A continuous derivative for real-valued functions

Abbas Edalat

Department of Computing, Imperial College London, UK

Abstract

We develop a notion of derivative of a real-valued function on a Banach space, called the L-derivative, which is constructed by introducing a generalization of Lipschitz constant of a map. As with the Clarke gradient, the values of the L-derivative of a function are non-empty weak* compact and convex subsets of the dual of the Banach space. The L-derivative, however, is shown to be upper semi continuous, a result which is not known to hold for the Clarke gradient. We also formulate the notion of primitive maps dual to the L-derivative, an extension of Fundamental Theorem of Calculus for the L-derivative and a domain for computation of real-valued functions on a Banach space with a corresponding notion of effectivity. For real-valued functions on finite dimensional Euclidean spaces, the L-derivative can be obtained within an effectively given continuous domain. We also show that in finite dimensions the L-derivative and the Clarke gradient coincide thus providing a computable representation for the latter in this case.

This paper is dedicated to the historical memory of Sharaf al-din Tusi (d. 1213), the Iranian mathematician who was the first to use the derivative systematically to solve for roots of cubic polynomials and find their maxima.

1 Introduction

The notion of derivative of functions has been the key fundamental concept in the advent and development of differential calculus and is at the basis of some of the most crucial branches of mathematics including ordinary and partial differential equations, dynamical systems, mathematical physics, differential geometry and differential topology. These comprise what is often referred to as continuous mathematics, one of the two main branches of mathematics, with discrete mathematics as the other distinguished branch.

The first systematic use of the derivative of functions was undertaken by the Iranian mathematician Sharaf al-din Tusi (d. 1213) who introduced a technique, algebraically equivalent to what we now call the Ruffini-Horner method, for finding the roots of cubic polynomials by an iterative process using the derivative of the polynomial [14]. Although he never put a name to it in Arabic, which like Latin later on in Europe was the language of scholarship in the Muslim world, he

1

also used the derivative to find the maxima of polynomials, which until recently historians of mathematics had attributed to the 16th century French mathematician François Viète [15, 19]. Sharaf al-din Tusi who died only six years before the cataclysmic Mongol invasion of Iran in 1219 is now considered to be the forerunner of algebraic geometry.

Nearly two centuries after the ground breaking work of Newton and Leibniz on the foundation of differential calculus in the 17th century, modern mathematical analysis was born with the introduction of mathematical limit by Cauchy in the nineteenth century, which provided a precise notion for the existence of the derivative of a Function at a point. This led to new and surprising results about the derivative. In 1872, based on what was by then a rigorous mathematical framework, Weierstrass constructed a continuous function which was nowhere differentiable.

In the early 20th century, the French mathematicians *Gâteaux* and *Fréchet* extended the notion of derivative in two distinct ways to functions of infinite dimensional Banach spaces. These higher dimensional derivatives have now applications in quantum field theory, but, like the classical derivative of a real-valued function of a single variable, they may not exist and when they do exist they may not give rise to continuous functions. A comprehensive modern account of the various notions of derivative in topological linear spaces is given in [27].

In 1980's, Frank Clarke, motivated by problems in non-smooth analysis and control theory, introduced the notion of generalized gradient of a function, which is now named after him [4]. Clarke's gradient of a locally Lipschitz real-valued function on a Banach space always exists and is a set-valued function: on finite dimensional Euclidean spaces it takes non-empty compact and convex subsets of the Euclidean space as its values and the gradient is upper semi-continuous. On an infinite dimensional Banach space, the Clarke gradient is a non-empty weak* compact and convex subset of the dual of the Banach space. It is however not known if Clarke's gradient is also upper semi-continuous on infinite dimensional Banach spaces [3].

A few decades earlier, following the seminal work of Alan Turing [24, 25] and the advent of computer science in 1930's, computable analysis took shape in 1950's with the work of Grzegorczyk [17, 18]. A fundamental thesis established in the subject is that a computable function is necessarily a continuous function [21, 26]. Indeed, if a function is to be computed at a real number, which is given as the limit of a sequence of rational numbers, then the continuity of the function is required to be able to compute the value of the function as the limit of its values at the elements of the sequence.

Since the derivative of functions plays a fundamental role in mathematics, one would expect a real interest in a notion of derivative which is always continuous in computability theory. However, surprisingly, no attempt was made to develop a continuous derivative for functions and the work of Clarke went unnoticed by researchers in computable analysis, who have only worked with the classical derivative of functions.

A new approach to differential calculus based on mathematical structures in computer science, called domains [6, 2, 16] was introduced in [9, 10] first for real-valued functions of a real variable and then for multivariable functions. The motivation here has arisen from computer science and computable analysis to formulate and use, in particular, a notion of continuous derivative for functions.

In the domain-theoretic framework, a continuous derivative for functions, a corresponding notion of primitive maps, an extension of fundamental theorem of calculus and a domain for differentiable functions have been developed. These have led to data types for presenting differentiable functions and solving ordinary differential equations [7, 11], a constructive version of the inverse and implicit function theorems [12] and a denotational semantics for hybrid systems [13].

The concept of derivative of a real-valued function which was developed in [10] depends, somewhat unsatisfactorily, on the choice of the coordinate system used. In fact, the value of the derivative of a locally Lipschitz real valued function on a finite Euclidean space turns out to be the smallest hyperrectangle, with edges parallel to the given coordinate axes, containing Clarke's gradient.

In this paper, inspired by the above domain-theoretic framework, we introduce a coordinate free approach to develop the notion of the *L-derivative* of a real-valued function on a Banach space; it is constructed by formulating a generalized Lipschitz property of functions. The local generalized Lipschitz properties of the function, which provide finitary information about the rate of growth of the function in local neighbourhoods, are used to define the L-derivative of the function globally. Like the Clarke gradient, the values of the L-derivative are non-empty weak* compact and convex subsets of the dual of the Banach space.

The L-derivative, developed here from the local to the global and from the discrete to the continuum, is shown to be upper semi-continuous for real-valued locally Lipschitz functions on any Banach space, a result which is not known for the Clarke gradient as we have already mentioned above.

For a C^1 function, i.e., one with a continuous Fréchet derivative, the L-derivative and the Fréchet derivative coincide. More generally, when the function fails to be C^1, the L-derivative contains the Clarke gradient, and also the Gâteaux and the Fréchet derivatives, whenever the latter two exist.

The L-derivative gives rise to an extension of the Fundamental Theorem of Calculus. The class of functions from the Banach space into the collection of non-empty weak* compact and convex subsets of the dual of the Banach space, which are generated by step functions, is dual via the L-derivative to families of real-valued locally Lipschitz functions on the Banach space. The L-derivative is also employed to construct a domain of computation for real-valued functions on Banach spaces which carries an effective structure when the space is separable. These results extend those for finite dimensions in [9, 10].

For functions on finite Euclidean spaces, the L-derivative is an element of a countably based continuous domain which can be given an effective structure that characterizes computable functions with computable L-derivatives. Any continuous function and its L-derivative can be obtained as the supremum of an increasing

sequence of pairs of finitary and consistent information about the function and its L-derivative.

Although they are defined using very different techniques, we show here that in finite dimensions the Clarke gradient and the L-derivative coincide. Thus, in finite dimensions the construction of the L-derivative provides a new computable representation for the Clarke gradient.

1.1 Background definitions

For the remainder of this section we will present the basic background definitions of the various notions of derivative which we will need in this paper.

Let X and Y be Banach spaces and let $U \subset X$ be an open subset. We recall that the (one sided) directional derivative of $f : U \to Y$ at $x \in U$ in the direction $v \in X$ is

$$F'(x; v) = \lim_{t \downarrow 0} \frac{f(x + tv) - f(x)}{t},$$

if the limit exists. If the above directional derivative exists for all $v \in X$, then $D(f)(x) : X \to Y$ with $D(f)(x)(v) := F'(x; v)$ is the *Gâteaux* derivative of f at x if $D(f)(x)$ is a bounded linear map [27, 20].

The *Fréchet* derivative [27] of a map $f : U \to Y$ at $x \in U$, when it exists, is a bounded linear map $T : X \to Y$ with

$$\lim_{\|x-y\| \to 0} \frac{\|f(x) - f(y) - T(x - y)\|}{\|x - y\|} = 0.$$

The linear map T is denoted by $f'(x)$. When the Fréchet derivative exists at x, so does the Gâteaux derivative and they are equal. However, the *Fréchet* derivative at x can fail to exist even if the Gâteaux derivative exists at x and is a bounded linear map.

From now on we will assume that $Y = \mathbb{R}$. We next aim to define the generalized (Clarke) gradient of a function [4, Chapter two] and explain its properties. Let $f : U \to \mathbb{R}$ be Lipschitz near $x \in U$ and $v \in X$. The *generalized directional derivative* of f at x in the direction of v is

$$f^\circ(x; v) = \limsup_{\substack{y \to x \\ t \downarrow 0}} \frac{f(y + tv) - f(y)}{t}.$$

Let us denote by X^* the dual of X, i.e. the set of real-valued continuous linear functions on X. Unless otherwise stated we will consider X^* with its weak* topology. Recall that the weak* topology is the weakest topology on X^* in which for any $x \in X$ the map $f \mapsto f(x) : X^* \to \mathbb{R}$ is continuous.

The *generalized gradient* of f at x, denoted by $\partial f(x)$ is the subset of X^* given by

$$\{A \in X^* : f^\circ(x; v) \geq A(v) \text{ for all } v \in X\}.$$

It is shown in [4, page 27] that

- $\partial f(x)$ is a non-empty, convex, weak* compact subset of X^*.

- For $v \in X$, we have:

$$f^\circ(x; v) = \max\{A(v) : A \in \partial f(x)\}.$$

There is an alternative characterization of the generalized gradient when X is finite dimensional, say $X = \mathbb{R}^n$. In this case, by Rademacher's theorem [5, p 148], a locally Lipschitz map $f : U \to \mathbb{R}$ is Fréchet differentiable almost everywhere with respect to the Lebesgue measure. If Ω_f is the nullset where f fails to be differentiable then:

$$\partial f(x) = \mathrm{Co}\{\lim f'(x_i) : x_i \to x, \ x_i \notin \Omega_f\}, \qquad (1)$$

where $\mathrm{Co}(S)$ is the convex hull of a subset $S \subset \mathbb{R}^n$ [4, page 63]. The above expression is interpreted as follows. Consider all sequences $(x_i)_{i \geq 0}$, with $x_i \notin \Omega_f$, for $i \geq 0$, which converge to x such that the limit $f'(x_i)$ exists. Then the generalized gradient is the convex hull of all such limits. Note that, in the above definition, since f is locally Lipschitz at x, it is differentiable almost everywhere in a neighbourhood of x and thus there are plenty of sequences $(x_m)_{m \geq 0}$ such that $\lim_{m \to \infty} x_m = x$ and $\lim_{m \to \infty} f'(x_m)$ exists.

Recall that for a Hausdorff space Z, we can define three topologies on the set of non-empty compact subsets of Z as follows. The *upper topology* has as a base the collection of subsets of the form $\Box O = \{C : C \subseteq O\}$, whereas the *lower topology* has as a subbase the collection of subsets of the form $\Diamond O = \{C : C \cap O \neq \emptyset\}$, where $O \subset Z$ is an open subset. The *Vietoris topology* is the refinement of the upper and lower topologies and is Hausdorff and, when Z is a metric space, it is equivalent to the topology induced by the *Hausdorff metric* d_h defined by $d_h(A, B) = \max(d(A, B), d(B, A))$, where for compact sets C and D, $d(C, D)$ is the infimum of positive numbers δ such that C is contained in the δ-*parallel body* of D defined as $D_\delta = \{x \in Z \mid \exists y \in D. \, d(x, y) \leq \delta\}$; see [23, page 737]. We write $\mathbf{U}(Z)$, $\mathbf{L}(Z)$ and $\mathbf{V}(Z)$, respectively, for the three topological spaces or *hyperspaces*, called respectively the *upper space*, the *lower space* and the *Vietoris space* of Z, obtained by considering, respectively, the upper topology, the lower topology and the Vietoris topology on the set of non-empty compact subsets of Z. The upper space and the lower space are non-Hausdorff.

In finite dimensions, the Clarke gradient is *upper semi-continuous*, i.e. it is continuous with respect to the *upper topology* on the space of the non-empty compact subsets of $\mathbb{R}^n$. It is not known if a similar result holds in infinite dimensions [3], i.e. if the Clarke gradient is continuous with respect to the upper topology on the space of non-empty weak* compact subsets of X^*.

For $X = \mathbb{R}^n$, we let ∇f denote the classical gradient of f, when it exists, i.e.,

$$(\nabla f)_i(x) = \frac{\partial f}{\partial x_i} =$$

$$\lim_{x_i' \to x_i} \frac{f(x_1, \ldots, x_i, \ldots, x_n) - f(x_1, \ldots, x_i', \ldots, x_n)}{x_i - x_i'},$$

for $1 \le i \le n$. Recall that, in finite dimensions, if the (Fréchet) derivative exists at a point then the gradient also exists at that point and is the same linear map.

We also recall that for a function $f : U \to \mathbb{R}$, where U is an open subset of $\mathbb{R}^n$, Dini's lower and upper partial derivatives, for $1 \le i \le n$, are defined respectively as

$$(\nabla f)_i^l(x) = \tag{2}$$

$$\liminf_{x_i' \to x_i} \frac{f(x_1, \ldots, x_i, \ldots, x_n) - f(x_1, \ldots, x_i', \ldots, x_n)}{x_i - x_i'}$$

$$(\nabla f)_i^u(x) = \tag{3}$$

$$\limsup_{x_i' \to x_i} \frac{f(x_1, \ldots, x_i, \ldots, x_n) - f(x_1, \ldots, x_i', \ldots, x_n)}{x_i - x_i'}.$$

Note that the Dini's lower and upper partial derivatives always exist as extended real numbers.

2 Some properties related to the dual of a Banach space

Let X be a Banach space. For an open subset $U \subset X$, let $U \to \mathbb{R}$ be the set of all continuous functions of type $U \to \mathbb{R}$ with respect to the norm topology on X. For $x \in X$ and $f \in X^*$, we write $f(x)$ for the real number obtained by the action of f on x and $x(f)$ for the same real number when x is considered as a linear functional on X^*.

The operator norm on X^* extends pointwise to an interval valued map on the weak* compact subsets of X^*. Note that a weak* compact subset is bounded with respect to the operator norm. If b is a non-empty weak* compact and convex subset of X^* then $\|b\| = \{\|\lambda\| : \lambda \in b\}$ with $\|b\| = [\|b\|^-, \|b\|^+]$ is a compact real interval. In particular, for $X = \mathbb{R}$, if $b \subset \mathbb{R}$ is a compact interval, then so also is $|b| = \{|r| : r \in b\}$.

We will consider the extension of the action of bounded linear operators on X (i.e. the mapping $X^* \times X \to \mathbb{R}$ given by $(f, x) \mapsto f(x)$) to the three hyperspaces to obtain three maps $E_U : \mathbf{U}(X^*) \times X \to \mathbf{U}(\mathbb{R})$, $E_L : \mathbf{L}(X^*) \times X \to \mathbf{L}(\mathbb{R})$ and $E_V : \mathbf{V}(X^*) \times X \to \mathbf{V}(\mathbb{R})$ which are defined with respect to the different three topologies but have the same action given by $(b, x) \mapsto \{f(x) : f \in b\}$. We write $b(x) = \{f(x) : f \in b\}$ which is a compact subset as b is weak* compact.

Proposition 2.1 *The three maps E_U, E_L and E_V are each continuous separately in their two arguments with respect to the norm topology on X.*

Proof To prove the continuity of the three maps when the second argument is fixed, let $x \in X$. It is sufficient to show that, for any open set $I \subset \mathbb{R}$, the preimage

of $\square I$ is open in $\mathbf{U}(X^*)$ and the preimage of $\Diamond I$ is open in $\mathbf{L}(X^*)$. First we prove that the preimage of $\square I$ is open in $\mathbf{U}(X^*)$. Let b be a non-empty weak* compact subset of X^* with $b(x) \subset I$. Let $\epsilon > 0$ be such that $(b(x))_\epsilon \subset I$. For any $f \in b$, the weak* open set $O(x, f, \epsilon) = \{g \in X^* : |f(x) - g(x)| < \epsilon\}$ gives an open neighbourhood of f. Put $O(x, \epsilon) = \bigcup_{f \in b} O(x, f, \epsilon)$. Then, we have $b \subset O(x, \epsilon)$. If c is a weak* compact set with $c \subset O(x, \epsilon)$, then by compactness there exist a finite number of functions $f_i \in b$ ($i = 1, \cdots, n$) such that $c \subset \bigcup_{1 \le i \le n} O(x, f_i, \epsilon)$. Then for any $g \in c$, there exists $i \in \{1, \cdots, n\}$ such that $f_i \in b$ with $|f_i(x) - g(x)| < \epsilon$ and thus $g(x) \in I$. It follows that $c(x) \subset I$ and therefore the preimage of $\square I$ is open in $\mathbf{U}(X^*)$. Next we consider the preimage of $\Diamond I$. Let b be a non-empty weak* compact subset of X^* with $b(x) \cap I \ne \emptyset$. Let $f \in b$ with $f(x) \in I$ take any $\epsilon > 0$ such that the ϵ open ball centred at $f(x)$ is contained in I. Then, $b \cap O(x, f, \epsilon) \ne \emptyset$ and for any weak* compact subset c of X^* with $c \cap O(x, f, \epsilon) \ne \emptyset$ we have $c(x) \cap I \ne \emptyset$, which shows that the pre-image of $\Diamond I$ is open in $\mathbf{L}(X^*)$. Finally, we prove the continuity of the three maps when the first argument b, say, is fixed. Let $b(x) \in O$ where O is either $\square I$ and $\Diamond I$ for any open subset $I \subset \mathbb{R}$. Since any weak* compact subset of X^* is bounded with respect the operator norm, $\|b\| \le K$ for some $K > 0$. Hence, for any given $\epsilon > 0$ and any $f \in b$, the relation $\|x - y\| < \epsilon/K$ implies $|f(x) - f(y)| \le \|f\|\|x - y\| \le K\epsilon/K = \epsilon$. From this property, the result follows easily. $\square$

As usual, we consider the upper space $\mathbf{U}(Z)$ of any Hausdorff space Z partially ordered with reverse inclusion so that $\mathbf{U}(Z)$ becomes a dcpo; we also include in this dcpo a least element represented by Z. Thus, the map $x \mapsto \{x\} : Z \to \mathbf{U}(Z)$ is a topological embedding onto the set of maximal elements of $\mathbf{U}(Z)$. We identify the input and output of this embedding and write $\{x\}$ simply as x.

We also recall that the *Scott* topology on any dcpo has as open sets those sets O which are upper sets (that is $x \in O$ and $x \sqsubseteq y$ implies $y \in O$) and which are inaccessible by directed sets, i.e. if $\sup_{i \in I} a_i \in O$ for a directed set $(a_i)_{i \in I}$ then there exists $i \in I$ such that $a_i \in O$ [22]. A function $f : D \to E$ of dcpo's D and E is continuous with respect to the Scott topologies on D and E iff it is monotone ($x \sqsubseteq y$ implies $f(x) \sqsubseteq f(y)$) and preserves the lubs of directed subsets, i.e. for any directed set $(a_i)_{i \in I}$ in D we have: $\sup_{i \in I} f(a_i) = f(\sup_{i \in I} a_i)$.

We then have the following:

Proposition 2.2 *[6, Propositions 3.1(iii) and 3.3]*

(i) *For any Hausdorff space Z, the Scott topology on $\mathbf{U}(Z)$ refines the upper topology.*

(ii) *If Z is locally compact then $\mathbf{U}(Z)$ is a continuous dcpo, on which the Scott topology and the upper topology coincide.* $\square$

Consider the poset, denoted by $\mathbf{C}(X^*)$, consisting of X^* and its nonempty weak* compact and convex subsets partially ordered by reverse inclusion so that it has

the least element $\perp = X^*$. Note that $\mathbf{C}(X^*)$ is a bounded complete dcpo and a sub-dcpo of the upper space $\mathbf{U}(X^*)$ of X^*.

When $X = \mathbb{R}$, we consider the sub-dcpo of $\mathbf{U}(\mathbb{R})$ denoted by $\mathbf{I}\mathbb{R}$ of all non-empty compact intervals of $\mathbb{R}$ ordered by reverse inclusion; it is a countably based bounded complete continuous domain. We now restrict the first component of G in Proposition 2.1 to convex subsets so that the range of G will become non-empty compact intervals. Restricting to continuity with respect to the Scott topology, Proposition 2.1 reduces to:

Corollary 2.3 *The map $G : \mathbf{C}(X^*) \times X \to \mathbf{I}\mathbb{R}$ is continuous separately in its two arguments with respect to the Scott topology on $\mathbf{C}(X^*)$ and $\mathbf{I}\mathbb{R}$ and the norm topology on X. In particular, for any directed set $b_i \in \mathbf{C}(X^*)$, $i \in I$ and $v \in X$, we have $\bigcap_{i \in I}(b_i(v)) = (\bigcap_{i \in I} b_i)(v)$. $\square$*

The following result plays a crucial role in the construction of the L-derivative.

Theorem 2.4 *Let S and T be disjoint non-empty convex subsets of the dual X^* of a Banach space X such that, with respect to the weak* topology, S is closed and T is compact. Then there exists a hyperplane in X^* induced by an element of X which separates S and T, i.e. there exist $x \in X$ and $c \in \mathbb{R}$ such that $x(f) < c$ for $f \in S$ and $x(f) > c$ for $f \in T$.*

Proof Since S and T are disjoint closed sets, for each $\lambda \in T$ there exists $\epsilon > 0$ and a finite number of elements $x_1, \cdots x_n \in X$ such that the open neighbourhood of λ defined by $\{\alpha : |x_j(\alpha) - x_j(\lambda)| < \epsilon$ for $1 \leq j \leq n\}$, is disjoint from S. By compactness of T we can find a finite number of functionals, say, $\lambda_i \in X^*$ $(1 \leq i \leq m)$ and elements $x_{i1}, x_{i2}, \cdots x_{in_i} \in X$ and $\epsilon_i > 0$ for $1 \leq i \leq m$ such that the open subsets

$$N_i = \{\alpha : |x_{ij}(\alpha) - x_{ij}(\lambda_i)| < \epsilon_i, \text{ for } 1 \leq j \leq n_i\},$$

for $1 \leq i \leq m$, are disjoint from S and cover T. Define $\Phi : X^* \to \mathbb{R}^{n_1 + n_2 + \cdots + n_m}$ by:

$$\alpha \mapsto$$

$$(x_{11}(\alpha), \cdots, x_{1n_1}(\alpha); \cdots; x_{i1}(\alpha), \cdots, x_{in_i}(\alpha); \cdots; x_{m1}(\alpha), \cdots, x_{mn_m}(\alpha)).$$

Then $\Phi(S)$ and $\Phi(T)$ are convex subsets of $\mathbb{R}^{n_1 + n_2 + \cdots, + n_m}$ and $\Phi(T)$ is compact. Let $P_i : \mathbb{R}^{n_1} \times \cdots \times \mathbb{R}^{n_i} \times \cdots \times \mathbb{R}^{n_m} \to \mathbb{R}^{n_i}$ be the projection onto $\mathbb{R}^{n_i}$. Note that $\Phi(T)$ is contained in the open subset of $\mathbb{R}^{n_1 + \cdots + n_m}$ given by the intersection of the m infinite open strips $\mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \times \cdots \times \mathbb{R}^{n_{i-1}} \times R_i \times \mathbb{R}^{n_{i+1}} \times \cdots \times \mathbb{R}^{n_m}$ where

$$R_i = \{z : \|z - P_i(\Phi(x_i))\| < \epsilon_i\},$$

with $\|v\|$ being the max norm of $v \in \mathbb{R}^{n_i}$. Therefore, $\Phi(T)$ is disjoint from the closure of $\Phi(S)$. Note that for any two disjoint convex subsets of $\mathbb{R}^k$, with one closed and the other compact, there exists a hyperplane in $\mathbb{R}^k$ which separates

them (the normal to such a hyperplane is given by the line through two boundary points of the sets which give the closest distance of the two sets). Thus, there is a hyperplane $\sum_{1 \leq i \leq m} \sum_{1 \leq j \leq n_i} c_{ij} z_{ij} = c$ in $\mathbb{R}^{n_1 + n_2 + \cdots + n_m}$, for some real numbers $c, c_{ij} \in \mathbb{R}$ ($1 \leq i \leq m$ and $1 \leq j \leq n_i$), which separates $\Phi(T)$ from the closure of $\Phi(S)$. Let $x = \sum_{1 \leq i \leq m} \sum_{1 \leq j \leq n_i} c_{ij} x_{ij} \in X$. It follows that the hyperplane $x(\alpha) = c$, where $\alpha \in X^*$, separates S and T in X^*. $\square$

3   Ties of functions

The local differential property of a function is formalized in our framework by the notion of an interval Lipschitz constant. Assume $U \subset X$ is an open subset of a Banach space X.

Definition 3.1 Let f be a real-valued function with domain $\operatorname{dom}(f) \subset U$. We say that $f : \operatorname{dom}(f) \to \mathbb{R}$ has *an interval Lipschitz constant* $b \in \mathbf{C}(X^*)$ in a convex open subset $a \subset \operatorname{dom}(f)$ if for all $x, y \in a$ we have: $b(x - y) \sqsubseteq f(x) - f(y)$. The *single tie* $\delta(a, b)$ of a with b is the collection of all real-valued partial functions f on U with $a \subset \operatorname{dom}(f) \subset U$ which have an interval Lipschitz constant b in a. We call a the *domain* of the single tie.

Since a single tie provides a local Lipschitz property for a family of functions, it is sufficient in Definition 3.1 to restrict the domain of a single tie to a convex open subset. As an example, if $X = \mathbb{R}^2$ and b is the compact rectangle $b_1 \times b_2$ (with compact intervals $b_1, b_2 \subset \mathbb{R}$), the information relation above reduces to:

$$b_1(x_1 - y_1) + b_2(x_2 - y_2) \sqsubseteq f(x) - f(y).$$

Lemma 3.2 *For $b \in \mathbf{C}(X^*)$ and $z \in X$, we have $|b(z)|^+ \leq \|b\|^+ \|z\|$.*

Proof We have $|b(z)|^+ = |\{f(z) : f \in b\}|^+ = \{|f(z)| : f \in b\}^+ \leq \{\|f\| : f \in b\}^+ \|z\| = \|b\|^+ \|z\|$. $\square$

Proposition 3.3 *If $f \in \delta(a, b)$ for $a \neq \emptyset$ and $b \neq \bot$, then $f : a \to \mathbb{R}$ is Lipschitz: for all $x, y \in a$ we have $|f(x) - f(y| \leq \|b\|^+ \|x - y\|$.*

Proof Suppose $f \in \delta(a, b)$ and $x, y \in a$. It follows from $f(x) - f(y) \sqsupseteq b(x - y)$ that $|f(x) - f(y)| \leq \|b\|^+ \|x - y\|$. $\square$

For any topological space Z and any dcpo D with bottom $\bot$, let $Z \to D$ be the dcpo of Scott continuous functions from Z to D. The domain of $f : Z \to D$ is defined as $\operatorname{dom}(f) = \{x : f(x) \neq \bot\}$. In particular, for any open subset $a \subset Z$ and any non-bottom $b \in D$, the *single-step function* $a \searrow b : Z \to D$, with $(a \searrow b)(x) = b$ if $x \in a$ and $(a \searrow b)(x) = \bot$ if $x \notin a$, is Scott continuous and has domain a. A *step function* is then the supremum of any finite set of consistent single-step functions. In the sequel, we consider the dcpo $U \to \mathbf{C}(X^*)$ of Scott continuous functions with $U \subset X$ equipped with its the norm topology.

9

The following proposition justifies our definition of the interval Lipschitz constant. Let a be a convex open subset of X.

Proposition 3.4 *If $f : a \to \mathbb{R}$ is $C^1(a)$ i.e., f is Fréchet differentiable and $f' : a \to X^*$ is continuous, then the following three conditions are equivalent: (i) $f \in \delta(a, b)$, (ii) $\forall z \in a.\ f'(z) \in b$ and (iii) $a \searrow b \sqsubseteq f'$.*

Proof (i) $\Rightarrow$ (ii). Suppose, for the sake of a contradiction that for some $z \in a$, we have $L := f'(z) \notin b$. By Theorem 2.4, there exists a unit vector $s \in X$ and $c \in \mathbb{R}$ such that $(s(L))^+ < c$ and $(s(b))^- > c$. From $f \in \delta(a, b)$, we obtain for sufficiently small h that $\frac{f(z+hs)-f(z)}{h} \in s(b)$. But by Fréchet differentiability at z we have:
$$\lim_{h \to 0} \left| \frac{f(z+hs) - f(z)}{h} - s(L) \right| = 0,$$
which is a contradiction.

(ii) $\Rightarrow$ (i). Assume $x, y \in a$. Then, since the convex set a contains the straight line from x to y, by the mean value theorem there exists $z \in a$ such that $f(x) - f(y) = f'(z)(x - y) \in b(x - y)$.

(iii) $\iff$ (ii). Obvious. $\square$

Note that the convexity of the domain of a single tie is crucial in establishing the equivalence in Proposition 3.4.

We will now show that ties have a dual property in relation to step functions of type $U \to \mathbf{C}(X^*)$.

Proposition 3.5 *Suppose $a \neq \emptyset$ and $b \neq \perp$. We have $\delta(a, b) \supseteq \delta(c, d)$ iff $c \supseteq a$ and $b \sqsubseteq d$.*

Proof The "if" part follows easily from the definition of $\delta(a, b)$. To show the "only if" part, we take any $f \in \delta(c, d)$ such that $\mathrm{dom}(f) = c$. Then, since $f \in \delta(a, b)$, we have $a \subset \mathrm{dom}(f) = c$. On the other hand if $b \sqsubseteq d$ does not hold, take $\gamma \in d \setminus b$ and consider the function $f : c \to \mathbb{R}$ with $f(x) = \gamma(x)$. Then, $f \in \delta(c, d) \setminus \delta(a, b)$. $\square$

Corollary 3.6 *Suppose $a, c \neq \emptyset$ and $b, d \neq \perp$. We have $\delta(a, b) = \delta(c, d) \iff a = c\ \&\ b = d$. Furthermore $\delta(a, b) \supseteq \delta(c, d)$ iff $a \searrow b \sqsubseteq c \searrow d$.* $\square$

For the rest of this section, we assume we are in an infinite dimensional Banach space or in the finite dimensional space $\mathbb{R}^n$ with $n \geq 2$. The case $n = 1$ is completely covered in [9].

Definition 3.7 A *tie* of partial real-valued functions on U is any intersection $\Delta = \bigcap_{i \in I} \delta(a_i, b_i)$, for an arbitrary indexing set I. The *domain* of a non-empty tie Δ is defined as $\mathrm{dom}(\Delta) = \bigcup_{i \in I} \{a_i \mid b_i \neq \perp\}$.

If a non-empty tie is given by the intersection of a finite number of single ties, then it gives us a family of functions with a *finite* set of consistent differential properties. Generally, a non-empty tie gives a family of functions with a consistent set of differential properties.

Similar to Proposition 3.3, we have the following result. Recall that a function $f : U \to \mathbb{R}$ defined on the open set $U \subseteq X$ is *locally Lipschitz* if it is Lipschitz in a neighbourhood of any point in U.

Proposition 3.8 *If Δ is a tie and $f \in \Delta$, then f is locally Lipschitz on $dom(\Delta)$.*

Proof Let $x \in \text{dom}(\Delta)$. Then there exists a tie $\delta(a, b)$ with $x \in a$ and $f \in \Delta \subseteq \delta(a, b)$ and the result follows from Proposition 3.3 $\square$.

We now collect some fundamental properties of ties, which we will use later. The next proposition, whose proof uses Theorem 2.4, is the key technical result for the development of our theory.

Proposition 3.9 *For any indexing set I, the family of step functions $(a_i \searrow b_i)_{i \in I}$ is consistent if $\bigcap_{i \in I} \delta(a_i, b_i) \neq \emptyset$.*

Proof Suppose $f \in \bigcap_{i \in I} \delta(a_i, b_i)$. We will show that every finite subfamily of $(a_i \searrow b_i)_{i \in I}$ is consistent, from which the result follows as $\mathbf{C}(X^*)$ is bounded complete. It suffices to prove that for any finite subset $J \subseteq I$ we have $\bigcap_{j \in J} b_j \neq \emptyset$ if $\bigcap_{j \in J} a_j \neq \emptyset$. This we will show by induction on the cardinality $|J|$ of J. For $|J| = 1$, there is nothing to prove. Suppose now $|J| > 1$ and $\bigcap_{j \in J} a_j \neq \emptyset$. Let $k \in J$. Then by the inductive hypothesis $\bigcap_{j \in J \setminus \{k\}} b_j \neq \emptyset$. If $\bigcap_{j \in J} b_j = \emptyset$, then by Theorem 2.4 there exists a vector $z \in X$ and $c \in \mathbb{R}$ such that such that the disjoint non-empty compact convex sets $\bigcap_{j \in J \setminus \{k\}} b_j$ and b_k are on the opposite sides of the affine space $\{\lambda : z(\lambda) = c\}$. Take elements $x, y \in \bigcap_{j \in J} a_j$ such that $x - y = lz$ for some $l > 0$. It follows that the two intervals $(\bigcap_{j \in J \setminus \{k\}} b_j)(x - y)$ and $b_k(x - y)$ are disjoint. But by our assumption that $f \in \bigcap_{i \in I} \delta(a_i, b_i) \subseteq \bigcap_{j \in J} \delta(a_j, b_j)$, we have $b_k(x - y) \sqsubseteq f(x) - f(y)$ and $b_j(x - y) \sqsubseteq f(x) - f(y)$ for $j \neq k$, which implies $(\bigcap_{j \in J \setminus \{k\}} b_j)(x - y) \sqsubseteq f(x) - f(y)$, a contradiction. $\square$

Corollary 3.10 *The family $(a_i \searrow b_i)_{i \in I}$ is consistent if for any finite subfamily $J \subseteq I$ we have $\bigcap_{i \in J} \delta(a_i, b_i) \neq \emptyset$.* $\square$

Proposition 3.11 *If $a \searrow b \sqsubseteq \sup_{i \in I} a_i \searrow b_i$, then $\delta(a, b) \supseteq \bigcap_{i \in I} \delta(a_i, b_i)$.*

Proof Let $b \sqsubseteq \bigcap_{a_i \supseteq a} b_i$ and assume $f \in \bigcap_{i \in I} \delta(a_i, b_i)$. Let $x, y \in a$. For each $i \in I$ with $a_i \supseteq a$ we have:

$$b_i(x - y) \sqsubseteq f(x) - f(y).$$

Therefore, we get:

$$b(x - y) \sqsubseteq \bigcap_{a_i \supseteq a} b_i(x - y) \sqsubseteq f(x) - f(y),$$

as required. $\square$

Corollary 3.12 *If* $\sup_{i \in I} a_i \searrow b_i \sqsubseteq \sup_{i \in J} a_i \searrow b_i$, *then* $\bigcap_{i \in I} \delta(a_i, b_i) \supseteq \bigcap_{i \in J} \delta(a_i, b_i)$. $\square$

Let $(\mathbf{T}(U), \supseteq)$ be the partial order of ties of $U \to X$ ordered by reverse inclusion.

Proposition 3.13 $(\mathbf{T}(U) \setminus \{\emptyset\}, \supseteq)$ *is a dcpo.*

Proof Suppose $(\Delta_j)_{j \in J}$ is a directed set in $(\mathbf{T}(U) \setminus \{\emptyset\}$ with respect to the partial order $\supseteq$, i.e. $\Delta_{j_1} \cap \Delta_{j_2} \neq \emptyset$ for $j_1, j_2 \in J$. Let $\Delta_j = \bigcap_{i \in I_j} \delta(a_i, b_i)$, where we assume $I_{j_1} \cap I_{j_2} = \emptyset$ for $j_1 \neq j_2$. Consider the collection $(\delta(a_i, b_i))_{i \in \bigcup_{j \in J} I_j}$. By Corollary 3.10, it suffices to show that any finite subfamily of this collection has non-empty intersection. Suppose $i_t \in \bigcup_{j \in J} I_j$ for $1 \leq t \leq n$. Then $\delta(a_{i_t}, b_{i_t}) \in \Delta_{j_t}$ for some $j_t \in J$ ($1 \leq t \leq n$). By assumption $\bigcap_{1 \leq t \leq n} \Delta_{j_t} \neq \emptyset$. Hence, $\bigcap_{1 \leq t \leq n} \delta(a_{i_t}, b_{i_t}) \supseteq \bigcap_{1 \leq t \leq n} \Delta_{j_t} \neq \emptyset$. $\square$

For any topological space Z and any dcpo D, let $Z \to_s D$ be the subset of $Z \to D$ consisting of Scott continuous functions which are supremums of step functions, i.e., $f = \sup_{i \in I} a_i \searrow b_i$ for a family $(a_i \searrow b_i)_{i \in I}$ of step functions with a_i an open subset of Z and $b_i \in D$.

Consider $U \to_s \mathbf{C}(X^*)$. Since any open set $a \subset X$ is the union of open balls, we can assume without loss of generality that the open subsets a_i ($i \in I$) in the expression for f above are convex. It is easy to check that $U \to_s \mathbf{C}(X^*)$ is a dcpo.

We now show that, for any Banach space X, the set of maximal elements of $U \to_s \mathbf{C}(X^*)$ contains the set of functions of type $U \to X^*$, which are continuous with respect to the norm topology on U and X^*. Recall that a metric space is separable if it has a countable dense subset.

Proposition 3.14
  (i) *If $f : U \to X^*$ is continuous with respect to the norm topologies on U and X^*, then $f \in U \to_s \mathbf{C}(X^*)$. Moreover, if X is separable with a countable dense subset $P \subset X$, then f is the lub of single step functions of the form $a \searrow b$ where a is an open ball centred at a point of P with rational radius whereas b is a closed ball centred at a point of P with a rational radius.*

  (ii) *If $f : U \to \mathbb{R}$ is continuous with respect to the norm topology on U, then $f \in U \to_s \mathbf{IR}$. Moreover, if X is separable with a countable dense subset $P \subset X$, then f is the lub of single step functions of the form $a \searrow b$ where a is an open ball centred at a point of P with rational radius whereas b is a rational compact interval.*

Proof (i) By continuity of f, for $x \in U$ and an open ball $B_\epsilon(f(x))$ of radius ϵ around $f(x)$, there exists an open neighbourhood a of x such that $f[a] \subset B_\epsilon(f(x))$. Since the closed ball $\overline{B_\epsilon(f(x))}$ is weak* compact by Alaoglu's theorem, it follows that $a \searrow \overline{B_\epsilon(f(x))} \sqsubseteq f$. Since $\epsilon > 0$ is arbitrary and $\bigcap_{\epsilon > 0} \overline{B_\epsilon(f(x))} = f(x)$, we conclude that f is the supremum of step functions below it. It is easy to check that the second statement also holds.

  (ii) This is proved similar to (i). $\square$

We are finally in a position to define the L-primitives of a Scott continuous function; in fact now we can do more and define:

Definition 3.15 The *L-primitive map* $\int : (U \to_s \mathbf{C}(X^*)) \to \mathbf{T}(U)$ is defined by

$$\int f = \bigcap_{a \searrow b \sqsubseteq f} \delta(a, b).$$

We call $\int f$ the *L-primitives* of f. The following result depends crucially on the fact that the domain of the L-primitive map is defined to be $U \to_s \mathbf{C}(X^*)$ rather than the bigger function space $U \to \mathbf{C}(X^*)$.

Proposition 3.16 *If* $f = \sup_{i \in I} a_i \searrow b_i$, *then* $\int f = \bigcap \delta(a_i, b_i)$.

Proof This follows easily from Corollary 3.12.

The above property leads us to believe that $U \to_s \mathbf{C}(X^*)$, respectively $Z \to D$, may have wider applications in Banach space theory, respectively abstract domain theory, beyond this paper.

Proposition 3.17 *The L-primitive map is continuous and onto the set of non-empty tie.*

Proof Clearly the primitive map is monotone. Let $(g_i)_{i \in I}$ be a directed set in $U \to_s \mathbf{C}(X^*)$ with $g_i = \sup_{j \in I_i} a_j \searrow b_j$. Then,

$$\int \sup_{i \in I} g_i = \int \sup_{i \in I} \sup_{j \in I_i} a_j \searrow b_j =$$

$$\bigcap_{i \in I} \bigcap_{j \in I_i} \delta(a_j, b_j) = \bigcap_{i \in I} \int g_i = \sup_{i \in I} \int g_i.$$

By Proposition 3.9, any non-empty tie is the L-primitive of some element. $\square$

If $X = \mathbb{R}^n$, for $n \geq 2$ or if X is infinite dimensional, the L-primitive map will have the empty tie in its range, a situation which does not occur for $n = 1$. This is similar to the situation in classical analysis in which a continuous vector field in $\mathbb{R}^n$ for $n > 1$ may not be an exact differential.

Example 3.18 Let $g \in \mathbb{R}^2 \to \mathbf{C}(R^2)$ be the maximal function given by $g(x, y) = (g_1(x, y), g_2(x, y))$ with $g_1(x, y) = 1$ and $g_2(x, y) = x$. Then $\frac{\partial g_1}{\partial y} = 0 \neq 1 = \frac{\partial g_2}{\partial x}$, and it will follow as in classical analysis that $\int g = \emptyset$.

4 The L-derivative

Given a Scott continuous function $f : U \to \mathbb{R}$, the relation $f \in \delta(a, b)$ provides, as we have seen, finitary information about the local interval Lipschitz properties of f. By collecting all such local information, we obtain the complete differential properties of f, namely its L-derivative.

Definition 4.1 The *L-derivative* of a continuous function $f : U \to \mathbb{R}$ is the map

$$\mathcal{L}f : U \to \mathbf{C}(X^*),$$

given by

$$\mathcal{L}f = \sup_{f \in \delta(a,b)} a \searrow b.$$

Theorem 4.2   (i) *The L-derivative is well-defined and Scott continuous.*

(ii) *If $f \in C^1(U)$ then $\mathcal{L}f = f'$.*

(iii) *$f \in \delta(a, b)$ iff $a \searrow b \sqsubseteq \mathcal{L}f$.*

Proof (i) Let the indexing set I be defined by $i \in I \iff f \in \delta(a_i, b_i)$. Then $\bigcap_{i \in I} \delta(a_i, b_i) \neq \emptyset$. Hence, $(a_i \searrow b_i)_{i \in I}$ is consistent by Proposition 3.9. Therefore, $\mathcal{L}f = \sup_{i \in I} a_i \searrow b_i$ is well-defined and is Scott continuous.

(ii) By Proposition 3.4, $f \in \delta(a, b) \iff a \searrow b \sqsubseteq f'$. Hence, $f' \sqsupseteq \sup_{f \in \delta(a,b)} a \searrow b$. To show equality, let $z \in U$ and put $L := f'(z)$. By the continuity of the Fréchet derivative $f' : U \to X^*$ at z, for each integer $n > 0$, there exists an open ball $a \subset U$ with $z \in a$ such that $f'(x) \in B_{1/n}(L)$ for $x \in a$, where $B_r(L)$ is the open ball of radius r and centre $L \in X^*$. By Proposition 3.4, we have $f' \sqsupseteq \mathcal{L}f \sqsupseteq a \searrow \overline{B_{1/n}(L)}$, where $\overline{B_r(L)}$ is the closed ball centred at L with radius r, which is convex and weak* compact by Alaoglu's Theorem. Since $\bigcap_{n \geq 0} \overline{B_{1/n}(L)} = f'(z)$ we conclude that $f' = \mathcal{L}f$.

(iii) Obvious. $\square$

Since the Scott topology refines the upper topology on $\mathbf{C}(X^*)$, we also obtain:

Corollary 4.3 *The L-derivative of any continuous function $X \to \mathbb{R}$ is upper semi-continuous.* $\square$

We now obtain the generalization of Theorem 4.2(iii) to ties, which provides a duality between the L-derivative and L-primitives and can be considered as a general version of the Fundamental Theorem of Calculus.

Theorem 4.4 (Fundamental Theorem of Calculus) *For any $g \in U \to_s \mathbf{C}(X^*)$,*

$$f \in \int g \iff g \sqsubseteq \mathcal{L}f.$$

14

Proof Let $g \in U \to_s \mathbf{C}(X^*)$. Then by Theorem 4.2(iii):

$$f \in \int g \iff f \in \bigcap_{a \searrow b \sqsubseteq g} \delta(a,b) \iff$$

$$a \searrow b \sqsubseteq \mathcal{L}f \text{ iff } a \searrow b \sqsubseteq g \iff g \sqsubseteq \mathcal{L}f. \quad \square$$

We will now show that the Gâteaux derivative, if it exists, is always in the L-derivative.

Lemma 4.5 *For any locally Lipschitz map $f : U \to \mathbb{R}$ and any $x, v \in X$ we have:*

$$(\mathcal{L}f(x))(v) = \bigcap\{b(v) : f \in \delta(a,b), x \in a\}.$$

Proof This follows immediately from Corollary 2.3. $\square$

Lemma 4.6 *Let $U \subset X$, $x \in U$ and $f : U \to \mathbb{R}$ be locally Lipschitz. Then, for any $v \in X$.*

$$\limsup_{y \to x \; t \downarrow 0} \frac{f(y + tv) - f(y)}{t} \leq (\mathcal{L}f(v))^+,$$

$$\liminf_{y \to x \; t \downarrow 0} \frac{f(y + tv) - f(y)}{t} \geq (\mathcal{L}f(v))^-,$$

Proof If $f \in \delta(a,b)$ with $x \in a$ then for y sufficiently close to x and $t > 0$ sufficiently small we have: $f(y + tv) - f(y) \in tb(v)$ and thus

$$\limsup_{y \to x \; t \downarrow 0} \frac{f(y + tv) - f(y)}{t} \leq (b(v))^+$$

which implies

$$\limsup_{y \to x \; t \downarrow 0} \frac{f(y + tv) - f(y)}{t} \leq \inf\{(b(v))^+ : f \in \delta(a,b), x \in a\}.$$

Since $\mathcal{L}f(x) = \bigcap\{b : f \in \delta(a,b), x \in a\}$, the proof of the first inequality follows from Lemma 4.5. The second inequality is proved in a similar way. $\square$

Corollary 4.7 *The Gâteaux derivative of f at x, when it exists, belongs to the L-derivative. Similarly for the Fréchet derivative.* $\square$

In order to obtain the next corollary we first need the following characterization of the generalized gradient.

Lemma 4.8 *For any locally Lipschitz function f we have: $A \in \partial f(x)$ iff for all $v \in X$,*

$$\liminf_{y \to x \; t \downarrow 0} \frac{f(y + tv) - f(y)}{t} \leq A(v) \leq \limsup_{y \to x \; t \downarrow 0} \frac{f(y + tv) - f(y)}{t}$$

15

Proof The "if" part follows by definition. For the "only if" part, the second inequality is just the definition of the generalized gradient. For the first inequality, assume $A \in \partial f(x)$ and $v \in X$. Then, by the definition of the generalized gradient, with v replaced by $-v$ we have:

$$-A(v) \leq \limsup_{\substack{y \to x \\ t \downarrow 0}} \frac{f(y - tv) - f(y)}{t}$$

or

$$A(v) \geq \liminf_{\substack{y \to x \\ t \downarrow 0}} \frac{-f(y - tv) + f(y)}{t}.$$

Setting $z = y - tv$ the latter inequality reduces to

$$A(v) \geq \liminf_{\substack{z \to x \\ t \downarrow 0}} \frac{f(z + tv) - f(z)}{t},$$

as required. □

Corollary 4.9 *The generalized (Clarke) gradient is contained in the L-derivative.*

Proof This follows from Lemma 4.6 and Lemma 4.8. □

We do not know if the L-derivative and the Clarke gradient coincide on an infinite dimensional Banach space. We do know however that in finite dimensions they are the same, as we will show in Section 8.

5 Domain for Lipschitz functions

We will construct a domain for locally Lipschitz functions and for $C^1(U)$. The idea is to use step functions in $U \to_s \mathbf{I}\mathbb{R}$ to represent the function and step functions in $U \to \mathbf{C}(X^*)$ to represent the differential properties of the function. Note that a continuous partial function f of type $U \to \mathbb{R}$, as we have considered in defining ties of functions in Section 3, can be regarded as an element $\hat{f}$ of $U \to_s \mathbf{I}\mathbb{R}$ with $\hat{f}(x) = f(x)$ if $f(x)$ is defined and $\hat{f}(x) = \bot = \mathbb{R}$ otherwise; we always identify f and $\hat{f}$. Furthermore, a function $f \in U \to \mathbf{I}\mathbb{R}$ is given by a pair of respectively lower and upper semi-continuous functions $f^-, f^+ : U \to \mathbb{R}$ with $f(x) = [f^-(x), f^+(x)]$.

Consider the *consistency* relation

$$\mathsf{Cons} \subset (U \to_s \mathbf{I}\mathbb{R}) \times (U \to_s \mathbf{C}(X^*)),$$

defined by $(f, g) \in \mathsf{Cons}$ if $\uparrow\!f \cap \int g \neq \emptyset$. For a consistent (f, g), we think of f as the *function part* or the *function approximation* and g as the *derivative part* or the *derivative approximation*. We will show that the consistency relation is Scott closed. The proofs of the rest of results in this section are essentially as in [9] for the case of $X = \mathbb{R}$. We will present them here for a general Banach space X for the sake of completeness.

16

Proposition 5.1 *Let $g \in U \to_s \mathbf{C}(X^*)$ and $(f_i)_{i \in I}$ be a non-empty family of functions $f_i : dom(g) \to \mathbb{R}$ with $f_i \in \int g$ for all $i \in I$. If $h_1 = \inf_{i \in I} f_i$ is real-valued then $h_1 \in \int g$. Similarly, if $h_2 = \sup_{i \in I} f_i$ is real-valued, then $h_2 \in \int g$.*

Proof Suppose h_1 is real-valued. Let $a \searrow b \sqsubseteq g$. We have $f_i(x) - f_i(y) \in b(x - y)$ for all $i \in I$. Thus, $(b(x - y))^- \leq f_i(x) - f_i(y) \leq (b(x - y))^+$. Thus, $\inf_{i \in I} f_i(x) \leq f_i(y) + (b(x - y))^+$. Taking infimum again, we obtain $\inf_{i \in I} f_i(x) \leq \inf_{i \in I} f_i(y) + (b(x-y))^+$ and hence: $h_1(x) - h_1(y) \leq (b(x-y))^+$. Similarly, $(b(x - y))^- \leq h_1(x) - h_1(y)$ and the result follows. The case of h_2 is similar. $\square$

Let $R[0, 1]$ be the set of partial maps of $[0, 1]$ into the extended real line. Consider the two dcpo's $(R[0, 1], \leq)$ and $(R[0, 1], \geq)$. Define the maps $s : (U \to_s \mathbf{I}\mathbb{R}) \times (U \to_s \mathbf{C}(X^*)) \to (R, \leq)$ and $t : (U \to_s \mathbf{I}\mathbb{R}) \times (U \to_s \mathbf{C}(X^*) \to (R, \geq)$ by

$$s : (f, g) \mapsto \inf\{h : dom(g) \to \mathbb{R} \mid h \in \int g \ \& \ h \geq f^-\}$$

$$t : (f, g) \mapsto \sup\{h : dom(g) \to \mathbb{R} \mid h \in \int g \ \& \ h \leq f^+\}.$$

We use the convention that the infimum and the supremum of the empty set are ∞ and $-\infty$, respectively. Note that given a connected component A of $dom(g)$ with $A \cap dom(f) = \emptyset$, then $s(f, g)(x) = -\infty$ and $t(s, f)(x) = \infty$ for $x \in A$. In words, $s(f, g)$ is the least primitive map of g that is greater than the lower part of f, whereas $t(f, g)$ is greatest primitive map of g less that the upper part of f.

Proposition 5.2 *The following are equivalent:*

   (i) $(f, g) \in$ Cons.

   (ii) $s(f, g) \leq t(f, g)$.

   (iii) *There exists a locally Lipschitz function $h : dom(g) \to \mathbb{R}$ with $g \sqsubseteq \mathcal{L}h$ and $f \sqsubseteq h$ on $dom(g)$.*

Proof If $dom(f) \cap dom(g) = \emptyset$, then the three statements hold trivially. So assume in the following proof that $dom(f) \cap dom(g) \neq \emptyset$.
(ii) $\Rightarrow$ (i). Suppose $s(f, g) \leq t(f, g)$. Then, $s(f, g) \in \uparrow f \cap \int g$ and hence $(f, g) \in$ Cons.
(i) $\Rightarrow$ (ii). Suppose $(f, g) \in$ Cons. Assume $h \in \uparrow f \cap \int g$. Then, the induced map $h : dom(g) \to \mathbb{R}$ satisfies $h \in \int g$. Hence, $f^- \leq h \leq f^+$ and thus $s(f, g) \leq t(f, g)$.
(ii) $\Rightarrow$ (iii). Suppose $s(f, g) \leq t(f, g)$. Put $h = s(f, g)$.
(iii) $\Rightarrow$ (ii). We have $s(f, g) \leq h \leq t(f, g)$. $\square$

Moreover, s and t are well-behaved:

Proposition 5.3 *The maps s and t are Scott continuous.*

Proof Consider the map s. If $f_1 \sqsubseteq f_2$ and $g_1 \sqsubseteq g_2$, then we have $\int g_1 \supseteq \int g_2$ and $f_1^- \leq f_2^-$ and it follows that $s(f_1, g_1) \leq s(f_2, g_2)$. Let $\{(f_i, g_i)\}_{i \in I}$ be a directed set and put $f = \sup_{i \in I} f_i$ and $g = \sup_{i \in I} g_i$. To show the continuity of s we need to show that $\sup_{i \in I} s(f_i, g_i) \geq s(f, g)$ on any connected component of $\mathrm{dom}(g) = \bigcup_{i \in I} \mathrm{dom}(g_i)$. Take any such connected component $A \subseteq \mathrm{dom}(g)$. If $A \cap \mathrm{dom}(f) = \emptyset$ then $s(f, g) = -\infty$ on A and the result follows. Assume that $A \cap \mathrm{dom}(f) \neq \emptyset$, i.e., $\mathrm{dom} f_{i_0} \cap \mathrm{dom} g_{i_0} \neq \emptyset$ for some $i_0 \in I$. If $s(f_i, g_i) = \infty$ on $A \cap \mathrm{dom}(g_i)$ for some $i \geq i_0$, then $\sup_{i \in I} s(f_i, g_i) = \infty$ on A and the result follows again. Otherwise, assume without loss of generality that $-\infty < s(f_i, g_i) < \infty$ on $A \cap \mathrm{dom}(g_i)$ for all $i \in I$. Then from $(s(f_i, g_i))\!\restriction_A \in \int g_i$ it follows that $\forall i \geq j.\ (s(f_i, g_i))\!\restriction_A \in \int g_j$, and hence, by Proposition 5.1, $(\sup_{i \in I}(s(f_i, g_i)\!\restriction_A)) \in \int g_j$. Thus $(\sup_{i \in I} s(f_i, g_i))\!\restriction_A \in \sup_j \int g_j = \int \sup g_j$. On the other hand, $s(f_i, g_i) \geq f_i^-$ on A implies $\sup_{i \in I} s(f_i, g_i) \geq f_i^-$ on A and hence $\sup_{i \in I} s(f_i, g_i) \geq f^-$ on A. This shows that s is continuous. Similarly t is continuous. $\square$

This enables us to deduce:

Corollary 5.4 *The relation* Cons *is Scott closed.*

Proof Let $(f_i, g_i)_{i \in I} \subset (U \to_s \mathbb{IR}) \times (U \to_s \mathbf{C}\mathbb{R}^n)$ be a directed set with $(f_i, g_i) \in$ Cons for all $i \in I$. Then, by Proposition 5.2, $s(f_i, g_i) \leq t(f_i, g_i)$ for all $i \in I$. Hence, $s(f, g) = \sup_{i \in I} s(f_i, g_i) \leq \inf_{i \in I} t(f_i, g_i) = t(f, g)$. $\square$

We can now sum up the situation for a consistent pair of function and derivative information.

Corollary 5.5 *Let $(f, g) \in$ Cons. Then in each connected component A of the domain of definition of g which intersects the domain of definition of f, there exist two locally Lipschitz functions $s : A \to \mathbb{R}$ and $t : A \to \mathbb{R}$ such that $s, t \in \mathord{\uparrow} f \cap \int g$ and for each $u \in \mathord{\uparrow} f \cap \int g$, we have with $s(x) \leq u(x) \leq t(x)$ for all $x \in A$.*

We now can define a basic construct of this paper:

Definition 5.6 Define

$$D^1(U) = \{(f, g) \in (U \to_s \mathbb{IR}) \times (U \to_s \mathbf{C}(X^*)) : (f, g) \in \mathsf{Cons}\}.$$

Proposition 5.7 *For any $f \in (U \to \mathbb{R})$ the element $(f, \mathcal{L}f)$ is a maximal element of $D^1(U)$.*

Proof By Corollary 4.4, we have $f \in \int \mathcal{L}f$ and thus, $(f, \mathcal{L}f) \in D^1 U$. We now show that $(f, \mathcal{L}f)$ is maximal. If $\mathcal{L}f \sqsubseteq g$ and $(f, g) \in D^1 U$ then we have $f \in \int g$, which implies $g \sqsubseteq \mathcal{L}f$, i.e., $g = \mathcal{L}f$ and $(f, \mathcal{L}f)$ is maximal. $\square$

For a locally Lipschitz function $f : U \to \mathbb{R}$ the L-derivative satisfies $\mathcal{L}f(x) \neq \perp$ for all $x \in U$, whereas for a piecewise C^1 function f we further have the property that $\mathcal{L}f(x)$ is maximal except for a finite set of points.

6 L-derivative in finite dimensions

Assume $X = \mathbb{R}^n$ and $U \subset \mathbb{R}^n$ is an open subset. Then we can identify $X^* = X = \mathbb{R}^n$. Moreover $\mathbf{C}(\mathbb{R}^n)$ and $U \to \mathbf{C}(\mathbb{R}^n)$ are both countably based bounded complete continuous dcpo's with $U \to_s \mathbf{C}(\mathbb{R}^n) = U \to \mathbf{C}(\mathbb{R}^n)$.

In the finite dimensional case, we are able to deduce the following proposition which relates the L-derivative to its classical counterpart. For any compact subset $c \subset \mathbb{R}^n$ we denote its diameter by $w(c)$. For a non-empty compact interval $c = [c^-, c^+] \subset \mathbb{R}$, we thus have $w(c) = c^+ - c^-$. The following result generalizes Theorem 4.2(i) in finite dimensions; we do not know if it can be extended to infinite dimensional Banach spaces.

Proposition 6.1 *If $\mathcal{L}f(y) \in \mathbf{C}(\mathbb{R}^n)$ is maximal for some $y \in U$, then the Fréchet derivative of f exists at y and $f'(y) = \mathcal{L}f(y)$.*

Proof Put $c := \mathcal{L}f(y) = \bigcap\{b \,|\, y \in a \ \& \ f \in \delta(a, b)\}$. Let $\epsilon > 0$ be given. Take $a \searrow b \sqsubseteq \mathcal{L}f$ with $y \in a$ and $w(b) < \epsilon$. Note that $b \sqsubseteq c$ and there exists $\delta > 0$ such that $\|x - y\| < \delta$ implies $x \in a$. We have $b(x - y) \sqsubseteq f(x) - f(y)$ for $x \in a$, and

$$w(b(x - y)) \leq w(b)|x - y| \leq \epsilon |x - y|.$$

Since $b(x - y) \sqsubseteq c(x - y)$, we obtain $|f(x) - f(y) - c(x - y)| \leq \epsilon |x - y|$ and the result follows by the definition of Fréchet derivative. $\square$

We can now obtain the following result in finite dimensions, which is simply the classical version of the Fundamental Theorem of Calculus.

Corollary 6.2 *Suppose $g : U \to \mathbb{R}$ is a continuous function. Then $f \in \int g$ implies that f' exists in U and we have $\mathcal{L}f = f' = g$.*

Proof By Theorem 4.4, $g \sqsubseteq \mathcal{L}f$ and thus $\mathcal{L}f = g$ since g is maximal. By Proposition 6.1, we also obtain $\mathcal{L}f = f'$. $\square$

We now consider a given Cartesian coordinate system denoted say by e with basis $(e_1, \cdots, e_n)$. Let $\mathbf{Q}_e(\mathbb{R}^n)$ be the collection of all n dimensional compact hyper-rectangles with edges parallel with e_i's and define the *rectangular L-derivative with respect to e* as

$$(\mathcal{L}f)^e = \sup\{a \searrow b : f \in \delta(a, b) \ \& \ b \in \mathbf{Q}_e(\mathbb{R}^n)\}.$$

It immediately follows from the definition that $(\mathcal{L}f)^e(y)$, at each point $y \in U$, is an n dimensional compact hyper-rectangle with edges parallel to the basis vectors e_i. Moreover, if E denotes the collection of all Cartesian coordinate systems in $\mathbb{R}^n$, we have:

Proposition 6.3 *For each point $y \in U$, we have:*

$$\mathcal{L}f(y) = \bigcap_{e \in E} (\mathcal{L}f)^e(y). \quad \square$$

For $b \in \mathbf{Q}_e(\mathbb{R}^n)$, i.e., $b = b_1 \times \ldots \times b_n$, the relation $b(x-y) \sqsubseteq f(x) - f(y)$, which defines the single tie $\delta(a, b)$, can be computed in the coordinate system e simply as $\sum_{j=1}^n b_j(x_j - y_j) \sqsubseteq f(x) - f(y)$. Furthermore, if $x_j = y_j$ for all $j \neq i$, the relation reduces to $b_i(x_i - y_i) \sqsubseteq f(x) - f(y)$. This suggests a characterization of $(\mathcal{L}f)^e(y)$ in terms of Dini's derivatives; in fact we can deduce the following result.

Proposition 6.4 *The components of the rectangular L-derivative with respect to the basis e are given by:*

$$(\mathcal{L}f)_i^e(x) = \left[\liminf_{y \to x}(\nabla f)_i^l(y), \limsup_{y \to x}(\nabla f)_i^u(y) \right],$$

when the two limits are finite for all $i = 1, \cdots, n$ and $\mathcal{L}f(x) = \perp$ otherwise.

Proof If $f \in \delta(a, b)$ for some $a \subseteq U$ with $x \in a$ and $b = [b_1^-, b_1^+] \times \cdots \times [b_n^-, b_n^+]$, then $b_i^- \leq (\nabla f)_i^l(y) \leq (\nabla f)_i^u(y) \leq b_i^+$ for $y \in a$, and thus $b_i^- \leq \liminf_{y \to x}(\nabla f)_i^l(y) \leq \limsup_{y \to x}(\nabla f)_i^u(y) \leq b_i^+$. It follows that

$$\left[\liminf_{y \to x}(\nabla f)_i^l(y), \limsup_{y \to x}(\nabla f)_i^u(y) \right] \subseteq (\mathcal{L}f)_i^e(x).$$

On the other hand, if $\liminf_{y \to x}(\nabla f)_i^l(y)$ and $\limsup_{y \to x}(\nabla f)_i^u(y)$ are finite for all $i = 1, \cdots, n$, then for any $\epsilon > 0$ there exists an open $a \subseteq U$ containing x such that, for all $y \in a$ and all $i = 1, \ldots, n$,

$$K_i(x, \epsilon) := \liminf_{y \to x}(\nabla f)_i^l(y) - \epsilon < (\nabla f)_i^l(y)$$

$$L_i(x, \epsilon) := \limsup_{y \to x}(\nabla f)_i^u(y) + \epsilon > (\nabla f)_i^u(y).$$

Let c be the interior of a hypercube containing x with $\overline{c} \subset a$ and fix i with $1 \leq i \leq n$. By the first inequality above, we can cover $\overline{c}$ with a finite number of open hyper-rectangles such that for any pair of points $y \geq z$ in each hyper-rectangle with $y_j = z_j$ for $j \neq i$ and $y_i \geq z_i$, we have $K_i(x, \epsilon)(y_i - z_i) \leq f(y) - f(z)$. It thus follows, by adding a finite number of inequalities one for each open hyper-rectangle, that for all $y, z \in c$ with $y_j = z_j$ for $j \neq i$ and $y_i \geq z_i$, we have $K_i(x, \epsilon)(y_i - z_i) \leq f(y) - f(z)$, and similarly, by using the second inequality above, $f(y) - f(z) \leq L_i(x, \epsilon)(y_i - z_i)$. Thus, for all $y, z \in c$ with $y_j = z_j$ for $j \neq i$, we have $b_i(y_i - z_i) \sqsubseteq f(y) - f(z)$ where $b_i = [K_i(x, \epsilon), L_i(x, \epsilon)]$. For any pair $y, z \in c$, consider the $n + 1$ points $y = p^0, p^1, p^2, \ldots p^{n-1}, p^n = z$, such that $p_i^i = z_i$ and $p_j^i = y_j$ for $j \neq i$. Therefore,

$$f(y) - f(z) =$$

$$(f(p^0) - f(p^1)) + (f(p^1) - f(p^2)) + \ldots + (f(p^i) + f(p^{i+1})) + \ldots (f(p^n) - f(p^{n+1})) \sqsupseteq$$

$$\sum_{i=1}^n b_i(y_i - z_i).$$

It follows that $f \in \delta(c, b)$. Since $\epsilon > 0$ is arbitrary, we conclude that

$$\left[\liminf_{y \to x}(\nabla f)_i^l(y), \limsup_{y \to x}(\nabla f)_i^u(y)\right] \supseteq (\mathcal{L}f)_i^e(x). \qquad \square$$

The domain-theoretic derivative developed in [10] is indeed $(\mathcal{L}f)^e$, the rectangular L-derivative with respect to a given coordinate axis e. We do not know if there is an analogue of the above Proposition for infinite dimensional separable Hilbert spaces.

7 Computability

Let Z be a topological space with a countable basis M of its open subsets, and D a dcpo with a countable subset $E \subset D$. Let $(f_i)_{i \geq 0}$ be an effective enumeration of the class of step functions of $Z \to D$ made from single-step functions $a \searrow b$ where $a \in M$ and $b \in E$. We say $f \in U \to_s D$ is *computable* with respect to this enumeration if there exists a total recursive function $\phi : \mathbb{N} \to \mathbb{N}$ such that $(f_{\phi(n)})_{n \geq 0}$ is an increasing sequence with $f = \sup_{n \geq 0} f_{\phi(n)}$.

When, in addition, Z is locally compact and D is a countably based continuous dcpo, then $Z \to D$ is a countably based continuous dcpo, which can be given an effective structure. In this case, we obtain the same class of computable elements with any effective change of a countable basis of D. In general however, the computable elements will depend on the enumeration of the countable subset E.

Suppose now that X is a separable Banach space, with a countable dense set $P \subset X$. Then the collection of open balls centred at points of P with rational radii provides a countable basis of the norm topology on X. We use the rational compact intervals as a countable basis of $\mathbb{IR}$ and the collection of closed balls of X^* with centres at points P with rational radii as a countable subset of $\mathbf{C}(X^*)$ to generate two countable sets, S_1 and S_2 say, of step functions for the two dcpo's $U \to_s \mathbb{IR}$ and $U \to_s \mathbf{C}(X^*)$. We then obtain an enumeration $(f_i)_{i \geq 0}$ of S_1 and an enumeration $(g_i)_{i \geq 0}$ of S_2.

By Proposition 3.14, we know that any continuous function $f : U \to \mathbb{R}$ and any function $g : U \to \mathbf{C}(X^*)$ continuous with respect to the norm topology on X and X^*, is the supremum of step functions in S_1 and S_2 respectively. We say that f is *computable* with respect to the enumeration $(f_i)_{i \geq 0}$, respectively g is *computable* with respect to $(g_i)_{i \geq 0}$, if f considered as an element of $U \to \mathbb{IR}$, respectively g considered as an element of $U \to \mathbf{C}(X^*)$, is computable with respect to the enumeration.

We then use an oracle to decide if $(f_i, g_j) \in \mathsf{Cons}$ for $i, j \geq 0$, which enables us to construct an enumeration $(h_i)_{i \geq 0}$ of a countable set, S_3 say, of step functions of $D^1(U)$, where $h_i = (f_{p(i)}, g_{q(i)})$ for $i \geq 0$ with $p, q : \mathbb{N} \to \mathbb{N}$ total recursive functions. By Proposition 3.14, we know that if $f : U \to \mathbb{R}$ is Fréchet differentiable then (f, f') is the lub of step functions in S_3. We thus say that f and its

Fréchet derivative f' are computable with respect to $(h_i)_{i \geq 0}$ if (f, f') considered as a maximal element of $D^1(U)$ is computable with respect to this enumeration.

As we will see in the next section, when X is finite dimensional $D^1(U)$ can be given an effective structure with respect to which Cons is decidable, obviating the need for an oracle.

7.1 An effectively given domain for Lipschitz functions

In the finite dimensional case, $X = \mathbb{R}^n$, the countably based bounded complete continuous dcpo's $U \to \mathbf{IR}$ and $U \to \mathbf{C}(\mathbb{R}^n)$ have each a canonical basis respectively made from single-step functions $a \searrow b$, where a is an open ball with a rational radius centred at a point in U with rational coordinates and b is a rational compact interval, respectively a convex compact polyhedra in $\mathbb{R}^n$ with vertices having rational coordinates.

In [10], it is shown that, when the rectangular L-derivative $(\mathcal{L}f)^e$ with respect to a given coordinate axis e is used, the corresponding consistency predicate Cons^e, defined on $(U \to \mathbf{IR}) \times (U \to \mathbf{IR}^n)$ by $(f, g) \in \mathsf{Cons}^e$ if there exists $h : U \to \mathbb{R}$ such that $f \sqsubseteq h$ and $g \sqsubseteq (\mathcal{L}h)^e$ is decidable on the basis elements. The proof of decidability is fairly simple for $n = 1$ with an algorithm to test consistency, which is linear in the total number of single-step functions in the function and derivative parts [7]. In higher dimensions, the existing proof of decidability in [10] is long and the algorithm to test consistency is super-exponential. First, one checks, by a generalization of Green's theorem, if $\int^e g \neq \emptyset$ where $\int^e$ is the primitive map dual to the rectangular L-derivative with respect to e, i.e., $f \in \int^e g$ if $g \sqsubseteq (\mathcal{L}f)^e$. If the test for integrability of g is positive then one checks if $s^e(f, g) \leq t^e(f, g)$ where s^e and t^e are defined as s and t in Section 5 except that $\int^e$ is used in their definitions.

Using the decidability of Cons^e on basis elements, we can provide an effective structure for the sub-domain of $D^1(U)$ which uses the rectangular L-derivative. In particular this will characterize real-valued functions on $U \subset \mathbb{R}^n$ which are computable and have a computable rectangular L-derivative as pairs $(f, \mathcal{L}f)$ for which there exists a total recursive function $\phi : \mathbb{N} \to \mathbb{N}$ with $(f, \mathcal{L}^e f) = \sup_{i \geq 0}(f_{p(\phi(i))}, g_{q(\phi(i))})$ in the notation of Section 7.

If $f : U \to \mathbb{R}$ is C^{m-1} for some open subset $U \subset \mathbb{R}^n$, i.e., if it has continuous Fréchet derivatives $f^{(d)}$ of order d with $1 \leq d \leq m - 1$, then the L-derivative of components of $f^{(m-1)}$ exists. One can extend the construction of $D^1(U)$ to higher derivatives and build a domain $D^m(U)$ for representing and approximating a function together with its $m - 1$ Fréchet derivatives and its mth L-derivative $\mathcal{L}f^{(m-1)}$. The basis of this domain will consist of $m+1$ step functions representing approximations to the function, its first $m - 1$ Fréchet derivatives and its mth L-derivative. We will discuss the question of decidability of consistency for basis elements of this domain in the final section.

8 Relation with generalized gradient

Recall that by Rademacher's theorem [5, p 148], a function $f : U \to \mathbb{R}$ which is Lipschitz in an open neighbourhood of $U \subset \mathbb{R}^n$ is differentiable almost everywhere with respect to the n-dimensional Lebesgue measure in that neighbourhood. Let $\Omega_f \subseteq U$ denote the set of points, where f is not differentiable.

We now establish the equality of the L-derivative and the generalized gradient in finite dimensions..

Theorem 8.1 *For any function $f : U \to \mathbb{R}$, the rectangular L-derivative with respect to a given Cartesian coordinate system, at a point where the function is locally Lipschitz is the smallest hyper-rectangle with sides parallel to the coordinate planes that contains the generalized gradient at that point.*

Proof Fix a Cartesian coordinate system e. By Corollary 4.9 and Proposition 6.3, we already know that

$$\partial f(x) \subseteq (\mathcal{L}f(x))^e. \tag{4}$$

We show that $(\mathcal{L}f(x))^e$ is the smallest hyper-rectangle with sides parallel to the coordinate planes, which contains $\partial f(x)$. Assume $\epsilon > 0$ is given, $1 \leq i \leq n$ and let $B \subset \mathbb{R}^n$ be the unit closed ball centred at the origin. For $1 \leq i \leq n$, let $\pi_i : \mathbb{R}^n \to \mathbb{R}$ be the projection to the i coordinate axis and consider the pointwise extension of π_i to compact subsets of $\mathbb{R}^n$. From Equation 1, we have:

$$(\pi_i(\partial f(x)))^+ = \limsup\{(\nabla f)_i(y) : y \to x,\ y \notin \Omega_f\}.$$

Thus, there exists $\delta > 0$ such that for all $y \in x + \delta B$, we have $(\nabla f)_i(y) \leq (\pi_i(\partial f(x)))^+ + \epsilon$. Consider the line segment $L_y = \{y + te_i : 0 < t < \delta/2\}$, where e_i is the unit vector in the direction of the ith coordinate axis. Since Ω_f has zero n-dimensional Lebesgue measure in $x + \delta B$, it follows from Fubini's theorem that for almost all $y \in x + \frac{\delta}{2}B$, the line segment L_y meets Ω_f in a set of zero one-dimensional Lebesgue measure. If y is such a point and $0 < t < \delta/2$, we obtain:

$$f(y + te_i) - f(y) = \int_0^t (\nabla f)_i(y + se_i)\, ds,$$

since, by Rademacher's theorem [5, p 148], f' exists almost everywhere on L_y. On the other hand, $(\nabla f)_i(y + se_i) \leq (\pi_i(\partial f(x)))^+ + \epsilon$, since $\|y + se_i - x\| < \delta$ for $0 < s < t$. Thus,

$$f(y + te_i) - f(y) \leq t(\pi_i(\partial f(x)))^+ + \epsilon). \tag{5}$$

Equation 5 holds for almost all y within $\delta/2$ of x and for all $t \in (0, \delta/2)$. Since f, being Lipschitz, is continuous it follows that Equation 5 holds for all y within $\delta/2$ of x and for all $t \in (0, \delta/2)$. Thus, $(\nabla f)_i^u(y) \leq (\pi_i(\partial f(x)))^+ + \epsilon$ for all y within $\delta/2$ of x and, using Proposition 6.4, we conclude that

$$((\mathcal{L}f(x))_i^e)^+ = \limsup_{y \to x}(\nabla f)_i^u(y) \leq (\pi_i(\partial f(x)))^+. \tag{6}$$

Similarly,

$$((\mathcal{L}f(x))_i^e)^- = \liminf_{y \to x}(\nabla f)_i^l(y) \geq$$

$$\liminf\{(\nabla f)_i(y) : y \to x, y \notin \Omega_f\} = (\pi_i(\partial f(x)))^-. \tag{7}$$

Comparing Equations (6) and (7) with Equation (4), it follows that $\partial f(x)$ touches all the $2n$ sides of the hyper-rectangle $(\mathcal{L}f(x))^e$ and the proof is complete. $\square$

Corollary 8.2 *For any locally Lipschitz map $f : U \to \mathbb{R}$, the L-derivative and the Clarke gradient are equal: $\mathcal{L}f = \partial f$. $\square$*

Thus, in finite dimensions, the L-derivative gives a new representation for the Clarke gradient and the construction of an effectively given domain for locally Lipschitz functions provides a new computational framework for its applications. We note that the proof of Theorem 8.1 uses Proposition 6.4, for which we do not know any infinite dimensional analogue.

9 Further work and open problems

As already pointed out, it remains an open question if the L-derivative coincides with the Clarke gradient on infinite dimensional Banach spaces. It is also unknown if the Clarke gradient is upper semi-continuous in infinite dimensions, a property which holds for the L-derivative as we have shown in this paper. On the other hand, it will be interesting to see if the L-derivative can be extended to functions from a Banach space to a finite dimensional Banach space, for example to the complex plane, a case which has applications in quantum field theory.

There are quite a few unsolved problems in finite dimensions. For $n = 1$, the algorithm for testing consistency of basis elements in $D^1(U)$ is linear as already mentioned. For $D^2(U)$, consistency on basis elements is decidable but the present algorithm to test it is super-exponential in the total number of single-step functions for the three approximations of the function part, the derivative part and the second derivative part. [1]. Decidability of consistency for $D^m(U)$ when $m > 2$ is unknown. For $n = 2$, consistency on basis elements for $D^1(U)$ is decidable but the algorithm to test it in [10] is super-exponential. The complexity of consistency test in this case is unknown as is the question of decidability of consistency of basis elements for $D^m(U)$ when $m > 1$.

Based on the domain-theoretic framework for differential calculus, one can embark on the task of constructing a domain for orientable Euclidean manifolds, which would extend the set-theoretic model for computational geometry and solid modelling presented in [8] to the piecewise smooth setting.

References

[1] S. Abolfathbeigi and M. Mahmoudi. Consistency for approximating twice differentiable functions, 2003. Manuscript in Persian, Department of Mathematical Sciences, Sharif University of Technology, Tehran, Iran.

[2] S. Abramsky and A. Jung. Domain theory. In S. Abramsky, D. M. Gabbay, and T. S. E. Maibaum, editors, *Handbook of Logic in Computer Science*, volume 3. Clarendon Press, 1994.

[3] F. H. Clarke. Private communications. Summer 2005.

[4] F. H. Clarke. *Optimization and Nonsmooth Analysis*. Wiley, 1983.

[5] F. H. Clarke, Yu. S. Ledyaev, R. J. Stern, and P. R. Wolenski. *Nonsmooth Analysis and Control Theory*. Springer, 1998.

[6] A. Edalat. Dynamical systems, measures and fractals via domain theory. *Information and Computation*, 120(1):32–48, 1995.

[7] A. Edalat, M. Krznarić, and A. Lieutier. Domain-theoretic solution of differential equations (scalar fields). In *Proceedings of MFPS XIX*, volume 83 of *Electronic Notes in Theoretical Computer Science*, 2003. Full paper in `www.doc.ic.ac.uk/~ae/papers/scalar.ps`.

[8] A. Edalat and A. Lieutier. Foundation of a computable solid modelling. *Theoretical Computer Science*, 284(2):319–345, 2002.

[9] A. Edalat and A. Lieutier. Domain theory and differential calculus (Functions of one variable). *Mathematical Structures in Computer Science*, 14(6):771–802, December 2004.

[10] A. Edalat, A. Lieutier, and D. Pattinson. A computational model for multivariable differential calculus. In V. Sassone, editor, *Proc. FoSSaCS 2005*, volume 3441, pages 505–519, 2005.

[11] A. Edalat and D. Pattinson. A domain-theoretic account of Picard's theorem. *LMS Journal of Computation and Mathematics*, 10:83–118, 2007.

[12] A. Edalat and D. Pattinson. Inverse and implicit functions in domain theory. In P. Panangaden, editor, *Proc. 20th IEEE Symposium on Logic in Computer Science (LICS 2005)*, pages 417–426, 2005.

[13] A. Edalat and D. Pattinson. Denotational semantics of hybrid automata. In L. Aceto and A. Ingofsdottir, editors, *Proc. FoSSaCS 2006*, volume 3921, pages 231–245, 2006.

[14] N. Fars. Aspects analytiques dans la mathematique de Sharaf al-din al-Tusi. *Historia Sc.*, 5(1), 1995.

[15] N. Fars. Le calcul du maximum et "derive" selon Sharaf al-din al-Tusi. *Arabic Sci. Philos.*, 5(2):219–237, 1995.

[16] G. Gierz, K. H. Hofmann, K. Keimel, J. D. Lawson, M. Mislove, and D. S. Scott. *Continuous Lattices and Domains*. Cambridge University Press, UK, 2003.

[17] A. Grzegorczyk. Computable functionals. *Fund. Math.*, 42:168–202, 1955.

[18] A. Grzegorczyk. On the definition of computable real continuous functions. *Fund. Math.*, 44:61–71, 1957.

[19] J. P. Hogendijk. Sharaf al-din al-Tusi on the number of positive roots of cubic equations. *Fund. Math.*, 16(1):69–85, 1989.

[20] K. Lau and C. Weil. Differentiability via directional derivatives. *Proceedings of American Mathematical Society*, 70(1):11–17, June 1978.

[21] M. B. Pour-El and J. I. Richards. *Computability in Analysis and Physics*. Springer-Verlag, 1988.

[22] D. S. Scott. Outline of a mathematical theory of computation. In *4th Annual Princeton Conference on Information Sciences and Systems*, pages 169–176, 1970.

[23] M. B. Smyth. Topology. In S. Abramsky, D. Gabbay, and T. S. E. Maibaum, editors, *Handbook of Logic in Computer Science*, chapter 5. Oxford University Press, 1992.

[24] A. Turing. On computable numbers with an application to the Entscheidungsproblem. *Proc. London Mathematical Soc.*, 42:230–265, 1936.

[25] A. Turing. On computable numbers with an application to the Entscheidungsproblem. *Proc. London Mathematical Soc.*, 43:544–546, 1937.

[26] K. Weihrauch. *Computable Analysis (An Introduction)*. Springer, 2000.

[27] S. Yamamuro. *Differential calculus in Topological Linear Spaces*, volume 374 of *Lecture Notes in Mathematics*. Springer-Verlag, 1970.