2.5.2 (iv), we may deduce that \( f \) is Fréchet differentiable almost everywhere on \( X \). \( \Box \)
Bibliography


Index

Approximate subdifferential
 sequential, 151
topological, 10
Asplund space, 15

Baire category
 first category, 17
 residual, 17
 theorem, 17

Banach space
 of class S, 15
 separable, 86
 smoothable, 7
 WCG, 21
 β-subdifferential, 8
 β-bounded, 5
 Fréchet, 5
 Gâteaux, 5
 Hadamard, 5

β-derivative
 sequential, 143
topological, 143

β-differentiable, 5
 Fréchet, 5
 Gâteaux, 5
 Hadamard, 5

C-Lipschitz extension, 121

C-Lipschitz separated spaces, 125
 Cantor function, 43
 Christensen’s theorem, 104
 Clarke’s Jacobian, 131
 Clarke’s subdifferential, 10
 Compact-open topology, 115
 Convex subdifferential, 12
 Cyclically normal cuso, 162

Dense minimizer function, 118
 of the second species, 118
 Dini-Hadamard subdifferential, 8
 superdifferential, 8

Directional derivative, 5
 Clarke’s, 11
 Dini, 31
 Dini-Hadamard type, 9
 Michel-Penot, 11
 Rockafellar’s, 10

Dual unit ball
 w*-metrizable, 30
 w*-sequentially compact, 21, 30

E-admissible ε-path, 155
 ε polygonal path, 155
 ε-net, 142
 Essentially smooth function, 24

Fréchet space, 115
Fubini's theorem
   Borwein-Moor Haar null version, 23

Gradient oscillation, 104, 144
Green’s theorem
   on arbitrary Banach space, 160
   on separable space, 158

Haar null set, 22
Hausdorff distance, 93

Ioffe’s G-subdifferential, 111

Jeyakumar’s convexificator, 44

Kenderov’s theorem, 144
Krein-Milman theorem converse, 19

Lattice operations, 88
Line integral, 154, 155
Lipschitz function
   directionally, 11
   locally, 5
Locally equi-Lipschitz functions, 87
Lower semi-Lipschitz function, 119

Maximal cyclically monotone, 14
Mazur’s theorem, 104

Mean-value theorem
   Borwein-Fitzpatrick-Giles, 19
   Lebesgue’s, 19
   Lebourg’s, 19
Michel-Penot subdifferential, 11
Minimal subdifferential, 23
Multifunction
   Borwein’s USC (CSC) hulls, 21
cuscò, 8
   minimal, 8
densely defined, 21
graph of, 8
locally bounded, 21
usco, 8

Nowhere monotone function, 31
nonmonotonic type, 32
oscillating on the right (left), 33
the first species, 31
the second species, 32

1-D almost everywhere set, 25

Parallel body, 92
Permutation-invariant function, 108
Polygonal path, 155
   closed, 155
Preiss' theorem, 7
Preiss-Phelps-Namioka theorem, 139
Prevalent set, 22
Proximal subdifferential, 9
Pseudo-regular function, 12
   partially pseudo-regular, 103

Quasi lower (upper) semicontinuous, 25

Rademacher’s theorem, 6
Regular function, 12
Robustly semicontinuous
   lower (upper), 95
Rockafellar type function, 34
Rockafellar’s theorem, 14
σ-minimal Clarke subdifferential, 144
Index

s-Hölder super-sub-differential, 9
   countable existence theorem, 61
s-minimal subdifferential, 148
   Approximate, 148
   Clarke, 150
Shur space, 16
Singular function, 42
Strictly differentiable
   Fréchet, 7
   Gâteaux, 7
Strong separation theorem, 18
Subdifferential integrable, 14
   strongly, 95
   weakly, 95

T-Lipschitz function, 87
The space of
   automorphism, 42
   bounded continuous function, 110
   continuous function, 44, 115
   K-nondecreasing function, 114
   nondecreasing continuous function, 40

Universally measurable
   Radon, 22

van der Waerden function, 51

Variational principle
   Borwein-Preiss, 27
   Ekeland’s, 27
   on finite dimensional subspaces, 28

Viscosity subdifferential, 9

Weierstrass function, 49

Y-strictly $\beta$-differentiable, 16
Zorn’s lemma, 18