LIPSCHITZ BEHAVIOR OF THE ROBUST REGULARIZATION

ADRIAN S. LEWIS AND C.H. JEFFREY PANG

ABSTRACT. To minimize or upper-bound the value of a function "robustly", we might instead minimize or upper-bound the "ϵ-robust regularization", defined as the map from a point to the maximum value of the function within an ϵ-radius. This regularization may be easy to compute: convex quadratics lead to semidefinite-representable regularizations, for example, and the spectral radius of a matrix leads to pseudospectral computations. For favorable classes of functions, we show that the robust regularization is Lipschitz around any given point, for all small $\epsilon > 0$, even if the original function is nonlipschitz (like the spectral radius). One such favorable class consists of the semi-algebraic functions. Such functions have graphs that are finite unions of sets defined by finitely-many polynomial inequalities, and are commonly encountered in applications.

CONTENTS

Key words: robust optimization, nonsmooth analysis, locally Lipschitz, regularization, semi-algebraic, pseudospectrum, robust control, semidefinite representable, prox-regularity.

AMS subject classifications: 93B35, 49K40, 65K10, 90C30, 15A18, 14P10.

Date: 1st October 2008.

1. INTRODUCTION

In the implementation of the optimal solution of an optimization model, one is not only concerned with the minimizer of the optimization model, but how numerical errors and perturbations in the problem description and implementation can affect the solution. We might therefore try to solve an optimization model in a robust manner. The issues of robust optimization, particularly in the case of linear and quadratic programming, are documented in [1].

A formal way to address robustness is to consider the "robust regularization" [15]. The notation "$\rightrightarrows$" denotes a set-valued map. That is, if $F : X \rightrightarrows Y$ and $x \in X$, then $F(x)$ is a subset of Y.

Definition 1.1. For $\epsilon > 0$ and $F : X \to \mathbb{R}^m$, where $X \subset \mathbb{R}^n$, the *set-valued robust regularization* $F_\epsilon : X \rightrightarrows \mathbb{R}^m$ is defined as

$$F_\epsilon(x) := \{F(x+e) \mid |e| \leq \epsilon, x+e \in X\}.$$

For the particular case of a real-valued function $f : X \to \mathbb{R}$, we define the *robust regularization* $\bar{f}_\epsilon : X \to \mathbb{R}$ of f by

$$\begin{aligned} \bar{f}_\epsilon(x) &:= \sup\{y \in f_\epsilon(x)\} \\ &= \sup\{y \mid \exists x' \in X \text{ such that } f(x') = y \text{ and } |x' - x| \leq \epsilon\}. \end{aligned}$$

In this paper, we restrict our attention to the real-valued robust regularization $\bar{f}_\epsilon : X \to \mathbb{R}$. The use of set-valued analysis is restricted to Section 4.

The minimizer of the robust regularization protects against small perturbations better, and might be a better solution to implement. We illustrate with the example

$$f(x) = \begin{cases} -x & \text{if } x < 0 \\ \sqrt{x} & \text{if } x \geq 0. \end{cases}$$

The robust regularization can be quickly calculated to be

$$\bar{f}_\epsilon(x) = \begin{cases} \epsilon - x & \text{if } x < \alpha(\epsilon) \\ \sqrt{\epsilon + x} & \text{if } x \geq \alpha(\epsilon), \end{cases}$$

where $\alpha(\epsilon) = \frac{1+2\epsilon-\sqrt{1+8\epsilon}}{2} > -\epsilon$. The minimizer of f is $\alpha(0)$, and f is not Lipschitz there. To see this, observe that $\frac{f(\delta)-f(0)}{\delta-0} \to \infty$ as $\delta \to 0$. But the robust regularization $\bar{f}_\epsilon$ is Lipschitz at its minimizer $\alpha(\epsilon)$; its left and right derivatives there are -1 and $\frac{1}{2\sqrt{\epsilon+\alpha(\epsilon)}}$, which are both finite.

The sensitivity of f at 0 can be attributed to the lack of Lipschitz continuity there. Lipschitz continuity is important in variational analysis, and is well studied in the recent books [23, 20]. The existence of a finite Lipschitz constant on f close to the optimizer can be important in the problems from which the optimization problem was derived.

There are two main aims in this paper. The first aim is to show that robust regularization has a regularizing property: Even if the original function f is not Lipschitz at a point x, the robust regularization can be Lipschitz there under various conditions. For example, in Corollary 4.6, we prove that if the set of points at which f is not Lipschitz is isolated, then the robust regularization $\bar{f}_\epsilon$ is Lipschitz at these points for all small $\epsilon > 0$. The second aim is to highlight the relationship between calmness and Lipschitz continuity, a topic important in the study of metric regularity, and studied in some generality for set-valued mappings (for example, in [17, Theorem 2.1], [21, Theorem 1.5]) but exploited less for single-valued mappings.

In Theorem 5.3, we prove that if $f : \mathbb{R}^n \to \mathbb{R}$ is semi-algebraic and continuous, then given any point in $\mathbb{R}^n$, the robust regularization $\bar{f}_\epsilon$ is Lipschitz there for all small $\epsilon > 0$. Semi-algebraic functions are functions whose graph can be defined by a finite union of sets defined by finitely many polynomial equalities and inequalities, and is a broad class of functions in applications. (For example, piecewise polynomial functions, rational functions and the mapping from a matrix to its eigenvalues are all semi-algebraic functions.) Moreover, the Lipschitz modulus of $\bar{f}_\epsilon$ at $\bar{x}$ is of order $o\left(\frac{1}{\epsilon}\right)$. This estimate of the Lipschitz modulus can be helpful for robust design.

Several interesting examples of robust regularization are tractable to compute and optimize. For example, the robust regularization of any strictly convex quadratic is a semidefinite -representable function, tractable via semidefinite programming: see Section 6. The robust regularizations of the spectral abscissa and radius of a nonsymmetric square matrix, which are the largest real part and the largest norm respectively of the eigenvalues of a matrix, are two more interesting examples. The robust regularization of the spectral abscissa and spectral radius are also known as the pseudospectral abscissa and the pseudospectral radius. The pseudospectral abscissa is important in the study of the system $\frac{d}{dt}u(t) = Au(t)$, and is easily calculated using the algorithm in [4], while the pseudospectral radius is important in the study of the system $u_{t+1} = Au_t$, and is easily calculated using the algorithm in [18]. We refer the reader to [27] for more details on the importance of the pseudospectral abscissa and radius in applications. The spectral abscissa is nonlipschitz whenever the eigenvalue with the largest real part has a nontrivial Jordan block. But for a fixed matrix, the pseudospectral abscissa is Lipschitz there for all $\epsilon \in (0, \bar{\epsilon})$ if $\bar{\epsilon} > 0$ is small enough [16]. We rederive this result here, using a much more general approach.

2. Calmness as an extension to Lipschitzness

We begin by discussing the relation between calmness and Lipschitz continuity, which will be important in the proofs in Section 5 later. Throughout the paper, we will limit ourselves to the single-valued case. For more on these topics and their set-valued extensions, we refer the reader to [23].

Definition 2.1. Let $F : X \to \mathbb{R}^m$ be a single-valued map, where $X \subset \mathbb{R}^n$.

(a) [23, Section 8F] Define the *calmness modulus* of F at $\bar{x}$ with respect to X to be

$$\begin{aligned} \text{calm}\, F(\bar{x}) &:= \inf\{\kappa \mid \text{There is a neighbourhood } V \text{ of } \bar{x} \text{ such that} \\ &\qquad |F(x) - F(\bar{x})| \leq \kappa |x - \bar{x}| \text{ for all } x \in V \cap X\} \\ &= \limsup_{x \xrightarrow[X]{} \bar{x}} \frac{|F(x) - F(\bar{x})|}{|x - \bar{x}|}. \end{aligned}$$

Here, $x \xrightarrow[X]{} \bar{x}$ means that $x \in X$ and $x \to \bar{x}$. The function F is *calm* at $\bar{x}$ with respect to X if $\text{calm}\, F(\bar{x}) < \infty$.

(b)[23, Definition 9.1] Define the *Lipschitz modulus* of F at $\bar{x}$ with respect to X to be

$$\begin{aligned} \text{lip}\, F(\bar{x}) &:= \inf\{\kappa \mid \text{There is a neighbourhood } V \text{ of } \bar{x} \text{ such that} \\ &\qquad |F(x) - F(x')| \leq \kappa |x - x'| \text{ for all } x, x' \in V \cap X\} \\ &= \limsup_{\substack{x, x' \xrightarrow[X]{} \bar{x} \\ x \neq x'}} \frac{|F(x) - F(x')|}{|x - x'|}. \end{aligned}$$

The function F is *Lipschitz* at $\bar{x}$ with respect to X if $\text{lip}\, F(\bar{x}) < \infty$. $\diamond$

The definitions differ slightly from that of [23]. As can be seen in the definitions, Lipschitz continuity is a more stringent form of continuity than calmness. In fact, they are related in the following manner.

Proposition 2.2. *Suppose that $F : X \to \mathbb{R}^m$ where $X \subset \mathbb{R}^n$.*

(a) $\limsup_{x \xrightarrow[X]{} \bar{x}} \text{calm}\, F(x) \leq \text{lip}\, F(\bar{x})$.

(b) If there is an open set U containing $\bar{x}$ such that $U \cap X$ is convex, then $\text{lip}\, F(\bar{x}) = \limsup_{x \xrightarrow[X]{} \bar{x}} \text{calm}\, F(x)$.

Proof. To simplify notation, let $\kappa := \limsup_{x \xrightarrow[X]{} \bar{x}} \text{calm}\, F(x)$.

(a) For any $\epsilon > 0$, we can find a point x_ϵ such that $|\bar{x} - x_\epsilon| < \epsilon$ and $\text{calm}\, F(x_\epsilon) > \kappa - \epsilon$. Then we can find a point $\tilde{x}_\epsilon$ such that $|x_\epsilon - \tilde{x}_\epsilon| < \epsilon$ and $|F(x_\epsilon) - F(\tilde{x}_\epsilon)| > (\kappa - \epsilon) |x_\epsilon - \tilde{x}_\epsilon|$. As ϵ can be made arbitrarily small, we have $\kappa \leq \text{lip}\, F(\bar{x})$ as needed.

(b) For every $\epsilon > 0$, there is some neighborhood of $\bar{x}$, say $\mathbb{B}_\delta(\bar{x})$, such that

$$\text{calm}\, F(x) \leq \kappa + \epsilon \text{ if } x \in \mathbb{B}_\delta(\bar{x}) \cap X.$$

For any $y, z \in \mathbb{B}_\delta(\bar{x}) \cap X$, consider the line segment joining y and z, which we denote $[y, z]$. As calm $F(\tilde{x}) \leq \kappa + \epsilon$ for all $\tilde{x} \in [y, z]$, there is a neighborhood around $\tilde{x}$, say $V_{\tilde{x}}$, such that $|F(\hat{x}) - F(\tilde{x})| \leq (\kappa + 2\epsilon)|\hat{x} - \tilde{x}|$ for all $\hat{x} \in V_{\tilde{x}} \cap X$.

As $[y, z]$ is compact, choose finitely many $\tilde{x}$ such that the union of $V_{\tilde{x}}$ covers $[y, z]$. We can add y and z into our choice of points and rename them as $\tilde{x}_1, \ldots, \tilde{x}_k$ in their order on the line segment $[y, z]$, with $\tilde{x}_1 = y$ and $\tilde{x}_k = z$. Also, we can find a point $\hat{x}_i$ between $\tilde{x}_i$ and $\tilde{x}_{i+1}$ such that $\hat{x}_i \in V_{\tilde{x}_i} \cap V_{\tilde{x}_{i+1}}$. Therefore, we add these $\hat{x}_i$ into $\tilde{x}_1, \ldots, \tilde{x}_k$ and get a new set $x_1, \ldots, x_K$, again in their order on the line segment and $x_1 = y$, $x_K = z$.

We have:

$$\begin{aligned} |F(y) - F(z)| &\leq \sum_{i=1}^{K-1} |F(x_i) - F(x_{i+1})| \\ &\leq \sum_{i=1}^{K-1} (\kappa + 2\epsilon)|x_i - x_{i+1}| \\ &\leq (\kappa + 2\epsilon)|y - z|, \end{aligned}$$

and as ϵ is arbitrary, lip $F(\bar{x}) \leq \kappa$ as claimed. ☐

Convexity is a strong assumption here, but some analogous condition is needed, as the following examples show.

Example 2.3. (a) Consider the set $X \subset \mathbb{R}$ defined by

$$X = \left(\bigcup_{i=1}^{\infty} \left[\frac{1}{3^i}, \frac{2}{3^i} \right] \right) \cup \{0\},$$

and define the function $F : X \to \mathbb{R}$ by

$$F(x) = \begin{cases} \frac{1}{3^i} & \text{if } \frac{1}{3^i} \leq x \leq \frac{2}{3^i}, \\ 0 & \text{if } x = 0. \end{cases}$$

It is clear that calm $F(x) = 0$ for all $x \in X \setminus \{0\}$ since F is constant on each component of X, and calm $F(0) = 1$. But

$$\begin{aligned} \text{lip}\, F(0) &= \lim_{i\to\infty} \frac{F\left(\frac{1}{3^i}\right) - F\left(\frac{2}{3^{i+1}}\right)}{\frac{1}{3^i} - \frac{2}{3^{i+1}}} \\ &= \lim_{i\to\infty} \frac{\frac{1}{3^i} - \frac{1}{3^{i+1}}}{\frac{1}{3^i} - \frac{2}{3^{i+1}}} \\ &= 2. \end{aligned}$$

Thus, $\limsup_{x\to 0} \text{calm}\, F(x) < \text{lip}\, F(0)$.

(b) Consider $X \subset \mathbb{R}^2$ defined by $X := \{(x_1, x_2) \mid x_2^2 = x_1^4\}$ and the function $F : \mathbb{R}^2 \to \mathbb{R}$ defined by $F(x_1, x_2) = x_2$. One can easily check that $\limsup_{x\to 0} \text{calm}\, F(x) = 0$ and $\text{lip}\, F(0,0) = 1$. This is an example of a semi-algebraic function where inequality holds.$\diamond$

Note that $\text{calm}\, F(\bar{x})$ can be strictly smaller than $\text{lip}\, F(\bar{x})$ even if X is convex, as demonstrated below.

Example 2.4. (a) Consider $F : \mathbb{R} \to \mathbb{R}$ defined by

$$F(x) = \begin{cases} 0 & \text{if } x = 0, \\ x^2 \sin\left(\frac{1}{x^2}\right) & \text{otherwise.} \end{cases}$$

Here, $\text{calm}\, F(0) = 0$, but $\text{lip}\, F(0) = \infty$.

(b) Consider $F : \mathbb{R}^2 \to \mathbb{R}$ defined by:

$$F(x_1, x_2) = \begin{cases} 0 & \text{if } x_1 \leq 0 \\ x_1 & \text{if } 0 \leq x_1 \leq x_2/2 \\ -x_1 & \text{if } 0 \leq x_1 \leq -x_2/2 \\ 2x_2 & \text{if } x_1 \geq |x_2|/2. \end{cases}$$

We can calculate $\text{calm}\, F(0,0) = 2/\sqrt{5}$, and $\text{lip}\, F(0,0) = 2$, so this gives $\text{calm}\, F(0,0) < \text{lip}\, F(0,0)$. This is an example of a semi-algebraic function where inequality holds.$\diamond$

At this point, we make a remark about subdifferentially regular functions. We recall the definition of subdifferential regularity.

Definition 2.5. [23, Definition 8.3] Consider a function $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ and a point $\bar{x}$ with $f(\bar{x})$ finite. For a vector $v \in \mathbb{R}^n$, one says that

(a) v is a *regular subgradient* of f at $\bar{x}$, written $v \in \hat{\partial} f(\bar{x})$, if

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

(b) v is a *(general) subgradient* of f at $\bar{x}$, written $v \in \partial f(\bar{x})$, if there are sequences $x^\nu \to \bar{x}$ and $v^\nu \in \hat{\partial} f(x^\nu)$ with $v^\nu \to v$ and $f(x^\nu) \to f(\bar{x})$.

(c) If f is Lipschitz continuous at $\bar{x}$, then f is *subdifferentially regular* if $\hat{\partial} f(\bar{x}) = \partial f(\bar{x})$.

Though the definition of subdifferential regularity differs from that given in [23, Definition 7.25], it can be deduced from [23, Corollary 8.11, Theorem 9.13 and Theorem 8.6] when f is Lipschitz, and is simple enough for our purposes. Subdifferentially regular functions are important and well-studied in variational analysis. The class of subdifferentially regular functions is closed under sums and pointwise maxima, and includes smooth functions and convex functions. It turns out that the calmness and Lipschitz moduli are equal for subdifferentially regular functions.

Proposition 2.6. *If* $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ *is Lipschitz continuous at* $\bar{x}$ *and subdifferentially regular there, then* calm $f(\bar{x}) =$ lip $f(\bar{x})$.

Proof. By [23, Theorem 9.13], lip $f(\bar{x}) = \max\{|v| \mid v \in \partial f(\bar{x})\}$. If $v \in \partial f(\bar{x})$, then $v \in \hat{\partial} f(\bar{x})$, and we observe that calm $f(\bar{x}) \geq |v|$ because

$$\begin{aligned} f(\bar{x} + tv) &\geq f(\bar{x}) + \langle v, tv\rangle + o(|t|) \\ &= f(\bar{x}) + |v|\,|tv| + o(|t|)\,. \end{aligned}$$

Therefore calm $f(\bar{x}) \leq$ lip $f(\bar{x}) = \max\{|v| \mid v \in \partial f(\bar{x})\} \leq$ calm $f(\bar{x})$, which implies that all three terms are equal. □

3. Calmness and robust regularization

Recall the definition of robust regularization in Definition 1.1. To study robust regularization, it is useful to study the dependence of $\bar{f}_\epsilon(x)$ on ϵ instead of on x. For a point $x \in X$, define $g_x : \mathbb{R}_+ \to \mathbb{R}$ by

$$g_x(\epsilon) = \bar{f}_\epsilon(x)\,.$$

To simplify notation, we write $g \equiv g_x$ if it is clear from context. Here are a few basic properties of g_x.

Proposition 3.1. *For* $f : X \to \mathbb{R}$ *and* g_x *as defined above, we have the following:*

(a) g_x *is monotonically nondecreasing.*

(b) If f *is continuous in a neighborhood of* x*, then* g_x *is continuous in a neighborhood of* 0.

Proof. Part (a) is obvious. For part (b), we prove the left and right limits separately. Suppose that $\epsilon_i \downarrow \epsilon$. There is a sequence of x_i such that $f(x_i) = \bar{f}_{\epsilon_i}(x)$, and $|x_i - x| \leq \epsilon_i$. We assume, by choosing a subsequence if needed, that $\lim_{i\to\infty} x_i = \tilde{x}$. We have $|\tilde{x} - x| \leq \epsilon$, and since f is continuous, $f(x_i) \to f(\tilde{x})$. This means that

$$\bar{f}_\epsilon(x) \geq f(\tilde{x}) = \lim_{i\to\infty} \bar{f}_{\epsilon_i}(x)\,,$$

which implies $g(\epsilon) \geq \limsup_{\tilde{\epsilon}\downarrow\epsilon} g(\tilde{\epsilon})$. The monotonicity of g tells us that $g(\epsilon) = \lim_{\tilde{\epsilon}\downarrow\epsilon} g(\tilde{\epsilon})$.

Next, suppose that ϵ_i increases monotonically to ϵ. Let $\hat{x}$ be such that $f(\hat{x}) = \bar{f}_\epsilon(x)$, with $|\hat{x} - x| \leq \epsilon$. Since f is continuous, for every $\delta_1 > 0$, there is a $\delta_2 > 0$ such that $|f(x') - f(\hat{x})| < \delta_1$ if $|x' - \hat{x}| < \delta_2$. This means that if $\epsilon - \epsilon_i < \delta_2$, then

$$\bar{f}_{\epsilon_i}(x) \geq f(\hat{x}) - \delta_1 = \bar{f}_\epsilon(x) - \delta_1.$$

As δ_1 can be made arbitrarily small, we conclude that $\lim_{\tilde{\epsilon}\uparrow\epsilon} \bar{f}_{\tilde{\epsilon}}(x) = \bar{f}_\epsilon(x)$, or $\lim_{\tilde{\epsilon}\uparrow\epsilon} g(\tilde{\epsilon}) = g(\epsilon)$. □

It turns out that calmness of the robust regularization is related to the derivative of g_x.

Proposition 3.2. *If* $f : X \to \mathbb{R}$ *and* $\epsilon > 0$*, then* calm $\bar{f}_\epsilon(x) \leq$ calm $g_x(\epsilon)$. *If in addition* $X = \mathbb{R}^n$ *and* g_x *is differentiable at* ϵ*, then*

$$\text{calm}\, \bar{f}_\epsilon(x) = \text{calm}\, g_x(\epsilon) = g_x'(\epsilon).$$

Proof. For the first part, we proceed to show that if $\kappa >$ calm $g_x(\epsilon)$, then $\kappa \geq$ calm $\bar{f}_\epsilon(x)$. If $|\tilde{x} - x| < \epsilon$, we have

$$\mathbb{B}_{\epsilon - |\tilde{x}-x|}(x) \subset \mathbb{B}_\epsilon(\tilde{x}) \subset \mathbb{B}_{\epsilon + |\tilde{x}-x|}(x),$$

which implies

$$\bar{f}_{\epsilon - |\tilde{x}-x|}(x) \leq \bar{f}_\epsilon(\tilde{x}) \leq \bar{f}_{\epsilon + |\tilde{x}-x|}(x).$$

Then note that if $\tilde{x}$ is close enough to x, we have

$$\bar{f}_\epsilon(\tilde{x}) \leq \bar{f}_{\epsilon + |\tilde{x}-x|}(x) = g_x(\epsilon + |\tilde{x} - x|) \leq g_x(\epsilon) + \kappa |\tilde{x} - x|,$$

and similarly

$$\bar{f}_\epsilon(\tilde{x}) \geq \bar{f}_{\epsilon - |\tilde{x}-x|}(x) = g_x(\epsilon - |\tilde{x} - x|) \geq g_x(\epsilon) - \kappa |\tilde{x} - x|,$$

which tells us that $\left|\bar{f}_\epsilon(\tilde{x}) - \bar{f}_\epsilon(x)\right| \leq \kappa |\tilde{x} - x|$, which is what we need.

For the second part, it is clear from the definition of the derivative that $g_x'(\epsilon) =$ calm $g_x(\epsilon)$. We prove that if $\kappa < g_x'(\epsilon)$, then $\kappa \leq$ calm $\bar{f}_\epsilon(x)$. By the differentiability of g_x, there is some $\bar{\delta} > 0$ such that for any $0 \leq \delta \leq \bar{\delta}$, we have

$$\begin{aligned} \bar{f}_{\epsilon+\delta}(x) &= g_x(\epsilon + \delta) \\ &> g_x(\epsilon) + \kappa\delta \\ &= \bar{f}_\epsilon(x) + \kappa\delta. \end{aligned}$$

For any $0 \leq \delta \leq \bar{\delta}$, there is some $\tilde{x}_\delta \in \mathbb{B}_{\epsilon+\delta}(x)$ such that $f(\tilde{x}_\delta) = \bar{f}_{\epsilon+\delta}(x)$. Let $\hat{x}_\delta = \frac{\delta}{|\tilde{x}_\delta - x|}(\tilde{x}_\delta - x) + x$. We have $\bar{f}_\epsilon(\hat{x}_\delta) = \bar{f}_{\epsilon+\delta}(x)$, which gives $\bar{f}_\epsilon(\hat{x}_\delta) - \bar{f}_\epsilon(x) > \kappa\delta$. Since $\hat{x}_\delta$ was chosen such that $\delta = |\hat{x}_\delta - x|$, we have $\bar{f}_\epsilon(\hat{x}_\delta) - \bar{f}_\epsilon(x) > \kappa |\hat{x}_\delta - x|$, which implies $\kappa \leq$ calm $\bar{f}_\epsilon(x)$ as needed. □

Remark 3.3. A similar statement can be made for $\epsilon = 0$, except that we change calmness to "calm from above" as defined in [23, Section 8F] in both parts.

We have the following corollary. The subdifferential "∂" was defined in Definition 2.5.

Corollary 3.4. *If* $f : \mathbb{R}^n \to \mathbb{R}$, $\epsilon > 0$ *and* g_x *is Lipschitz at* ϵ*, then*

$$\text{calm}\, \bar{f}_\epsilon(x) \leq \text{lip}\, g_x(\epsilon) = \sup\{|y| \mid y \in \partial g_x(\epsilon)\}.$$

Proof. It is clear that calm $\bar{f}_\epsilon(x) \leq$ calm $g_x(\epsilon) \leq$ lip $g_x(\epsilon)$. The formula lip $g_x(\epsilon) = \sup\{|y| \mid y \in \partial g_x(\epsilon)\}$ follows from [23, Theorem 9.13, Definition 9.1]. □

In general, the robust regularization is calm.

Proposition 3.5. *For a continuous function $f : X \to \mathbb{R}$, there is an $\bar{\epsilon} > 0$ such that $\bar{f}_\epsilon$ is calm at x for all $0 < \epsilon \leq \bar{\epsilon}$ except on a subset of $(0, \bar{\epsilon}]$ of measure zero.*

Proof. By Proposition 3.1(b), since f is continuous at x, g_x is continuous in $[0, \bar{\epsilon}]$ for some $\bar{\epsilon} > 0$. Since g_x is monotonically nondecreasing, it is differentiable in all $[0, \bar{\epsilon}]$ except for a set of measure zero. The derivative $g'_x(\epsilon)$ equals calm $\bar{f}_\epsilon(x)$ by Proposition 3.2. □

Remark 3.6. In general, the above result cannot be improved. For an example, let $c : [0,1] \to [0,1]$ denote the Cantor function, commonly used in real analysis texts as an example of a function that is not absolutely continuous and not satisfying the Fundamental Theorem of Calculus. Then calm $\bar{c}_\epsilon(0) = \infty$ for all ϵ lying in the Cantor set.$\diamond$

4. ROBUST REGULARIZATION IN GENERAL

In this section, in Corollary 4.6, we prove that if lip $f(x) < \infty$ for x close to but not equal to $\bar{x}$, then lip $\bar{f}_\epsilon(\bar{x}) < \infty$ for all small $\epsilon > 0$, even when lip $f(\bar{x}) = \infty$. To present the details of the proof, we need a short foray into set-valued analysis.

Definition 4.1. [23, Example 4.13] For two sets $C, D \subset \mathbb{R}^m$, the *Pompieu-Hausdorff distance* between C and D, denoted by $\mathbf{d}(C,D)$, is defined by

$$\mathbf{d}(C,D) := \inf\{\eta \geq 0 \mid C \subset D + \eta\mathbb{B}, D \subset C + \eta\mathbb{B}\}.$$

Definition 4.2. [23, Definitions 9.26, 9.28] A mapping $S : X \rightrightarrows \mathbb{R}^m$ is *Lipschitz continuous* on its domain $X \subset \mathbb{R}^n$, if it is nonempty-closed-valued on X and there exists $\kappa \geq 0$, a Lipschitz constant, such that

$$\mathbf{d}(S(x'), S(x)) \leq \kappa |x' - x| \text{ for all } x, x' \in X,$$

or equivalently, $S(x') \subset S(x) + \kappa |x' - x| \mathbb{B}$ for all $x, x' \in X$. The *Lipschitz modulus* is defined as

$$\text{lip}\, S(\bar{x}) := \limsup_{\substack{x,x' \xrightarrow[X]{} \bar{x} \\ x \neq x'}} \frac{\mathbf{d}(S(x'), S(x))}{|x' - x|},$$

and is the infimum of all κ such that there exists a neighborhood U of $\bar{x}$ such that S is Lipschitz continuous with constant κ in $U \cap X$.$\diamond$

For $F : X \to \mathbb{R}^m$, we may write the robust regularization $F_\epsilon : X \rightrightarrows \mathbb{R}^m$ by $F_\epsilon = F \circ \Phi_\epsilon$, where $\Phi_\epsilon : X \rightrightarrows X$ is defined by $\Phi_\epsilon(x) = \mathbb{B}_\epsilon(x) \cap X$. For reasons that will be clear later in Section 7, we consider the extension $\tilde{\Phi}_\epsilon : \mathbb{R}^n \rightrightarrows X$ defined by $\tilde{\Phi}_\epsilon(x) = \mathbb{B}_\epsilon(x) \cap X$. It is clear that $\tilde{\Phi}_\epsilon |_X = \Phi_\epsilon$ using our previous notation, and it follows straight from the definitions that $\operatorname{lip} \Phi_\epsilon(x) \leq \operatorname{lip} \tilde{\Phi}_\epsilon(x)$ for $x \in X$.

Definition 4.3. We say that $X \subset \mathbb{R}^n$ is *peaceful at* $\bar{x} \in X$ if $\operatorname{lip} \Phi_\epsilon(\bar{x})$ is finite for all small $\epsilon > 0$. If in addition $\limsup_{\epsilon \downarrow 0} \operatorname{lip} \tilde{\Phi}_\epsilon(\bar{x}) \leq \kappa$ for all small $\epsilon > 0$, we say that X is peaceful with modulus κ at $\bar{x}$, or *κ-peaceful at $\bar{x}$*.

When $\bar{x}$ lies in the interior of X and ϵ is small enough, then $\tilde{\Phi}_\epsilon$ is Lipschitz with constant 1. In section 7, we will find weaker conditions on X for the Lipschitz continuity of $\tilde{\Phi}_\epsilon$. We will see that convex sets are 1-peaceful, but for now, we remark that if X is convex, then Φ_ϵ is globally Lipschitz in X.

Proposition 4.4. *If X is a convex set, then $\Phi_\epsilon(x) \subset \Phi(x') + |x - x'| \mathbb{B}$ for all $x, x' \in X$.*

Proof. The condition we are required to prove is equivalent to

$$\mathbb{B}_\epsilon(x) \cap X \subset (\mathbb{B}_\epsilon(x') \cap X) + |x - x'| \mathbb{B} \text{ for } x, x' \in X.$$

For any point $\tilde{x} \in \mathbb{B}_\epsilon(x) \cap X$, the line segment $[x', \tilde{x}]$ lies in X, and is of length at most $|\tilde{x} - x| + |x - x'|$. The ball $\mathbb{B}_\epsilon(x')$ can contain the line segment $[x', \tilde{x}]$, in which case $\tilde{x} \in \mathbb{B}_\epsilon(x') \cap X$, or the boundary of $\mathbb{B}_\epsilon(x')$ may intersect $[x', \tilde{x}]$ at a point, say $\hat{x}$. Since X is a convex set, we have $\hat{x} \in \mathbb{B}_\epsilon(x') \cap X$. Furthermore

$$\begin{aligned} |\tilde{x} - \hat{x}| &= |\tilde{x} - x'| - \epsilon \\ &\leq |\tilde{x} - x| + |x - x'| - \epsilon \\ &\leq |x - x'|, \end{aligned}$$

so $\tilde{x} \in (\mathbb{B}_\epsilon(x') \cap X) + |x - x'| \mathbb{B}$. ☐

We remark that if X is nearly radial at $\bar{x}$ as introduced in [15], then X is 1-peaceful: see Section 7. The set X is *nearly radial at $\bar{x}$* if

$$\operatorname{dist}(\bar{x}, x + T_X(x)) \to 0 \text{ as } x \to \bar{x} \text{ in } X.$$

The set X is *nearly radial* if it is nearly radial at all points in X. The notation $T_X(x)$ refers to the *(Bouligand) tangent cone* (or "contingent cone") to X at $x \in X$, formally defined as

$$T_X(\bar{x}) = \{\lim t_r^{-1}(x_r - \bar{x}) : t_r \downarrow 0, \ \ x_r \to \bar{x}, \ \ x_r \in X\}$$

(see, for example, [23, Definition 6.1]). Many sets are nearly radial, including for instance semi-algebraic sets, amenable sets and smooth manifolds.

We now present a result on the regularizing property of robust regularization. In Proposition 4.5 below, condition (i) allows us to evaluate the Lipschitz modulus of functions whose domains are not necessarily convex. One situation where (i) is interesting is when X is a smooth manifold.

Proposition 4.5. *For $F : X \to \mathbb{R}^m$, suppose that either (i) or (ii) holds.*

(i) X is peaceful and $\operatorname{lip} \tilde{F}(x) < \infty$ *for all x close to but not equal to $\bar{x}$. Here, $\tilde{F} : \mathbb{R}^n \to \mathbb{R}^m$ is an extension of F on $\mathbb{R}^n$ such that $\tilde{F}|_X = F$.*

(ii) X is convex and $\operatorname{lip} F(x) < \infty$ *for all $x \in X$ close to but not equal to $\bar{x}$.*

Then $\operatorname{lip} F_\epsilon(\bar{x})$ *is finite for all small $\epsilon > 0$.*

Proof. The proof for both conditions are similar, so they will be treated together. One notes that $\operatorname{lip} F(x) \leq \operatorname{lip} \tilde{F}(x)$ always by the definition of these Lipschitz moduli, so we assume $\operatorname{lip} F(x) < \infty$ for all $x \in X$ close to but not equal to $\bar{x}$ until we have to distinguish these cases.

First, we prove that $\operatorname{lip} F : X \to \mathbb{R}_+$ is upper semicontinuous. This result is just a slight modification of the first part of [23, Theorem 9.2], but we include the proof for completeness. Suppose that $x_i \to x$. By the definition of $\operatorname{lip} F$, we can find $x_{i,1}$, $x_{i,2} \in X$ such that

$$\begin{aligned} \frac{|F(x_{i,1}) - F(x_{i,2})|}{|x_{i,1} - x_{i,2}|} &> \operatorname{lip} F(x_i) - |x_i - x|, \\ \text{and } |x_{i,j} - x_i| &< |x_i - x| \text{ for } j = 1, 2. \end{aligned}$$

Taking limits as $i \to \infty$, we see that $x_{i,1}, x_{i,2} \to x$, and it follows that

$$\begin{aligned} \operatorname{lip} F(x) &\geq \limsup_{i \to \infty} \frac{|F(x_{i,1}) - F(x_{i,2})|}{|x_{i,1} - x_{i,2}|} \\ &= \limsup_{i \to \infty} \operatorname{lip} F(x_i). \end{aligned}$$

Thus $\operatorname{lip} F : X \to \mathbb{R}_+$ is upper semicontinuous.

So for ϵ_1 small enough, choose $\epsilon_2 < \epsilon_1$ such that $\operatorname{lip} F$ is bounded above in $C_1 = (\mathbb{B}_{\epsilon_1+\epsilon_2}(\bar{x}) \backslash \mathbb{B}_{\epsilon_1-\epsilon_2}(\bar{x})) \cap X$, say by the constant κ_1. Then for any $\kappa_2 > \kappa_1$ and any $x \in C_1$, there is an ϵ_x such that F is Lipschitz continuous on $\mathbb{B}_{\epsilon_x}(x) \cap X$ with constant κ_2 with respect to X. Thus $\cup_{x \in C_1} \{\mathbb{B}_{\epsilon_x}(x)\}$ is an open cover of C_1.

By the Lebesgue Number Lemma, there is a constant δ such that if x_1, x_2 lie in C_1 and $|x_1 - x_2| \leq \delta$, then the line segment $[x_1, x_2]$ lies in one of the open balls $\mathbb{B}_{\epsilon_x}(x)$ for some $x \in C_1$. We may assume that $\delta < \epsilon_2$.

Also, since X is peaceful at $\bar{x}$, choose ϵ_1 small enough so that $\operatorname{lip} \Phi_{\epsilon_1}(\bar{x})$ is finite, say $\operatorname{lip} \Phi_{\epsilon_1}(\bar{x}) < K$. If X is convex, then this is possible due to

Proposition 4.4. We can assume that $K > 2$. Therefore, there is an open set U about $\bar{x}$ such that Φ_{ϵ_1} is Lipschitz in $U \cap X$ with constant K, that is $\Phi_{\epsilon_1}(x) \subset \Phi_{\epsilon_1}(x') + K|x - x'|\mathbb{B}$ for all $x, x' \in U \cap X$.

So, for $x, x' \in U \cap \mathbb{B}_{\frac{\delta}{2K}}(\bar{x}) \cap X$, we want to show that

$$F_{\epsilon_1}(x) \subset F_{\epsilon_1}(x') + K\kappa_2 |x - x'|\mathbb{B}.$$

Suppose that $y \in F_{\epsilon_1}(x)$. So $y = F(\tilde{x})$ for some $\tilde{x} \in \mathbb{B}_{\epsilon_1}(x) \cap X$. If $\tilde{x} \in \mathbb{B}_{\epsilon_1 - \frac{\delta}{2K}}(\bar{x})$, then $\tilde{x} \in \mathbb{B}_{\epsilon_1}(x') \cap X$ because $|x' - \bar{x}| \le \frac{\delta}{2K}$. So $y \in F_{\epsilon_1}(x')$. Otherwise $\tilde{x} \in \left(\mathbb{B}_{\epsilon_1 + \frac{\delta}{2K}}(\bar{x}) \backslash \mathbb{B}_{\epsilon_1 - \frac{\delta}{2K}}(\bar{x})\right) \cap X$.

We have $\Phi_{\epsilon_1}(x) \subset \Phi_{\epsilon_1}(x') + K|x - x'|\mathbb{B}$. So there is some $\hat{x} \in \Phi_{\epsilon_1}(x')$ such that

$$|\hat{x} - \tilde{x}| \le K|x - x'| \le K\frac{\delta}{2K} = \frac{\delta}{2}.$$

Furthermore,

$$|\hat{x} - \bar{x}| \le |\tilde{x} - x| + |x - \bar{x}| + |\hat{x} - \tilde{x}| \le \epsilon_1 + \frac{\delta}{2K} + \frac{\delta}{2} \le \epsilon_1 + \frac{3\delta}{4} < \epsilon_1 + \epsilon_2,$$

and

$$|\hat{x} - \bar{x}| \ge |\tilde{x} - x| - |x - \bar{x}| - |\hat{x} - \tilde{x}| \ge \epsilon_1 - \frac{\delta}{2K} - \frac{\delta}{2} \ge \epsilon_1 - \frac{3\delta}{4} > \epsilon_1 - \epsilon_2.$$

Hence $\hat{x} \in (\mathbb{B}_{\epsilon_1 + \epsilon_2}(\bar{x}) \backslash \mathbb{B}_{\epsilon_1 - \epsilon_2}(\bar{x})) \cap X$. We now proceed to prove the inequality $|F(\tilde{x}) - F(\hat{x})| < \kappa_2 |\hat{x} - \tilde{x}|$ for the two cases.

Condition (i): Since $|\hat{x} - \tilde{x}| < \delta$, the line segment $[\hat{x}, \tilde{x}]$ lies in $\mathbb{B}_{\epsilon_x}(x)$ for some $x \in X$. Since the line segment $[\hat{x}, \tilde{x}]$ is convex and lip $\tilde{F}$ is bounded from above by κ_2 there, we have

$$\begin{aligned} |F(\tilde{x}) - F(\hat{x})| &= \left|\tilde{F}(\tilde{x}) - \tilde{F}(\hat{x})\right| \\ &< \kappa_2 |\tilde{x} - \hat{x}| \end{aligned}$$

by [23, Theorem 9.2].

Condition (ii): The proof is similar, except that $[\hat{x}, \tilde{x}] \subset X$, and lip F is bounded above by κ_2.

On establishing $|F(\tilde{x}) - F(\hat{x})| < \kappa_2 |\hat{x} - \tilde{x}|$, we note that

$$\begin{aligned} F(\tilde{x}) &\in F(\hat{x}) + \kappa_2 |\hat{x} - \tilde{x}|\mathbb{B} \\ &\subset F_{\epsilon_1}(x') + \kappa_2 |\hat{x} - \tilde{x}|\mathbb{B} \\ &\subset F_{\epsilon_1}(x') + K\kappa_2 |x - x'|\mathbb{B}, \end{aligned}$$

and we are done. □

We are now ready to relate lip $\bar{f}_\epsilon(\bar{x})$ to lip $f(\bar{x})$. We remind the reader that in the proof of Corollary 4.6 below, $f_\epsilon : X \rightrightarrows \mathbb{R}$ is a set-valued map as introduced in Definition 1.1, which is similar to $\bar{f}_\epsilon$ but maps to intervals in $\mathbb{R}$.

Corollary 4.6. *For $f : X \to \mathbb{R}$, if either condition (i) or condition (ii) in Proposition 4.5 for $F : X \to \mathbb{R}$ taken to be f holds, then* $\operatorname{lip} \bar{f}_\epsilon(\bar{x}) < \infty$ *for all small $\epsilon > 0$.*

Proof. By Proposition 4.5, we have $\operatorname{lip} f_\epsilon(\bar{x}) < \infty$ with the given conditions. It remains to prove that $\operatorname{lip} \bar{f}_\epsilon(\bar{x}) \leq \operatorname{lip} f_\epsilon(\bar{x})$. We can do this by proving that $\operatorname{lip} \bar{S}(\bar{x}) \leq \operatorname{lip} S(\bar{x})$, where $S : X \rightrightarrows \mathbb{R}$ is a set-valued map, and $\bar{S} : X \to \mathbb{R}$ is defined by $\bar{S}(x) = \sup\{y \mid y \in S(x)\}$. Note that if $S = f_\epsilon$, then $\bar{S} = \overline{(f_\epsilon)} = \bar{f}_\epsilon$.

For any $\kappa > \operatorname{lip} S(x)$, we have $\mathbf{d}(S(\tilde{x}), S(\hat{x})) \leq \kappa|\tilde{x} - \hat{x}|$ for $\tilde{x}, \hat{x} \in X$ close enough to x by [23, Definition 9.26]. The definition of the Pompeiu-Hausdorff distance tells us that $S(\tilde{x}) \subset S(\hat{x}) + \kappa|\tilde{x} - \hat{x}|$, which implies $\bar{S}(\tilde{x}) \leq \bar{S}(\hat{x}) + \kappa|\tilde{x} - \hat{x}|$. By reversing the roles of $\tilde{x}$ and $\hat{x}$, we obtain $\left|\bar{S}(\tilde{x}) - \bar{S}(\hat{x})\right| \leq \kappa|\tilde{x} - \hat{x}|$. So $\kappa > \operatorname{lip} \bar{S}(x)$, and since κ is arbitrary, we have $\operatorname{lip} \bar{S}(x) \leq \operatorname{lip} S(x)$ as needed. □

5. SEMI-ALGEBRAIC ROBUST REGULARIZATION

In this section, in Theorem 5.3, we prove that if $f : \mathbb{R}^n \to \mathbb{R}$ is continuous and semi-algebraic, then at any given point, the robust regularization is locally Lipschitz there for all sufficiently small $\epsilon > 0$. This theorem is more appealing than Corollary 4.6 because the required condition is weaker. The condition $\operatorname{lip} f(x) < \infty$ for all x close to but not equal to $\bar{x}$ in Corollary 4.6 is a strong condition because if a function is not Lipschitz at a point $\bar{x}$, it is likely that it is not Lipschitz at some points close to $\bar{x}$ as well. For example in $f : \mathbb{R}^2 \to \mathbb{R}$ defined by $f(x_1, x_2) = \left|\sqrt{x_1}\right|$, f is not Lipschitz at all points where $x_1 = 0$.

We proceed to prove the main theorem of this section in the steps outlined below.

Proposition 5.1. *For $f : X \to \mathbb{R}$, where $X \subset \mathbb{R}^n$ is convex, define $G : X \times \mathbb{R}_+ \to \mathbb{R}_+ \cup \{\infty\}$ by*

$$G(x, \epsilon) := \limsup_{\tilde{\epsilon} \to \epsilon} \operatorname{lip} \bar{f}_{\tilde{\epsilon}}(x).$$

If f is semi-algebraic, then the maps $(x, \epsilon) \mapsto$ $\operatorname{calm} \bar{f}_\epsilon(x)$, $(x, \epsilon) \mapsto \operatorname{lip} \bar{f}_\epsilon(x)$ *and G are semi-algebraic.*

Proof. The semi-algebraic nature is a consequence of the Tarski-Seidenberg quantifier elimination. □

The semi-algebraicity of $(x, \epsilon) \mapsto \operatorname{calm} \bar{f}_\epsilon(x)$ gives us an indication of how the map $\epsilon \mapsto \operatorname{calm} \bar{f}_\epsilon(x)$ behaves asymptotically.

Proposition 5.2. *Suppose that $f : X \to \mathbb{R}$ is continuous and semi-algebraic, where $X \subset \mathbb{R}^n$. Fix $x \in X$. Then* calm $\bar{f}_\epsilon(x) = o\left(\frac{1}{\epsilon}\right)$ *as $\epsilon \searrow 0$. Hence $\bar{f}_\epsilon$ is calm at x for all small $\epsilon > 0$.*

Proof. The map g_x is semi-algebraic because it can be written as a composition of semi-algebraic maps $\epsilon \mapsto (x, \epsilon) \mapsto \bar{f}_\epsilon(x)$. Thus g_x is differentiable on some open interval of the form $(0, \bar{\epsilon})$ for $\bar{\epsilon} > 0$. Recall that calm $g_x(\epsilon) = g'_x(\epsilon)$ by Proposition 3.2.

We show that for any $K > 0$, we can reduce $\bar{\epsilon}$ if necessary so that the map $\epsilon \mapsto$ calm $\bar{f}_\epsilon(x)$ is bounded from above by $\epsilon \mapsto \frac{K}{\epsilon}$ on $\epsilon \in [0, \bar{\epsilon}]$. For any $K > 0$, there exists an $\bar{\epsilon} > 0$ such that either $g'_x(\epsilon) \leq \frac{K}{\epsilon}$ for all $0 < \epsilon < \bar{\epsilon}$, or $g'_x(\epsilon) \geq \frac{K}{\epsilon}$ for all $0 < \epsilon < \bar{\epsilon}$. The latter cannot happen, otherwise for any $0 < \epsilon < \bar{\epsilon}$,

$$\begin{aligned} \bar{f}_\epsilon(x) - f(x) &= \int_0^\epsilon g'_x(s)\, ds \\ &\geq \int_0^\epsilon \frac{K}{s} ds = \infty. \end{aligned}$$

This contradicts the continuity of g_x. If ϵ is small enough, the derivatives of g_x exist for all small $\epsilon > 0$ and $g'_x(\epsilon) =$ calm $\bar{f}_\epsilon(x)$ by Proposition 3.2. This gives us the required result. ☐

Consider $f : [0, 1] \to \mathbb{R}$ defined by $f(x) = x^{1/k}$. Then $g_0(\epsilon) = \epsilon^{1/k}$, so calm $\bar{f}_\epsilon(0) = g'_0(\epsilon) = \frac{1}{k}\epsilon^{(1/k)-1}$. As $k \to \infty$, we see that the bound above is tight.

We are now ready to state the main theorem of this paper. In the particular case of $X = \mathbb{R}^n$, we have the following theorem.

Theorem 5.3. *Consider any continuous semi-algebraic function $f : \mathbb{R}^n \to \mathbb{R}$. At any fixed point $\bar{x} \in \mathbb{R}^n$, the robust regularization $\bar{f}_\epsilon$ is Lipschitz at $\bar{x}$, and its calmness and Lipschitz moduli,* calm $\bar{f}_\epsilon(\bar{x})$ *and* lip $\bar{f}_\epsilon(\bar{x})$, *agree for sufficiently small ϵ and behave like $o\left(\frac{1}{\epsilon}\right)$ as $\epsilon \downarrow 0$.*

Proof. In view of Proposition 5.2, we only need to prove the there is some $\bar{\epsilon} > 0$ such that lip $\bar{f}_\epsilon(\bar{x}) =$ calm $\bar{f}_\epsilon(\bar{x})$ for all $\epsilon \in (0, \bar{\epsilon}]$. We can assume that $g_{\bar{x}}$ is twice continuously differentiable in $(0, \bar{\epsilon}]$. The graph of $G : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}_+$ as defined in Proposition 5.1 is semi-algebraic, so by the decomposition theorem [10, Theorem 6.7], there is a finite partition of definable $\mathcal{C}^2$ manifolds $C_1, \ldots, C_l$ such that $G\mid_{C_i}$ is $\mathcal{C}^2$.

If the segment $\{\bar{x}\} \times (0, \bar{\epsilon}]$ lies in the (relative) interior of one definable manifold, then

$$\begin{aligned} \text{lip}\, \bar{f}_\epsilon(\bar{x}) &= \limsup_{\tilde{x}\to\bar{x}} \text{calm}\, \bar{f}_\epsilon(\tilde{x}) \text{ (by Proposition 2.2)} \\ &= \limsup_{\tilde{x}\to\bar{x}} g'_{\tilde{x}}(\epsilon) \text{ (by Proposition 3.2)} \\ &= g'_{\bar{x}}(\epsilon) \\ &= \text{calm}\, \bar{f}_\epsilon(\bar{x}), \end{aligned}$$

and we have nothing to do. Therefore, assume that the segment is on the boundary of two or more of the C_i.

Since G is semi-algebraic, the map $\epsilon \mapsto \limsup_{\alpha\to\epsilon} \text{lip}\, \bar{f}_\alpha(\bar{x})$ is semi-algebraic, so we can reduce $\bar{\epsilon} > 0$ as necessary such that either

(1) $\limsup_{\alpha\to\epsilon} \text{lip}\, \bar{f}_\alpha(\bar{x}) < \text{calm}\, \bar{f}_\epsilon(\bar{x})$ for all $\epsilon \in (0, \bar{\epsilon}]$, or
(2) $\limsup_{\alpha\to\epsilon} \text{lip}\, \bar{f}_\alpha(\bar{x}) = \text{calm}\, \bar{f}_\epsilon(\bar{x})$ for all $\epsilon \in (0, \bar{\epsilon}]$, or
(3) $\limsup_{\alpha\to\epsilon} \text{lip}\, \bar{f}_\alpha(\bar{x}) > \text{calm}\, \bar{f}_\epsilon(\bar{x})$ for all $\epsilon \in (0, \bar{\epsilon}]$.

Case (1) cannot hold because lip $\bar{f}_\epsilon(\bar{x}) \geq$ calm $\bar{f}_\epsilon(\bar{x})$. Case (2) is what we seek to prove, so we proceed to show that case (3) cannot happen by contradiction.

We can choose $\tilde{\epsilon}, M_1, M_2 > 0$ such that $0 < \tilde{\epsilon} < \bar{\epsilon}$ and

$$\text{calm}\, \bar{f}_\epsilon(\bar{x}) < M_2 < M_1 < \limsup_{\alpha\to\epsilon} \text{lip}\, \bar{f}_\alpha(\bar{x}) \text{ for all } \epsilon \in [\tilde{\epsilon}, \bar{\epsilon}].$$

We state and prove a lemma important to the rest of the proof before continuing.

Lemma 5.4. *There exists an interval (ϵ_1, ϵ_2) contained in $(\tilde{\epsilon}, \bar{\epsilon}]$ and a manifold $T_1 \subset \mathbb{R}^n \times \mathbb{R}_+$ such that*

(1) $\{\bar{x}\} \times (\epsilon_1, \epsilon_2) \subset$ cl (T_1).
(2) T_1 is an open $\mathcal{C}^2$ manifold.
(3) $H : \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}$, defined by $H(x, \epsilon) = \bar{f}_\epsilon(x)$, is $\mathcal{C}^2$ in T_1.
(4) For all $(x, \epsilon) \in T_1$, we have $M_1 \leq g'_x(\epsilon) < \infty$.
(5) $(x, \epsilon) \mapsto g'_x(\epsilon)$ is continuous in T_1.

Proof. Consider the set

$$T := \{(x, \epsilon) \mid M_1 \leq g'_x(\epsilon) < \infty\}.$$

First, we prove that $\{\bar{x}\} \times [\tilde{\epsilon}, \bar{\epsilon}] \subset$ cl T. It suffices to show that for all $\epsilon \in (\tilde{\epsilon}, \bar{\epsilon}]$, $(\bar{x}, \epsilon) \in$ cl T. This can in turn be proven by showing that for all $\delta > 0$, we can find x', ϵ' such that $|\bar{x} - x'| < \delta$, $|\epsilon - \epsilon'| < \delta$ such that $(x', \epsilon') \in T$, or equivalently, $M_1 \leq g'_{x'}(\epsilon') < \infty$.

Since $\limsup_{\alpha\to\epsilon} \text{lip}\, \bar{f}_\alpha(\bar{x}) > M_1$, there is some ϵ° such that $|\epsilon^\circ - \epsilon| < \frac{\delta}{2}$ and lip $\bar{f}_{\epsilon^\circ}(\bar{x}) > M_1$.

Next, since

$$\limsup_{x\to\bar{x}} |\partial g_x(\epsilon^\circ)| \geq \limsup_{x\to\bar{x}} \text{calm}\, \bar{f}_{\epsilon^\circ}(x) = \text{lip}\, \bar{f}_{\epsilon^\circ}(\bar{x}),$$

there is some x' such that $|\bar{x} - x'| < \delta$ and $|\partial g_{x'}(\epsilon^\circ)| > \frac{1}{2}\text{lip}\, \bar{f}_{\epsilon^\circ}(\bar{x}) + \frac{1}{2}M_1$.

Finally, since $g_{x'}(\cdot)$ is semi-algebraic, we can find some ϵ' such that $|\epsilon' - \epsilon^\circ| < \frac{\delta}{2}$, $g'_{x'}(\epsilon')$ is well defined and finite, and

$$g'_{x'}(\epsilon') > |\partial g_{x'}(\epsilon^\circ)| - \frac{1}{2}\left(\text{lip}\, \bar{f}_{\epsilon^\circ}(\bar{x}) - M_1\right) > M_1.$$

This choice of x' and ϵ' are easily verified to satisfy the requirements stated.

By the decomposition theorem [10, Theorem 6.7], T can be decomposed into a finite disjoint union of $\mathcal{C}^2$ smooth manifolds $T_1, T_2, \ldots, T_p$ on which H is $\mathcal{C}^2$. Since $\{\bar{x}\} \times [\tilde{\epsilon}, \bar{\epsilon}] \subset \text{cl}\, T$, there must be some T_i and (ϵ_1, ϵ_2) such that $\{\bar{x}\} \times (\epsilon_1, \epsilon_2) \subset \text{cl}\, T_i$. Without loss of generality, let one such T_i be T_1.

Conditions (1), (2), (3) and (4) are automatically satisfied. Note that $g'_x(\epsilon)$ is exactly the derivative of $H(\cdot,\cdot)$ with respect to the second coordinate, and so Property (5) is satisfied. This concludes the proof of the lemma. □

We now continue with the rest of the proof of the theorem. If T_1 is of dimension one, then we have $T_1 \supset \{\bar{x}\} \times (\epsilon_1, \epsilon_2)$. Recall that if the derivative $g'_{\bar{x}}(\epsilon)$ exists, then $g'_{\bar{x}}(\epsilon) = \text{calm}\, \bar{f}_\epsilon(\bar{x})$ by Proposition 3.2. This would mean that calm $\bar{f}_\epsilon(\bar{x}) \geq M_2$, which contradicts our earlier assumption of calm $\bar{f}_\epsilon(\bar{x}) < M_2$. Therefore, the manifold T_1 is of dimension at least two.

Using Lemma 5.7 which we will prove later, we can construct the map $\varphi : [0,1)\times(\hat{\epsilon}_1, \hat{\epsilon}_2) \to \text{cl}T_1$, such that its derivative with respect to the second variable exists and is continuous, and $\varphi(0,\epsilon) = (\bar{x}, \epsilon)$ for all $\epsilon \in (\hat{\epsilon}_1, \hat{\epsilon}_2)$.

For each $0 < \delta < 1$, consider the path $\tilde{x}_\delta : [\hat{\epsilon}_1, \hat{\epsilon}_2] \to \mathbb{R}^n$ defined by $\tilde{x}_\delta(\epsilon) := \varphi(\delta, \epsilon)$. We have

$$\begin{aligned}
& \bar{f}_{\hat{\epsilon}_2}(\tilde{x}_\delta(\hat{\epsilon}_2)) - \bar{f}_{\hat{\epsilon}_1}(\tilde{x}_\delta(\hat{\epsilon}_1)) \\
= & \int_{\hat{\epsilon}_1}^{\hat{\epsilon}_2} \nabla H(\tilde{x}_\delta(s), s) \cdot (\tilde{x}'_\delta(s), 1)\, ds \\
= & \int_{\hat{\epsilon}_1}^{\hat{\epsilon}_2} \nabla_x H(\tilde{x}_\delta(s), s) \cdot \tilde{x}'_\delta(s)\, ds + \int_{\hat{\epsilon}_1}^{\hat{\epsilon}_2} \nabla_s H(\tilde{x}_\delta(s), s)\, ds,
\end{aligned}$$

where $H(x,\epsilon) = \bar{f}_\epsilon(x)$. The second component of $\nabla H(\tilde{x}_\delta(s), s)$ is simply $g'_{\tilde{x}_\delta(s)}(s)$. The first component can be analyzed as follows:

$$\begin{aligned} & \nabla_x H(\tilde{x}_\delta(s), s) \cdot \tilde{x}'_\delta(s) \\ = \; & \lim_{t\to 0} \frac{1}{t}\left(H(\tilde{x}_\delta(s) + t\tilde{x}'_\delta(s), s) - H(\tilde{x}_\delta(s), s)\right) \\ = \; & \lim_{t\to 0} \frac{1}{t}\left(\bar{f}_s(\tilde{x}_\delta(s) + t\tilde{x}'_\delta(s)) - \bar{f}_s(\tilde{x}_\delta(s))\right). \end{aligned}$$

Provided that $t\left|\tilde{x}'_\delta(s)\right| < s$, $\mathbb{B}_{s-t\left|\tilde{x}'_\delta(s)\right|}(\tilde{x}_\delta(s)) \subset \mathbb{B}_s(\tilde{x}_\delta(s) + t\tilde{x}'_\delta(s))$, and so

$$\begin{aligned} & \nabla_x H(\tilde{x}_\delta(s), s) \cdot \tilde{x}'_\delta(s) \\ \geq \; & \lim_{t\to 0} \frac{1}{t}\left(\bar{f}_{s-t\left|\tilde{x}'_\delta(s)\right|}(\tilde{x}_\delta(s)) - \bar{f}_s(\tilde{x}_\delta(s))\right) \\ = \; & \left|\tilde{x}'_\delta(s)\right| \lim_{t\to 0} \frac{1}{t\left|\tilde{x}'_\delta(s)\right|}\left(\bar{f}_{s-t\left|\tilde{x}'_\delta(s)\right|}(\tilde{x}_\delta(s)) - \bar{f}_s(\tilde{x}_\delta(s))\right) \\ = \; & -\left|\tilde{x}'_\delta(s)\right| g'_{\tilde{x}_\delta(s)}(s). \end{aligned}$$

Hence,

$$\begin{aligned} & \bar{f}_{\hat{\epsilon}_2}(\tilde{x}_\delta(\hat{\epsilon}_2)) - \bar{f}_{\hat{\epsilon}_1}(\tilde{x}_\delta(\hat{\epsilon}_1)) \\ = \; & \int_{\hat{\epsilon}_1}^{\hat{\epsilon}_2} \nabla_x H(\tilde{x}_\delta(s), s) \cdot \tilde{x}'_\delta(s)\, ds + \int_{\hat{\epsilon}_1}^{\hat{\epsilon}_2} \nabla_s H(\tilde{x}_\delta(s), s)\, ds \\ \geq \; & \int_{\hat{\epsilon}_1}^{\hat{\epsilon}_2} \left(1 - \left|\tilde{x}'_\delta(s)\right|\right) g'_{\tilde{x}_\delta(s)}(s)\, ds. \end{aligned}$$

Since the derivatives of φ are continuous, $\tilde{x}'_\delta(s) \to \tilde{x}'_0(s) = 0$ as $\delta \to 0$ for $\hat{\epsilon}_1 < s < \hat{\epsilon}_2$. In fact, the term $\left|\tilde{x}'_\delta(s)\right|$ converges to zero uniformly in $[\hat{\epsilon}_1, \hat{\epsilon}_2]$. To see this, recall that $\tilde{x}'_\delta(s)$ is a partial derivative of φ. Since φ is $\mathcal{C}^1$, $\tilde{x}'_\delta(s)$ is continuous with respect to s and δ. For any $\beta > 0$ and $s \in [\hat{\epsilon}_1, \hat{\epsilon}_2]$, there exists γ_s such that

$$\left|\tilde{x}'_\delta(\tilde{s})\right| < \beta \text{ if } \delta < \gamma_s \text{ and } |\tilde{s} - s| < \gamma_s.$$

The existence of γ such that

$$\left|\tilde{x}'_\delta(s)\right| < \beta \text{ if } \delta < \gamma \text{ and } s \in [\hat{\epsilon}_1, \hat{\epsilon}_2]$$

follows by the compactness of $[\hat{\epsilon}_1, \hat{\epsilon}_2]$. So we may choose δ small enough so that

$$\left(1 - \left|\tilde{x}'_\delta(s)\right|\right) > \frac{M_1 + M_2}{2M_1} \text{ for all } s \in [\hat{\epsilon}_1, \hat{\epsilon}_2].$$

Now, for δ small enough and $i = 1, 2$, we have $g'_{\bar{x}}(\hat{\epsilon}_i) < M_2$, so this gives us calm $\bar{f}_{\hat{\epsilon}_i}(\bar{x}) = g'_{\bar{x}}(\hat{\epsilon}_i) < M_2$ by Proposition 3.2. Therefore, if δ is small enough,

$$\left|\bar{f}_{\hat{\epsilon}_i}\left(\tilde{x}_\delta\left(\hat{\epsilon}_i\right)\right) - \bar{f}_{\hat{\epsilon}_i}\left(\bar{x}\right)\right| \leq M_2 \left|\tilde{x}_\delta\left(\hat{\epsilon}_i\right) - \bar{x}\right|.$$

Recall that if the derivative $g'_{\bar{x}}(\epsilon)$ exists, then $g'_{\bar{x}}(\epsilon) = \text{calm}\,\bar{f}_\epsilon(\bar{x})$ by Proposition 3.2. On the one hand, we have

$$\bar{f}_{\hat{\epsilon}_2}(\bar{x}) - \bar{f}_{\hat{\epsilon}_1}(\bar{x}) = \int_{\hat{\epsilon}_1}^{\hat{\epsilon}_2} g'_{\bar{x}}(s)\, ds \leq \int_{\hat{\epsilon}_1}^{\hat{\epsilon}_2} M_2 ds = M_2\left(\hat{\epsilon}_2 - \hat{\epsilon}_1\right).$$

But on the other hand, $\tilde{x}_\delta(s) \in T_1$ for $0 < \delta < 1$, and so $g'_{\tilde{x}_\delta(s)}(s) \geq M_1$ by Lemma 5.4. If δ is small enough, we have

$$\begin{aligned}
& \left|\bar{f}_{\hat{\epsilon}_2}(\bar{x}) - \bar{f}_{\hat{\epsilon}_1}(\bar{x})\right| \\
\geq\ & \left|\bar{f}_{\hat{\epsilon}_2}\left(\tilde{x}_\delta\left(\hat{\epsilon}_2\right)\right) - \bar{f}_{\hat{\epsilon}_1}\left(\tilde{x}_\delta\left(\hat{\epsilon}_1\right)\right)\right| \\
& \quad - \left(\left|\bar{f}_{\hat{\epsilon}_2}\left(\tilde{x}_\delta\left(\hat{\epsilon}_2\right)\right) - \bar{f}_{\hat{\epsilon}_2}(\bar{x})\right| + \left|\bar{f}_{\hat{\epsilon}_1}\left(\tilde{x}_\delta\left(\hat{\epsilon}_1\right)\right) - \bar{f}_{\hat{\epsilon}_1}(\bar{x})\right|\right) \\
\geq\ & \int_{\hat{\epsilon}_1}^{\hat{\epsilon}_2} \left(1 - \left|\tilde{x}'_\delta(s)\right|\right) g'_{\tilde{x}_\delta(s)}(s)\, ds \\
& \quad - M_2\left(\left|\tilde{x}_\delta\left(\hat{\epsilon}_2\right) - \bar{x}\right| + \left|\tilde{x}_\delta\left(\hat{\epsilon}_1\right) - \bar{x}\right|\right) \\
\geq\ & \int_{\hat{\epsilon}_1}^{\hat{\epsilon}_2} \left(1 - \left|\tilde{x}'_\delta(s)\right|\right) M_1 ds - M_2\left(\left|\tilde{x}_\delta\left(\hat{\epsilon}_2\right) - \bar{x}\right| + \left|\tilde{x}_\delta\left(\hat{\epsilon}_1\right) - \bar{x}\right|\right) \\
\geq\ & \int_{\hat{\epsilon}_1}^{\hat{\epsilon}_2} \left(\frac{M_1 + M_2}{2}\right) ds - M_2\left(\left|\tilde{x}_\delta\left(\hat{\epsilon}_2\right) - \bar{x}\right| + \left|\tilde{x}_\delta\left(\hat{\epsilon}_1\right) - \bar{x}\right|\right) \\
=\ & \left(\frac{M_1 + M_2}{2}\right)\left(\hat{\epsilon}_2 - \hat{\epsilon}_1\right) - M_2\left(\left|\tilde{x}_\delta\left(\hat{\epsilon}_2\right) - \bar{x}\right| + \left|\tilde{x}_\delta\left(\hat{\epsilon}_1\right) - \bar{x}\right|\right).
\end{aligned}$$

As δ is arbitrarily small and the terms $\left|\tilde{x}_\delta\left(\hat{\epsilon}_i\right) - \bar{x}\right| \to 0$ as $\delta \to 0$ for $i = 1, 2$, we have $\left|\bar{f}_{\hat{\epsilon}_2}(\bar{x}) - \bar{f}_{\hat{\epsilon}_1}(\bar{x})\right| \geq \left(\frac{M_1+M_2}{2}\right)\left(\hat{\epsilon}_2 - \hat{\epsilon}_1\right)$. This is a contradiction, and thus we are done. □

Before we prove Lemma 5.7 below, we need to recall the definition of simplicial complexes from [11, Section 3.2.1]. A *simplex* with vertices $a_0, \ldots, a_d$ is

$$\begin{aligned}
\left[a_0, \ldots, a_d\right] &= \{x \in \mathbb{R}^n \mid \exists \lambda_0, \ldots, \lambda_d \in [0, 1], \\
& \qquad \sum_{i=0}^{d} \lambda_i = 1 \text{ and } x = \sum_{i=0}^{d} \lambda_i a_i.\}
\end{aligned}$$

The corresponding *open simplex* is

$$\begin{aligned}
\left(a_0, \ldots, a_d\right) &= \{x \in \mathbb{R}^n \mid \exists \lambda_0, \ldots, \lambda_d \in (0, 1), \\
& \qquad \sum_{i=0}^{d} \lambda_i = 1 \text{ and } x = \sum_{i=0}^{d} \lambda_i a_i.\}
\end{aligned}$$

We shall denote by int (σ) the open simplex corresponding to the simplex σ. A face of the simplex $\sigma = [a_0, \ldots, a_d]$ is a simplex $\tau = [b_0, \ldots, b_e]$ such that

$$\{b_0, \ldots, b_e\} \subset \{a_0, \ldots, a_d\}.$$

A *finite simplicial complex* in $\mathbb{R}^n$ is a finite collection $K = \{\sigma_1, \ldots, \sigma_p\}$ of simplices $\sigma_i \subset \mathbb{R}^n$ such that, for every $\sigma_i, \sigma_j \in K$, the intersection $\sigma_i \cap \sigma_j$ is either empty or is a common face of σ_i and σ_j. We set $|K| = \cup_{\sigma_i \in K} \sigma_i$; this is a semi-algebraic subset of $\mathbb{R}^n$. We recall a result on relating semi-algebraic sets to simplicial complexes.

Theorem 5.5. [11, Theorem 3.12] *Let $S \subset \mathbb{R}^n$ be a compact semi-algebraic set, and $S_1, \ldots, S_p$, semi-algebraic subsets of S. Then there exists a finite simplicial complex K in $\mathbb{R}^n$ and a semi-algebraic homeomorphism $h : |K| \to S$, such that each S_k is the image by h of a union of open simplices of K.*

We need yet another result for the proof of Lemma 5.7.

Proposition 5.6. *Suppose that $\phi : (0,1)^2 \to \mathbb{R}$, not necessarily semi-algebraic, is continuous in $(0,1)^2$. Let* gph $\phi \subset (0,1)^2 \times \mathbb{R}$ *be the graph of ϕ. Then for any $t \in (0,1)$,* cl (gph ϕ) $\cap (0,t) \times \mathbb{R}$ *is either a single point or a connected line segment.*

Proof. Suppose that $((0,t), a_1)$ and $((0,t), a_2)$ lie in cl (gph ϕ). We need to show that for any $\alpha \in (a_1, a_2)$, $((0,t), \alpha)$ lies in cl (gph ϕ).

For any $\epsilon > 0$, we can find points $p_1, p_2 \in (0,1)^2$ such that the points $(p_1, \tilde{a}_1), (p_2, \tilde{a}_2) \in$ gph ϕ are such that $|\tilde{a}_i - a_i| < \epsilon$ and $|p_i - (0,t)| < \epsilon$ for $i = 1, 2$. Recall that by definition $\tilde{a}_i = \phi(p_i)$ for $i = 1, 2$. Choose ϵ such that $\tilde{a}_1 + \epsilon < \tilde{a}_2 - \epsilon$. By the intermediate value theorem, for any $\alpha \in (\tilde{a}_1 + \epsilon, \tilde{a}_2 - \epsilon)$, there exists a point p in the line segment $[p_1, p_2]$ such that $\phi(p) = \alpha$. Moreover, $|p - (0,t)| < \max_{i=1,2} |p_i - (0,t)|$. Letting $\epsilon \to 0$, we see that $((0,t), \alpha) \in$ cl (gph ϕ) as needed. ☐

We now prove our last result important for the proof of Theorem 5.3. The proof of the lemma below is similar to the proof of the Curve Selection Lemma in [11, Theorem 3.13].

Lemma 5.7. *Let $S \subset \mathbb{R}^n$ be a semi-algebraic set, and $\tau : [\epsilon_1, \epsilon_2] \to \mathbb{R}^n$ be a semi-algebraic curve such that $\tau([\epsilon_1, \epsilon_2]) \cap S = \emptyset$ and $\tau([\epsilon_1, \epsilon_2]) \subset$* cl (S)*. Then there exists a function $\varphi : [0,1] \times [\hat{\epsilon}_1, \hat{\epsilon}_2] \to \mathbb{R}^n$, with $[\hat{\epsilon}_1, \hat{\epsilon}_2] \neq \emptyset$ and $[\hat{\epsilon}_1, \hat{\epsilon}_2] \subset [\epsilon_1, \epsilon_2]$, such that*

(1) $\varphi(0, \epsilon) = \tau(\epsilon)$ for $\epsilon \in [\hat{\epsilon}_1, \hat{\epsilon}_2]$ and $\varphi((0,1] \times [\hat{\epsilon}_1, \hat{\epsilon}_2]) \subset S$.

(2) The partial derivative of φ with respect to the second variable, which we denote by $\frac{\partial}{\partial \epsilon}\varphi$, exists and is continuous in $[0,1] \times [\hat{\epsilon}_1, \hat{\epsilon}_2]$.

Proof. Replacing S with its intersection with a closed bounded set containing $\tau([\epsilon_1, \epsilon_2])$, we can assume S is bounded. Then $\mathrm{cl}(S)$ is a compact semi-algebraic set. By Theorem 5.5, there is a finite simplicial complex K and a semi-algebraic homeomorphism $h : |K| \to \mathrm{cl}(S)$, such that S and $\tau([\epsilon_1, \epsilon_2])$ are images by h of a union of open simplices in K . In particular, this means that there is an open interval $(\hat{\epsilon}_1, \hat{\epsilon}_2) \subset [\epsilon_1, \epsilon_2]$ such that $\tau((\hat{\epsilon}_1, \hat{\epsilon}_2))$ is an image by h of a 1-dimensional open simplex in K. Since $h^{-1} \circ \tau((\hat{\epsilon}_1, \hat{\epsilon}_2))$ is in $\mathrm{cl}(S)$ but not in S, there is a simplex σ of K which has $h^{-1} \circ \tau([\hat{\epsilon}_1, \hat{\epsilon}_2])$ lying in the boundary of σ, and $h(\mathrm{int}(\sigma)) \subset S$.

Let $\hat{\sigma}$ be the barycenter of σ. Define the map $\delta : [0,1] \times [\hat{\epsilon}_1, \hat{\epsilon}_2] \to \mathbb{R}^n$ by

$$\delta(t, \epsilon) = (1-t)\, h^{-1} \circ \tau(\epsilon) + t\hat{\sigma}.$$

The map above satisfies $\delta((0,1] \times (\hat{\epsilon}_1, \hat{\epsilon}_2)) \subset \mathrm{int}(\sigma)$. By contracting the interval $[\hat{\epsilon}_1, \hat{\epsilon}_2]$ slightly, $\varphi = h \circ \delta$ satisfies property (1).

By contracting the interval $[\hat{\epsilon}_1, \hat{\epsilon}_2]$ if necessary and applying the decomposition theorem [10, Theorem 6.7], we can assume that φ is $\mathcal{C}^1$ in the set $(0, \bar{t}] \times [\hat{\epsilon}_1, \hat{\epsilon}_2]$ for some $\bar{t} \in (0,1)$.

Since τ is semi-algebraic, we contract the interval $[\hat{\epsilon}_1, \hat{\epsilon}_2]$ again if necessary so that τ is $\mathcal{C}^1$ there. Therefore, $\frac{\partial}{\partial\epsilon}\varphi$ exists in $[0, \bar{t}] \times [\hat{\epsilon}_1, \hat{\epsilon}_2]$. It remains to show that $\frac{\partial}{\partial\epsilon}\varphi$ is continuous in $[0, \bar{t}] \times [\hat{\epsilon}_1, \hat{\epsilon}_2]$. We do this by showing that $\frac{\partial}{\partial\epsilon}\varphi_i : [0, \bar{t}] \times [\hat{\epsilon}_1, \hat{\epsilon}_2] \to \mathbb{R}$, the ith component of the derivative with respect to the second variable, is continuous for each i.

Since $\frac{\partial}{\partial\epsilon}\varphi_i$ is continuous in $(0, \bar{t}] \times [\hat{\epsilon}_1, \hat{\epsilon}_2]$, it remains to show that it is continuous at every point in $\{0\} \times [\hat{\epsilon}_1, \hat{\epsilon}_2]$. The graph of $\frac{\partial}{\partial\epsilon}\varphi_i$ corresponding to the domain $(0, \bar{t}] \times [\hat{\epsilon}_1, \hat{\epsilon}_2]$, which we denote by gph $\left(\frac{\partial}{\partial\epsilon}\varphi_i\right)$, is a subset of $(0, \bar{t}] \times [\hat{\epsilon}_1, \hat{\epsilon}_2] \times \mathbb{R}$. We show that $\left((0,\epsilon), \frac{\partial}{\partial\epsilon}\varphi_i(0,\epsilon)\right) \in \mathrm{cl}\left(\mathrm{gph}\left(\frac{\partial}{\partial\epsilon}\varphi_i\right)\right)$. For small $t_1, t_2 > 0$, consider $\varphi_i(t_1, \epsilon - t_2)$ and $\varphi_i(t_1, \epsilon + t_2)$. By the intermediate value theorem, there is some $\tilde{\epsilon} \in (\epsilon - t_2, \epsilon + t_2)$ such that

$$\frac{\partial}{\partial\epsilon}\varphi_i(t_1, \tilde{\epsilon}) = \frac{1}{2t_2}\left(\varphi_i(t_1, \epsilon + t_2) - \varphi_i(t_1, \epsilon - t_2)\right).$$

If t_2 were chosen such that

$$\left|\frac{1}{2t_2}\left(\varphi_i(0, \epsilon + t_2) - \varphi_i(0, \epsilon - t_2)\right) - \frac{\partial}{\partial\epsilon}\varphi_i(0,\epsilon)\right|$$

is small and t_1 is chosen such that

$$\left|\frac{1}{2t_2}\left(\varphi_i(t_1, \epsilon + t_2) - \varphi_i(t_1, \epsilon - t_2)\right) - \frac{1}{2t_2}\left(\varphi_i(0, \epsilon + t_2) - \varphi_i(0, \epsilon - t_2)\right)\right|$$

is small, then $\left|\frac{\partial}{\partial\epsilon}\varphi_i(t_1, \tilde{\epsilon}) - \frac{\partial}{\partial\epsilon}\varphi_i(0,\epsilon)\right|$ is small. Taking $t_2 \to 0$ and $t_1 \to 0$, we have $\left((0,\epsilon), \frac{\partial}{\partial\epsilon}\varphi_i(0,\epsilon)\right) \in \mathrm{cl}\left(\mathrm{gph}\left(\frac{\partial}{\partial\epsilon}\varphi_i\right)\right)$ as desired.

Recall that the graph gph $\left(\frac{\partial}{\partial\epsilon}\varphi_i\right)$ is taken corresponding to the domain $(0,\bar{t}] \times [\hat{\epsilon}_1, \hat{\epsilon}_2]$, and is a manifold of dimension 2 in $\mathbb{R}^3$. Its boundary is of dimension 1 [11, Proposition 3.16], so the intersection of cl $\left(\text{gph }\left(\frac{\partial}{\partial\epsilon}\varphi_i\right)\right)$ with $\{0\} \times [\hat{\epsilon}_1, \hat{\epsilon}_2] \times \mathbb{R}$ is of dimension 1 as well, and is homeomorphic to a closed line segment. There cannot be an interval $[\tilde{\epsilon}_1, \tilde{\epsilon}_2] \subset [\hat{\epsilon}_1, \hat{\epsilon}_2]$ on which cl $\left(\text{gph }\left(\frac{\partial}{\partial\epsilon}\varphi_i\right)\right) \cap \{0\} \times \{\epsilon\} \times \mathbb{R}$ has more than one value for all $\epsilon \in [\tilde{\epsilon}_1, \tilde{\epsilon}_2]$ because by appealing to Proposition 5.6, this implies that the dimension cannot be 1. We note however that it is possible that there exists an $\bar{\epsilon} \in [\hat{\epsilon}_1, \hat{\epsilon}_2]$ such that cl $\left(\text{gph }\left(\frac{\partial}{\partial\epsilon}\varphi_i\right)\right) \cap \{0\} \times \{\bar{\epsilon}\} \times \mathbb{R}$ is a 1-dimensional line segment. This can only happen for only finitely many $\bar{\epsilon} \in [\hat{\epsilon}_1, \hat{\epsilon}_2]$ due to semi-algebraicity.

In any case, we can contract the interval $[\hat{\epsilon}_1, \hat{\epsilon}_2]$ if necessary so that cl $\left(\text{gph }\left(\frac{\partial}{\partial\epsilon}\varphi_i\right)\right) \cap \{0\} \times \{\epsilon\} \times \mathbb{R}$ is a single point for all $\epsilon \in [\hat{\epsilon}_1, \hat{\epsilon}_2]$. This means that for any $(t, \tilde{\epsilon}) \to (0, \epsilon)$, we have $\frac{\partial}{\partial\epsilon}\varphi_i(t, \tilde{\epsilon}) \to \frac{\partial}{\partial\epsilon}\varphi_i(0, \epsilon)$, establishing the continuity of $\frac{\partial}{\partial\epsilon}\varphi_i(\cdot, \cdot)$ on $[0, \bar{t}] \times [\hat{\epsilon}_1, \hat{\epsilon}_2]$. A reparametrization allows us to assume that $\bar{t} = 1$, and we are done. □

6. Quadratic examples

In this section, we show how the robust regularization can be calculated for quadratic examples, which are more-or-less standard in the spirit of [3, 1]. We write $A \succeq 0$ for a real symmetric matrix A if A is positive semidefinite.

Theorem 6.1. *(Euclidean norm) For any real $m \times n$ matrix A and vector $b \in \mathbb{R}^m$, consider the function $g : \mathbb{R}^n \to \mathbb{R}$ defined by*

$$g(x) = \|Ax + b\|_2,$$

Then the following properties are equivalent for any point $(x, t) \in \mathbb{R}^n \times \mathbb{R}$:

(i) $t \geq \bar{g}_\epsilon(x)$

(ii) *there exists a real μ such that*

$$\begin{bmatrix} tI_m & Ax + b & \epsilon A \\ (Ax+b)^T & t - \mu & 0 \\ \epsilon A^T & 0 & \mu I_n \end{bmatrix} \succeq 0.$$

Proof. Applying [1, Thm 4.5.60] shows $t \geq \bar{g}_\epsilon(x)$ holds if and only if there exist real s and μ satisfying

$$\begin{aligned} t - s &\geq 0 \\ \begin{bmatrix} sI_m & Ax + b & \epsilon A \\ (Ax+b)^T & s - \mu & 0 \\ \epsilon A^T & 0 & \mu I_n \end{bmatrix} &\succeq 0. \end{aligned}$$

and the result now follows immediately. □

Since the matrix in property (ii) above is an affine function of the variables x, t and μ, it follows that the robust regularization $\bar{g}_\epsilon$ is "semidefinite-representable", in the language of [1]. This result allows us to use $\bar{g}_\epsilon$ in building tractable representations of convex optimization problems as semidefinite programs.

An easy consequence of the above result is a representation for the robust regularization of any strictly convex quadratic function.

Corollary 6.2. *(quadratics) For any real positive definite n-by-n matrix H, vector $c \in \mathbb{R}^n$, and scalar d, consider the function $h : \mathbb{R}^n \to \mathbb{R}$ defined by*

$$h(x) = x^T Hx + 2c^T x + d.$$

Then the following properties are equivalent for any point $(x, t) \in \mathbb{R}^n \times \mathbb{R}$:

(i) $t \geq \bar{h}_\epsilon(x)$;

(ii) *there exist reals s and μ such that*

$$\begin{array}{rcl} t - s^2 + c^T H^{-1} c - d & \geq & 0 \\ \begin{bmatrix} sI_n & H^{1/2}x + H^{1/2}c & \epsilon H^{1/2} \\ (H^{1/2}x + H^{-1/2}c)^T & s - \mu & 0 \\ \epsilon H^{1/2} & 0 & \mu I_n \end{bmatrix} & \succeq & 0. \end{array}$$

Proof. Clearly $t \geq \bar{h}_\epsilon(x)$ if and only if

$$\|y - x\|_2 \leq \epsilon \;\;\Rightarrow\;\; \|H^{1/2}y + H^{-1/2}c\|_2^2 \leq t - d + c^T H^{-1} c.$$

This property in turn is equivalent to the existence of a real s satisfying

$$\begin{array}{rcl} s^2 & \leq & t - d + c^T H^{-1} c \text{ and} \\ \|y - x\|_2 \leq \epsilon & \Rightarrow & \|H^{1/2}y + H^{-1/2}c\|_2 \leq s, \end{array}$$

and the result now follows from the preceding theorem. □

Since the quadratic inequality

$$t - s^2 + c^T H^{-1} c - d \geq 0$$

is semidefinite-representable, so is the robust regularization $\bar{h}_\epsilon$.

7. 1-PEACEFUL SETS

In this section, we prove that $X \subset \mathbb{R}^n$ is nearly radial implies X is 1-peaceful using the Mordukhovich Criterion [23, Theorem 9.40], which relates the Lipschitz modulus of set-valued maps to normal cones of its graph. The next section discusses further properties of nearly radial sets and how they are common in analysis.

The Mordukhovich Criterion requires the domain of the set-valued map to be $\mathbb{R}^n$, so we recall the map $\tilde{\Phi}_\epsilon : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ by $\tilde{\Phi}_\epsilon(x) = \mathbb{B}_\epsilon(x) \cap X$. Recall that $\tilde{\Phi}_\epsilon|_X = \Phi_\epsilon$ and $\operatorname{lip} \Phi_\epsilon(x) \leq \operatorname{lip} \tilde{\Phi}_\epsilon(x)$ for all $x \in X$. Let us recall the definitions of normal cones, the Aubin property and the graphical modulus.

Definition 7.1. [23, Definition 6.3] Let $X \subset \mathbb{R}^n$ and $\bar{x} \in X$. A vector v is normal to X at $\bar{x}$ in the regular sense, or a *regular normal*, written $v \in \hat{N}_X(\bar{x})$, if

$$\langle v, x - \bar{x} \rangle \leq o(|x - \bar{x}|) \text{ for } x \in X.$$

It is normal to X at $\bar{x}$ in the general sense, or simply a *normal* vector, written $v \in N_X(\bar{x})$, if there are sequences $x^\nu \underset{X}{\to} \bar{x}$ and $v^\nu \underset{X}{\to} v$ with $v^\nu \in \hat{N}_X(x^\nu)$.

Definition 7.2. [23, Definition 9.36] For $X \subset \mathbb{R}^n$, a mapping $S : X \rightrightarrows \mathbb{R}^m$ has the *Aubin property at $\bar{x}$ for $\bar{u}$*, where $\bar{x} \in X$ and $\bar{u} \in S(\bar{x})$, if $\operatorname{gph} S$ is locally closed at $(\bar{x}, \bar{u})$ and there are neighborhoods V of $\bar{x}$ and W of $\bar{u}$ such that

$$S(x') \cap W \subset S(x) + \kappa |x' - x| \mathbb{B} \text{ for all } x, x' \in X \cap V.$$

The *graphical modulus of S at $\bar{x}$ for $\bar{u}$* is

$$\begin{aligned} \operatorname{lip} S(\bar{x} \mid \bar{u}) \ := \ \inf\{\kappa \mid & \text{ There are neighbourhoods} \\ & V \text{ of } \bar{x}, W \text{ of } \bar{u} \text{ such that} \\ & S(x') \cap W \subset S(x) + \kappa |x' - x| \mathbb{B} \\ & \text{for all } x, x' \in X \cap V\}. \end{aligned}$$

If S is single-valued at $\bar{x}$, then in keeping with the notation of lip in Definition 2.1, we write $\operatorname{lip} S(\bar{x})$ instead of $\operatorname{lip} S(\bar{x} \mid S(\bar{x}))$. Note that this equals $\operatorname{lip} S(\bar{x})$ if S is continuous at $\bar{x}$. $\diamond$

A set-valued map S is *locally compact* around $\bar{x}$ if there exist a neighborhood V of $\bar{x}$ and a compact set $C \subset Y$ such that $S(V) \subset C$. This is equivalent to $S(V)$ being a bounded set, which is the case when S is outer semicontinuous and $S(\bar{x})$ is bounded. If S is outer semicontinuous and locally compact at $\bar{x}$, then by [20, Theorem 1.42], the Lipschitz modulus and the Aubin property are related by

$$\operatorname{lip} S(\bar{x}) = \max_{\bar{u} \in S(\bar{x})} \{\operatorname{lip} S(\bar{x} \mid \bar{u})\}.$$

In finite dimensions, we need $S(\bar{x})$ to be bounded and S to be outer semicontinuous for the formula above to hold.

Here is a lemma on convex cones.

Lemma 7.3. *Given any two convex cones C_1 and C_2 polar to each other and any vector x, we have*

$$(d(x, C_1))^2 + (d(x, C_2))^2 = \|x\|^2$$

Proof. This is a simple consequence of [23, Exercise 12.22] □

We now present our result on the relation between 1-peaceful sets and nearly radial sets.

Theorem 7.4. *If X is nearly radial at $\bar{x}$, then X is 1-peaceful at $\bar{x}$. The converse holds if X is subdifferentially regular for all points in a neighborhood around $\bar{x}$.*

Proof. The graph of $\tilde{\Phi}_\epsilon$ is the intersection of $\mathbb{R}^n \times X$ and the set $D \subset \mathbb{R}^n \times \mathbb{R}^n$ defined by

$$D := \{(x,y) \mid \|x-y\| \le \epsilon\}.$$

By applying a rule on the normal cones of products of sets [23, Proposition 6.41], we infer that $N_{\mathbb{R}^n \times X}(x,y) = \{\mathbf{0}\} \times N_X(y)$. Define the real valued function $g_0 : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}_+$ by $g_0(x,y) := \frac{1}{2}\|x-y\|^2$. Then the gradient of g_0 is $\nabla g_0(x,y) = (x-y, y-x)$.

From this point, we assume that $\|x-y\| = \epsilon$. The normal cone of D at (x,y) is $N_D(x,y) = \mathbb{R}_+\{(x-y,y-x)\}$ using [23, Exercise 6.7]. On applying a rule on the normal cones of intersections [23, Theorem 6.42], we get

$$N_{\text{gph}\,\tilde{\Phi}_\epsilon}(x,y) \subset (\{\mathbf{0}\} \times N_X(y)) + \mathbb{R}_+\{(x-y,y-x)\}. \tag{7.1}$$

Furthermore, if X is subdifferentially regular at y, the above set inclusion is an equation. By the Mordukhovich criterion [23, Theorem 9.40], $\tilde{\Phi}_\epsilon$ has the Aubin Property at (x,y) if and only if the graphical modulus lip $\tilde{\Phi}_\epsilon(x \mid y)$ is finite. It can be calculated by appealing to the formulas for the coderivative D^* [23, Definition 8.33] and outer norm $|\cdot|^+$ [23, Section 9D] below.

$$\begin{aligned}
\text{lip}\,\tilde{\Phi}_\epsilon(x \mid y) &= \left|D^*\tilde{\Phi}_\epsilon(x \mid y)\right|^+ \text{ (by [23, Theorem 9.40])} \\
&= \sup_{w\in\mathbb{B}} \sup_{z\in D^*\tilde{\Phi}_\epsilon(w)} \|z\| \text{ (by [23, Section 9D])} \\
&= \sup\left\{\|z\| \mid (w,z) \in \text{gph}\,D^*\tilde{\Phi}_\epsilon, \|w\| \le 1\right\} \\
&= \sup\left\{\|z\| \mid (-z,w) \in N_{\text{gph}\,\tilde{\Phi}_\epsilon}(x,y), \|w\| \le 1\right\} \\
&\qquad \text{(by [23, Definition 8.33])} \\
&\le \sup\{\|z\| \mid (-z,w) \in (\{\mathbf{0}\} \times N_X(y)) \\
&\qquad\quad +\mathbb{R}_+\{(x-y,y-x)\}, \|w\| \le 1.
\end{aligned} \tag{7.2}$$

We can assume that $z = y-x$ with a rescaling, and $w = y-x+v$ for some $v \in N_X(y)$. Since $(\{\mathbf{0}\} \times N_X(y)) + \mathbb{R}_+\{(x-y,y-x)\}$ is positively homogeneous set, we could find the supremum of $\frac{\|z\|}{\|w\|}$ in the same set and the formula reduces to

$$\begin{aligned}\text{lip}\,\tilde{\Phi}_\epsilon(x \mid y) &\le \sup_{v\in N_X(y)} \frac{\|y-x\|}{\|y-x+v\|} \\ &= \sup_{v\in N_X(y)} \frac{\|x-y\|}{\|(x-y)-v\|} \\ &= \frac{\|x-y\|}{d(x-y,N_X(y))}. \qquad (7.3)\end{aligned}$$

For a fixed $x \neq y$, say $\bar{x}$, we have $1/\text{lip}\,\tilde{\Phi}_\epsilon(\bar{x} \mid y) \ge \frac{d(\bar{x}-y,N_X(y))}{\|\bar{x}-y\|}$. First, we prove that for any open set W about $\bar{x}$, we have

$$\inf_{\substack{y\in W\cap X\\ y\neq\bar{x}}} \frac{d(\bar{x}-y,N_X(y))}{\|\bar{x}-y\|} = \inf_{\substack{y\in W\cap X\\ y\neq\bar{x}}} \frac{d\left(\bar{x}-y,\hat{N}_X(y)\right)}{\|\bar{x}-y\|}. \qquad (7.4)$$

It is clear that "$\le$" holds because $\hat{N}_X(y) \subset N_X(y)$, so we proceed to prove the other inequality. Consider $d(\bar{x}-y,N_X(y))$. Let $v \in P_{N_X(y)}(\bar{x}-y)$, the projection of $(\bar{x}-y)$ onto $N_X(y)$. Then $v \in N_X(y)$, and so there exists $y_i \to y$, with $y_i \in W \cap X$, and $v_i \to v$ such that $v_i \in \hat{N}_X(y_i)$. So

$$\begin{aligned} d(\bar{x}-y,N_X(y)) &= d(\bar{x}-y,\mathbb{R}_+(v)) \\ &= \lim_{i\to\infty} d(\bar{x}-y,\mathbb{R}_+(v_i)) \\ &= \lim_{i\to\infty} d(\bar{x}-y_i,\mathbb{R}_+(v_i)) \\ &\ge \limsup_{i\to\infty} d\left(\bar{x}-y_i,\hat{N}_X(y_i)\right) \\ \Rightarrow \frac{d(\bar{x}-y,N_X(y))}{\|\bar{x}-y\|} &\ge \limsup_{i\to\infty} \frac{d\left(\bar{x}-y_i,\hat{N}_X(y_i)\right)}{\|\bar{x}-y_i\|}. \end{aligned}$$

Thus equation 7.4 holds. Therefore

$$\liminf_{y\to\bar{x}} \frac{d\left(\bar{x}-y,\hat{N}_X(y)\right)}{\|\bar{x}-y\|} \ge 1 \text{ implies } \limsup_{y\to\bar{x}} \text{lip}\,\tilde{\Phi}_{\|\bar{x}-y\|}(\bar{x} \mid y) \le 1,$$

so we may now consider only regular normal cones.

By Lemma 7.3, we deduce the following:

$$d\left(\bar{x}-y,\hat{N}_X(y)\right)^2 + d\left(\bar{x}-y,\hat{N}_X(y)^*\right)^2 = \|\bar{x}-y\|^2 \text{ for } y \in X.$$

Since $T_X(y)^* = \hat{N}_X(y)$ always [23, Theorem 6.28(a)], we apply Lemma 7.3 and get

$$d\left(\bar{x}-y,\hat{N}_X(y)\right)^2 + d\left(\bar{x}-y,T_X(y)^{**}\right)^2 = \|\bar{x}-y\|^2 \text{ for } y\in X.$$

As $T_X(y) \subset T_X(y)^{**}$ [23, Corollary 6.21], this implies that

$$d\left(\bar{x}-y,\hat{N}_X(y)\right)^2 + d\left(\bar{x}-y,T_X(y)\right)^2 \geq \|\bar{x}-y\|^2 \text{ for } y\in X. \tag{7.5}$$

Note that if X is nearly radial at $\bar{x}$, then $\frac{1}{\|\bar{x}-y\|}d\left(\bar{x}-y,T_X(y)\right) \to 0$ as $\epsilon = \|\bar{x}-y\| \downarrow 0$, $y\in X$. This means that

$$1/\text{lip}\,\tilde{\Phi}_{\|\bar{x}-y\|}\left(\bar{x}\mid y\right) \geq \frac{1}{\|\bar{x}-y\|}d\left(\bar{x}-y,\hat{N}_X(y)\right) \to 1,$$

so

$$\limsup_{y\xrightarrow[X]{}\bar{x},y\neq\bar{x}} \text{lip}\,\tilde{\Phi}_{\|\bar{x}-y\|}\left(\bar{x}\mid y\right) \leq 1,$$

where $y \xrightarrow[X]{} \bar{x}$ means $y\in X$ and $y\to\bar{x}$.

Recall that $\tilde{\Phi}_\epsilon$ has closed graph, and hence it is outer semicontinuous [23, Theorem 5.7(a)]. It is also locally bounded, so

$$\text{lip}\,\tilde{\Phi}_\epsilon(\bar{x}) = \max_{y\in S_\epsilon(\bar{x})} \text{lip}\,\tilde{\Phi}_\epsilon\left(\bar{x}\mid y\right)$$

by [20, Theorem 1.42]. This gives us $\limsup_{\epsilon\to 0}\text{lip}\,\tilde{\Phi}_\epsilon(\bar{x}) \leq 1$, or X is 1-peaceful at $\bar{x}$, as needed.

If we assume that X is regular in a neighborhood of $\bar{x}$, then Formula (7.5) is an equation. Furthermore, (7.1), (7.2) and (7.3) are all equations. Thus if $\lim_{\epsilon\to 0}\text{lip}\,\tilde{\Phi}_\epsilon(\bar{x}) = 1$, then

$$\frac{1}{\|\bar{x}-y\|}d\left(\bar{x}-y,\hat{N}_X(y)\right) = 1/\text{lip}\,\tilde{\Phi}_{\|\bar{x}-y\|}\left(\bar{x}\mid y\right) \to 1 \text{ as } y\xrightarrow[X]{}\bar{x},\ y\neq\bar{x}.$$

and we have $\frac{1}{\|\bar{x}-y\|}d\left(\bar{x}-y,T_X(y)\right) \to 0$ as $y\xrightarrow[X]{}\bar{x}$ and $y\neq\bar{x}$, which means that X is nearly radial at $\bar{x}$. □

Finally, $1-$peaceful sets are interesting in robust regularization for another reason. The Lipschitz modulus of the robust regularization over 1-peaceful sets have Lipschitz modulus bounded above by that of the original function, as the following result shows.

Proposition 7.5. *If X is 1-peaceful and $F: X\to\mathbb{R}^n$ is locally Lipschitz at $\bar{x}$, then*

$$\limsup_{\epsilon\to 0}\text{lip}\,F_\epsilon(\bar{x}) \leq \text{lip}\,F(\bar{x}).$$

Proof. We use a set-valued chain rule [23, Exercise 10.39]. Recall the formula $F_\epsilon = \left(F \circ \tilde{\Phi}_\epsilon\right)|_X$. The mapping $(x,u) \mapsto \tilde{\Phi}_\epsilon(x) \cap F^{-1}(u)$ is locally bounded because the map $x \mapsto \tilde{\Phi}_\epsilon(x)$ is locally bounded. Thus

$$\text{lip}\, F_\epsilon(\bar{x}) \leq \text{lip}\, \tilde{\Phi}_\epsilon(\bar{x}) \cdot \max_{x \in \tilde{\Phi}_\epsilon(\bar{x})} \text{lip}\, F(x).$$

By Theorem 7.4, $\lim_{\epsilon\to 0} \text{lip}\, \tilde{\Phi}_\epsilon(\bar{x}) \leq 1$. Also, since lip $F : \mathbb{R}^n \to \mathbb{R}_+$ is upper semicontinuous, $\limsup_{\epsilon\to 0} \max_{x\in\tilde{\Phi}_\epsilon(\bar{x})} \text{lip}\, F(x) \leq \text{lip}\, F(\bar{x})$. Taking limits to both sides gives us what we need. □

8. NEARLY RADIAL SETS

As highlighted in Section 7, nearly radial sets are 1-peaceful. In this section, we study the properties of nearly radial sets and give examples of nearly radial sets to illustrate their abundance in analysis.

We contrast the definition of nearly radial sets given before Proposition 4.5 with a stronger property introduced by [25], which is the uniform version of the same idea. This idea was called $o(1)$-*convexity* in [25].

Definition 8.1. (nearly convex sets) A set $X \subset \mathbb{R}^n$ is *nearly convex* at a point $\bar{x} \in X$ if

$$\text{dist}\,(y, x + T_X(x)) = o(\|x - y\|) \text{ as } x, y \to \bar{x} \text{ in } X$$

The set X is *nearly convex* if it is nearly convex at every point X. ◇

Clearly if a set is nearly convex at a point, then it is nearly radial there, but the class of nearly radial sets is considerably broader. For example, the set

$$X = \{x \in \mathbb{R}^2 : x_1 x_2 = 0\}$$

is nearly radial at the origin but not nearly convex there, since as $n \to \infty$ the points $x_n = (n^{-1}, 0)$ and $y_n = (0, n^{-1})$ approach the origin in X and yet

$$\text{dist}(y_n, x_n + T_X(x_n)) = n^{-1} \neq o(\|x_n - y_n\|).$$

It is immediate that convex sets are nearly convex, and hence nearly radial. A straightforward exercise shows that smooth manifolds are also nearly convex, and hence again nearly radial. These observations are both special cases of the following result, rather analogous to [25, Theorem 2.2]. A set $X \subset \mathbb{R}^n$ is *amenable* [23, Section 10F] at a point $\bar{x} \in X$ if there is an open neighborhood V of $\bar{x}$, a $\mathcal{C}^1$ mapping $F : V \to \mathbb{R}^m$, and a closed convex set $D \subset \mathbb{R}^m$, such that

$$X \cap V = \{x \in V : F(x) \in D\}$$
$$\text{and} \quad N_D(F(\bar{x})) \cap N(\nabla F(\bar{x})^*) = \{\mathbf{0}\}, \tag{8.6}$$

where $N_D(\cdot)$ denotes the normal cone to D, and $N(\cdot)$ denotes null space. If in fact F is $\mathcal{C}^2$ then we call X *strongly amenable* [23, Definition 10.23] at $\bar{x}$.

Theorem 8.2. *(amenable implies nearly radial) Suppose the set $X \subset \mathbb{R}^n$ is amenable at the point $\bar{x} \in X$. Then X is nearly convex (and hence nearly radial) at $\bar{x}$.*

Proof. Since X is amenable at $\bar{x}$, we can suppose property (8.6) holds. Suppose without loss of generality $\bar{x} = \mathbf{0}$, and consider a sequences of points $x_r, y_r \to \mathbf{0}$ in the set $X \cap V$. We want to show

$$\text{dist}(y_r, x_r + T_X(x_r)) = o(\|x_r - y_r\|).$$

Without loss of generality we can suppose $x_r \neq y_r$ for all r, and denote the unit vectors $\|x_r - y_r\|^{-1}(x_r - y_r)$ by z_r. We want to prove

$$d_r = \min\{\|w + z_r\| : w \in T_X(x_r)\} \to 0.$$

The unique minimizer $w_r \in T_X(x_r)$ in the above projection problem satisfies

$$\begin{aligned} d_r &= \|w_r + z_r\| \\ w_r + z_r &\in -N_X(x_r) = -\nabla F(x_r)^* N_D(F(x_r)) \\ \langle w_r, w_r + z_r \rangle &= 0, \end{aligned}$$

by [23, Exercise 10.26(d)]. Choose vectors $u_r \in -N_D(F(x_r))$ such that

$$w_r + z_r = \nabla F(x_r)^* u_r.$$

We next observe that the sequence of vectors $\{u_r\}$ is bounded. Otherwise, we could choose a subsequence $\{u_{r'}\}$ satisfying $\|u_{r'}\| \to \infty$, and then any limit point of the sequence of unit vectors $\{\|u_{r'}\|^{-1}u_{r'}\}$ must lie in the set $-N_D(F(\mathbf{0})) \cap N(\nabla F(\mathbf{0})^*)$, contradicting property (8.6).

We now have

$$\begin{aligned} 0 \;\leq\; d_r^2 \;&=\; \langle z_r, \nabla F(x_r)^* u_r \rangle \;=\; \langle \nabla F(x_r) z_r, u_r \rangle \\ &= \; \left\langle \nabla F(x_r) z_r - \|x_r - y_r\|^{-1}[F(x_r) - F(y_r)], u_r \right\rangle \\ &\qquad + \left\langle \|x_r - y_r\|^{-1}[F(x_r) - F(y_r)], u_r \right\rangle. \end{aligned}$$

The first term converges to zero, using the smoothness of the mapping F and the boundedness of the sequence $\{u_r\}$. On the other hand, since the set D is convex, we have $F(y_r) - F(x_r) \in T_D(F(x_r))$, and $u_r \in -N_D(F(x_r))$ by assumption, so the second term is nonpositive, and the result follows. □

It is worth comparing these notions to a property that is slightly stronger still: *prox-regularity* (in the terminology of [23, Section 13F]), or $O(2)$-*convexity* [25].

Definition 8.3. (prox-regular sets) A set $X \subset \mathbb{R}^n$ is *prox-regular* at a point $\bar{x} \in X$ if

$$\text{dist}\,(y, x + T_X\,(x)) = O\left(\|x - y\|^2\right) \text{ as } x, y \to \bar{x} \text{ in } X. \diamond$$

Theorem 8.2 (amenable implies nearly radial) is analogous to the fact that strong amenability implies prox-regularity [23, Proposition 13.32] (and also to [25, Proposition 2.3]).

The class of nearly radial sets is very broad, as the following easy result (which fails for nearly convex sets) emphasizes.

Proposition 8.4. *(unions) If the sets $X_1, X_2, \ldots, X_n$ are each nearly radial at the point $\bar{x} \in \cap_j X_j$, then so is the union $\cup_j X_j$.*

Proof. If the result fails, there is a sequence of points $x_r \to \bar{x}$ in $\cup_j X_j$ and real $\epsilon > 0$ such that

$$\text{dist}\left(\frac{\bar{x} - x_r}{\|\bar{x} - x_r\|}, T_{\cup_j X_j}(x_r)\right) \geq \epsilon \text{ for all } r. \tag{8.7}$$

By taking a subsequence, we can suppose that there is an index i such that $x_r \in X_i$ for all r. But then we know

$$\text{dist}\left(\frac{\bar{x} - x_r}{\|\bar{x} - x_r\|}, T_{X_i}(x_r)\right) \to 0,$$

which contradicts inequality (8.7), since $T_{X_i}(x_r) \subset T_{\cup_j X_j}(x_r)$. □

A key concept in variational analysis is the idea of Clarke regularity (see for example [8, 9, 23]). We make no essential use of this concept in our development, but it is worth remarking on the relationship (or lack of it) between the nearly radial property and Clarke regularity. Note first that nearly radial sets need not be Clarke regular: the union of the two coordinate axes in $\mathbb{R}^2$ is nearly radial at the origin, for example, but it is not Clarke regular there.

On the other hand, Clarke regular sets need not be nearly radial.

Example 8.5. Consider the function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} 2^{-n} - 2^{-n-1}(2 - 2^{n+1}|x|)^{1+2^{-n}} & \text{if } 2^{-n-1} \leq |x| \leq 2^{-n} \ (n \in \mathbb{N}) \\ 0 & \text{if } x = 0. \end{cases}$$

The function f is even, and its graph consists of concave segments on each interval $x \in [2^{-n-1}, 2^{-n}]$, passing through the point $2^{-n}(1,1)$ with left derivative zero, and through the point $2^{-n-1}(1,1)$ with right derivative $1+2^{-n}$. A routine calculation now shows that this function is everywhere regular, and hence its epigraph epi f is everywhere Clarke regular. However, epi f

is not nearly radial at the origin. To see this, observe that for each $n \in \mathbb{N}$, if we consider the sequence $x_n = 2^{-n}(1,1) \to (0,0)$, then we have

$$T_{\text{epi}\,f}(x_n) = \left\{(x,y) : y \geq (1+2^{1-n}) \max\{x,0\}\right\},$$

so

$$\text{dist}(0, x_n + T_{\text{epi}\,f}(x_n)) = \frac{\|x_n\|}{\sqrt{2}},$$

contradicting the definition of a nearly radial set. $\diamond$

This is yet another attractive property for semi-algebraic sets.

Theorem 8.6. *(semi-algebraic sets) Semi-algebraic sets are nearly radial.*

Proof. Suppose the origin lies in a semi-algebraic set $X \subset \mathbb{R}^n$. We will show that X is nearly radial at the origin.

If the result fails, then there is a real $\delta > 0$ and a sequence of points $y_r \to \mathbf{0}$ in X such that

$$\left\| u + \frac{y_r}{\|y_r\|} \right\| > \delta \text{ for all } u \in T_X(y_r).$$

Hence for each index r there exists a real $\gamma_r > 0$ such that

$$\left\| \frac{z-y_r}{\|z-y_r\|} + \frac{y_r}{\|y_r\|} \right\| > \delta \text{ for all } z \in X \text{ such that } 0 < \|z-y_r\| < \gamma_r.$$

Consequently, each point y_r lies in the set

$$\begin{aligned} X_0 \quad = \quad & \Big\{ y \in X \mid \exists \gamma > 0 \text{ so } \left\| \frac{z-y}{\|z-y\|} + \frac{y}{\|y\|} \right\| > \delta \\ & \qquad \forall z \in X \setminus \{y\} \text{ with } \|z-y\| < \gamma \Big\}, \end{aligned}$$

so $\mathbf{0} \in \text{cl}\, X_0$.

By quantifier elimination (see for example the discussion of the Tarski-Seidenberg Theorem in [2, p. 62]), the set X_0 is semi-algebraic. Hence the Curve Selection Lemma (see [2, p. 98] and [19]) shows that there is a real-analytic path $p : [0,1] \to \mathbb{R}^n$ such that $p(0) = \mathbf{0}$ and $p(t) \in X_0$ for all $t \in (0,1]$. For some positive integer k and nonzero vector $g \in \mathbb{R}^n$ we have, for small $t > 0$,

$$\begin{aligned} p(t) &= gt^k + O(t^{k+1}) \\ p'(t) &= kgt^{k-1} + O(t^k), \end{aligned}$$

and in particular both $p(t)$ and $p'(t)$ are nonzero. For any such t we know

$$\left\| \frac{z-p(t)}{\|z-p(t)\|} + \frac{p(t)}{\|p(t)\|} \right\| > \delta$$

for any point $z \in X \setminus \{p(t)\}$ close to $p(t)$. Hence for any real $s \neq t$ close to t we have

$$\left\| \frac{p(s) - p(t)}{\|p(s) - p(t)\|} + \frac{p(t)}{\|p(t)\|} \right\| > \delta.$$

Taking the limit as $s \uparrow t$ shows

$$\left\| \frac{p(t)}{\|p(t)\|} - \frac{p'(t)}{\|p'(t)\|} \right\| \geq \delta$$

for all small $t > 0$. But since

$$\lim_{t \downarrow 0} \frac{p(t)}{\|p(t)\|} = \frac{g}{\|g\|} = \lim_{t \downarrow 0} \frac{p'(t)}{\|p'(t)\|},$$

this is a contradiction. □

By contrast, semi-algebraic sets need not be nearly convex. For example, the union of the two coordinate axes in $\mathbb{R}^2$ is semi-algebraic, but it is not nearly convex at the origin.

Acknowledgement. Thanks to Jim Renegar for helpful discussions concerning Theorem 8.6 (semi-algebraic sets), and to Jon Borwein for the statement of the definition of nearly radial sets before Proposition 4.5. The first author's research was supported in part by U.S. NSF grant DMS-080657.

REFERENCES

[1] A. Ben-Tal and A. Nemirovski. Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications. SIAM, Philadelphia, 2001.

[2] R. Benedetti and J.-J. Risler. Real Algebraic and Semi-algebraic Sets. Hermann, Paris, 1990.

[3] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. Linear Matrix Inequalities in System and Control Theory. SIAM, Philadelphia, 1994.

[4] J.V. Burke, A.S. Lewis, and M.L. Overton. Optimal stability and eigenvalue multiplicity. Foundations of Computational Mathematics, 1:205-225, 2001.

[5] J.V. Burke, A.S. Lewis, and M.L. Overton. Optimization over pseudospectra. SIAM Journal on Matrix Analysis 25:80-104,2003.

[6] J.V. Burke, A.S. Lewis and M.L. Overton, Robust stability and a criss-cross algorithm for pseudospectra IMA Journal of Numerical Analysis 23 (2003), pp. 359-375.

[7] F.H. Clarke, A new approach to Lagrange multipliers, Mathematics of Operations Research, 1, 165-174.

[8] F.H. Clarke. Optimization and Nonsmooth Analysis. Wiley, New York, 1983. Republished as Vol. 5, Classics in Applied Mathematics, SIAM, 1990.

[9] F.H. Clarke, Yu.S. Ledyaev, R.J. Stern, and P.R. Wolenski. Nonsmooth Analysis and Control Theory. Springer-Verlag, New York, 1998.

[10] M. Coste, An Introduction to O-minimal geometry,1999.

[11] M. Coste, An Introduction to Semialgebraic Geometry, 2002.

[12] A. Daniilidis and P. Georgiev, Approximate Convexity and Submonotonicity, J. Math. Anal. Appl., 291 (2004), 292-301.

[13] L. El Ghaoui and S.-I. Niculescu. Robust decision problems in engineering: a linear matrix inequality approach. In L. El Ghaoui and S.-I. Niculescu, editors, Advances in Linear Matrix Inequality Methods in Control, pages 3-37. SIAM, Philadelphia, 2000.

[14] J.-B. Hiriart-Urruty. The deconvolution operation in convex analysis: an introduction. Cybernetics Systems Analysis, 30:555-560, 1994.

[15] A. S. Lewis, Robust Regularization, preprint, 2002.

[16] A. S. Lewis and C. H. J. Pang, Variational analysis of pseudospectra, 2008, to appear in SIAM Journal on Optimization

[17] W. Li, Sharp Lipschitz constants for basic optimal solutions and basic feasible solutions of linear programs, SIAM J. Control Optim., 32 (1994), pp. 140?153.

[18] E. Mengi and M.L. Overton. Algorithms for the computation of the pseudospectral radius and the numerical radius of a matrix, IMA Journal of Numerical Analysis 25 (2005) pp. 648-669.

[19] J.W. Milnor. Singular Points of Complex Hypersurfaces. Princeton University Press, Princeton, N.J., 1968.

[20] B.S. Mordukhovich, Variational Analysis and Generalized Differentiation I and II. Springer, Berlin, 2006.

[21] S.M. Robinson, Solution continuity in monotone affine variational inequalities, SIAM J. Optim. Volume 18, Issue 3, pp. 1046-1060 (2007)

[22] R.T. Rockafellar. Favorable classes of Lipschitz continuous functions in subgradient optimization in nondifferentiable optimization (1982), E. Nurminski (eds.), Pergamon Press, New York.

[23] R.T. Rockafellar and R.J.-B. Wets. Variational Analysis. Springer, Berlin, 1998.

[24] B. Rustem and M. Howe, Algorithms for Worst-Case Design and Applications to Risk Management, Princeton, NJ, 2003.

[25] A. Shapiro. Existence and differentiability of metric projections in Hilbert space. SIAM Journal on Optimization, 4:130-141, 1994.

[26] A. Shapiro and F. Al-Khayyal. First-order conditions for isolated locally optimal solutions. Journal of Optimization Theory and Applications, 77:189-196, 1993.

[27] L.N. Trefethen and M. Embree, Spectra and Pseudospectra, Princeton, 2006.

ADRIAN S. LEWIS, SCHOOL OF OPERATIONS RESEARCH AND INFORMATION ENGINEERING, CORNELL UNIVERSITY, ITHACA, NY, 14853.

E-mail address: `aslewis@orie.cornell.edu`

C. H. JEFFREY PANG, CENTER FOR APPLIED MATHEMATICS, CORNELL UNIVERSITY, ITHACA, NY, 14853.

E-mail address: `cp229@cornell.edu`