Set-Valued Analysis **4**: 375–398, 1996.

Approximate Subgradients and Coderivatives in R^n

D. BORWEIN
Department of Mathematics, University of Western Ontario, London, Ontario N6A 5B7, Canada. e-mail: dborwein@uwo.ca.

J. M. BORWEIN*
Department of Mathematics and Statistics, Simon Fraser University, Burnaby, BC, V5A 1S6, Canada. e-mail: jborwein@cecm.sfu.ca.

XIANFU WANG
Department of Mathematics and Statistics, Simon Fraser University, Burnaby, BC, V5A 1S6, Canada. e-mail: xwang@cecm.sfu.ca.

(Received: 4 March 1996; in final form: 5 August 1996)

Abstract. We show that in two dimensions or higher, the Mordukhovich–Ioffe approximate subdifferential and Clarke subdifferential may differ almost everywhere for real-valued Lipschitz functions. Uncountably many Fréchet differentiable vector-valued Lipschitz functions differing by more than constants can share the same Mordukhovich–Ioffe coderivatives. Moreover, the approximate Jacobian associated with the Mordukhovich–Ioffe coderivative can be nonconvex almost everywhere for Fréchet differentiable vector-valued Lipschitz functions. Finally we show that for vector-valued Lipschitz functions the approximate Jacobian associated with the Mordukhovich–Ioffe coderivative can be almost everywhere disconnected.

Mathematics Subject Classifications (1991). Primary 49J52, Secondary 26A27, 26B12, 49J50, 52A20.

Key words: subgradient, coderivative, generalized Jacobian, Lipschitz function, bump function, gauge, nowhere dense set, Lebesgue measure, disconnectedness.

1. Introduction

In finite-dimensional spaces, for real-valued functions Mordukhovich and Ioffe [10, 15] have introduced the notion of *approximate subdifferential*. Rockafellar [20] has introduced the notion of *basic subdifferential* in terms of proximal subdifferential. The approximate and basic subdifferentials are in fact the same. The approximate subdifferential is an important notion because it is always nonconvex and contained in the *Clarke subdifferential*. For vector-valued functions, Mordukhovich and Ioffe [10, 16] have introduced the *approximate coderivative*. Being nonconvex-valued, the approximate coderivative is not dual to any tangentially generated derivative construction. Following a process identical to that carried out in the real-valued case, Clarke has introduced the *generalized Jacobian* which is always convex compact valued. Because the approximate coderivative

* Research supported by NSERC and the Shrum Endowment at Simon Fraser University.

and generalized Jacobian enjoy a rich calculus [7, 10, 18] they are widely used in set-valued analysis.

The aim of the present paper is to show how bump functions with prescribed gradient images can be used to build concrete examples so as to show starkly the difference between the approximate subdifferential and Clarke subdifferential, and the difference between the coderivatives and generalized Jacobians.

2. Basic Definitions and Preliminaries

Let h: $R^n \to [-\infty, +\infty]$ be proper and lower semicontinuous.

DEFINITION 1. A vector $z \in R^n$ is called a *proximal subgradient* of h at $\bar{x}$ if $h(\bar{x})$ is finite and for some $r \geqslant 0$ and $\delta > 0$ one has

$$h(x) \geqslant h(\bar{x}) + \langle z, x - \bar{x} \rangle - \tfrac{1}{2} r \|x - \bar{x}\|^2 \quad \text{when} \quad \|x - \bar{x}\| \leqslant \delta.$$

We denote $\partial_p h(x)$ the set of proximal subgradients of h at x.

DEFINITION 2. The Rockafellar *basic subdifferential* of h at $\bar{x}$ is defined to be the set

$$\partial_b h(\bar{x}) := \Big\{ \lim_{\nu \to \infty} z_\nu \colon\ z_\nu \in \partial_p h(x_\nu), x_\nu \to \bar{x}, h(x_\nu) \to h(\bar{x}) \Big\},$$

and the Rockafellar *singular basic subdifferential* is

$$\partial_b^\infty h(\bar{x}) := \Big\{ \lim_{\nu \to \infty} \lambda_\nu z_\nu \colon\ z_\nu \in \partial_p h(x_\nu), x_\nu \to \bar{x}, h(x_\nu) \to h(\bar{x}), \lambda_\nu \downarrow 0 \Big\}.$$

If $|h(x)| < \infty$ then

$$h^-(x; v) := \liminf_{\substack{t \downarrow 0 \\ u \to v}} \frac{h(x + tu) - h(x)}{t},$$

$$\partial^- h(x) := \{x^* \colon\ \langle x^*, v \rangle \leqslant h^-(x; v), \forall v \in R^n\},$$

are called the *lower Dini derivative* in direction v and the *Dini subdifferential* of h at x respectively. Set $U(h, \bar{x}, \delta) = \{z \colon\ \|z - \bar{x}\| < \delta,\ |h(z) - h(\bar{x})| < \delta\}$.

DEFINITION 3. The Mordukhovich–Ioffe *approximate subdifferential* of h at $\bar{x}$ is defined as follows:

$$\partial_a h(\bar{x}) := \bigcap_{\delta > 0} \bigcup_{z \in U(h, \bar{x}, \delta)} \partial^- h(z) = \limsup_{\substack{z \to \bar{x} \\ h(z) \to h(\bar{x})}} \partial^- h(z).$$

The Mordukhovich–Ioffe *singular approximate subdifferential* of h at $\bar{x}$ is

$$\partial_a^\infty h(\bar{x}) := \limsup_{\substack{z \to \bar{x} \\ h(z) \to h(\bar{x}) \\ \lambda \to 0+}} \lambda \partial^- h(z).$$

Note that for h being locally Lipschitz around $\bar{x}$ it is necessary and sufficient that $\partial_a^\infty h(\bar{x}) = \{0\}$ [18]. Throughout the paper we use μ to denote the Lebesgue measure. For any set A, conv(A), clconv(A), and cl(A) are used to denote the convex hull, closed convex hull and closure of A respectively.

DEFINITION 4. Suppose h is Lipschitz near $\bar{x}$, and suppose S is any set of Lebesgue measure 0 in R^n. The Clarke subdifferential of h at $\bar{x}$ is defined as

$$\partial_c h(x) := \text{conv}\{\lim \nabla h(x_i)\colon\ x_i \to x, x_i \notin S, x_i \notin \Omega_h\},$$

where the set of points at which h fails to be differentiable is denoted by Ω_h.

Note that in R^n for every lower semicontinuous function h, $\partial^- h(x) \neq \emptyset$ on a dense subset of its domain [11], this fails if we do not assume h to be lower semicontinuous. In fact, the Serpinski Theorem [5] shows that given any $t_n \downarrow 0$, there exists a finite-valued function $h\colon\ R \to R$ such that

$$\lim_{n\to\infty} \frac{h(x+t_n) - h(x)}{t_n} \equiv -\infty,$$

so $\partial^- h(x) = \partial_a h(x) = \emptyset$ for every x. The Clarke subdifferential definition given here is for Lipschitz functions, and it does not apply to continuous functions since even continuous functions in R can be nowhere differentiable. As immediately follows from the definitions, these subdifferentials have the following upper semicontinuity property: $\partial_\sharp h(\bar{x}) = \limsup_{y\to\bar{x}, h(y)\to h(\bar{x})} \partial_\sharp h(y)$, where $\sharp$ stands for b, a, and c. It turns out that the first two are sensitive to Lebesgue null sets while the third is not. On the real line two Lipschitz functions have the same Clarke subdifferential if and only if they have the same approximate subdifferential [3], but this is no longer so in higher dimensions. We summarize the relationship among these subdifferentials in the following proposition. Proofs may be found in [10, 20]. A more elementary proof of (i), based on a penalization scheme rather than using normal cones, is given in [4].

PROPOSITION 1. *Let* $h\colon\ R^n \to [-\infty, +\infty]$ *be lower semicontinuous and* $|h(x)| < +\infty$. *Then* (i) $\partial_a h(x) = \partial_b h(x)$, *and* (ii) $\partial_c h(x) = \text{clconv}[\partial_a h(x) + \partial_a^\infty h(x)]$.

Let $F\colon\ R^n \to R^m$ be a vector-valued Lipschitz function, written in terms of component functions as $F(x) := [f_1(x), f_2(x), \ldots, f_m(x)]$.

DEFINITION 5. The Clarke *generalized Jacobian* of F at x is the convex hull of all $m \times n$ matrices Z obtained as the limit of a sequence of the form $JF(x_i)$, where $x_i \to x$ and $x_i \notin \Omega_F$, the set of points at which F fails to be differentiable. That is,

$$\partial_{cJ} F(x) := \text{conv}\{\lim JF(x_i)\colon\ x_i \to x, x_i \notin \Omega_F\}.$$

It follows that $\partial_c F$ is a cusco [7] and

$$\partial_{cJ}F(x) \subset \partial_c f_1(x) \times \partial_c f_2(x) \times \cdots \times \partial_c f_m(x) \quad \text{for every } x \in \text{dom}(F).$$

It was shown by Warga [23] that $\partial_{cJ}F(x)$ is indifferent to the exclusion of x_i in the definition from an arbitrary set of measure 0.

Set $\text{gph}\, F := \{(x, F(x))\colon\ x \in \text{dom}(F)\}$ and let $\delta_{\text{gph}\, F}$ denote the indicator function of $\text{gph}\, F$, that is, the one equal to 0 on $\text{gph}\, F$ and $+\infty$ outside of $\text{gph}\, F$.

DEFINITION 6. The Mordukhovich–Ioffe *approximate coderivative* of F at x in the direction y^* is defined by

$$D^*F(x)(y^*) := \{x^* \in R^n\colon\ (x^*, -y^*) \in \partial_a \delta_{\text{gph}\, F}(x, F(x))\}.$$

We put $D^*F(x)(y^*) = \emptyset$ if $x \notin \text{dom}\, F$.

DEFINITION 7. The Aubin *contingent coderivative* of F at x in the direction y^* is defined as

$$\widehat{D}^*F(x)(y^*) := \{x^* \in R^n\colon\ \langle x^*, v\rangle \leqslant \langle y^*, u\rangle, \forall (v, u) \in \text{gph}\, \widehat{D}F(x)\},$$

where

$$\widehat{D}F(x)(v) := \left\{u \in R^m\colon\ \exists t_k \downarrow 0, \frac{F(x + t_k v) - F(x)}{t_k} \to u\right\}.$$

The Mordukhovich–Ioffe coderivative is robust in the sense that

$$D^*F(x)(y^*) = \limsup_{\substack{\tilde{x} \to x \\ \tilde{y} \to y^*}} D^*F(\tilde{x})(\tilde{y}),$$

and

$$D^*F(x)(y^*) = \limsup_{\substack{\tilde{x} \to x \\ \tilde{y} \to y^*}} \widehat{D}^*F(\tilde{x})(\tilde{y}) \quad \text{for every } y^* \in R^m.$$

In particular, when F is Fréchet differentiable at x we have

$$\widehat{D}^*F(x)(y^*) = \{JF(x)^t y^*\} \quad \text{for every } y^* \in R^m. \tag{1}$$

When F is Fréchet differentiable in a neighborhood of x we have

$$D^*F(x)(y^*) = \{\lim JF(\tilde{x})^t(\tilde{y})\colon\ \tilde{x} \to x, \tilde{y} \to y^*\}. \tag{2}$$

In [10, 17] it was shown that the approximate coderivative can be expressed in terms of the approximate subdifferential of the scalarization $\langle y^*, F\rangle(x) := \langle y^*, F(x)\rangle$. That is, $D^*F(x)(y^*) = \partial_a\langle y^*, F\rangle(x)$. Since $\partial_c\langle y^*, F(x)\rangle = \partial_{cJ}F(x)^t y^*$ [7], Proposition 1 shows that the relationship between the Clarke generalized Jacobian and the Mordukhovich–Ioffe coderivative is

$$\partial_{cJ}F(x)^t y^* = \text{conv}(\partial_a\langle y^*, F\rangle(x)) = \text{conv}(D^*F(x)(y^*)) \quad \text{for all } y^* \in R^m.$$

Similarly following Ioffe's definition for real-valued case:

DEFINITION 8. We call

$$\partial^- F(x) := [\partial^- f_1(x), \partial^- f_2(x), \ldots, \partial^- f_m(x)],$$

the *Dini Jacobian* of F at x.

$$\partial_{aJ} F(x) := \bigcap_{\delta>0} \bigcup_{y \in U(F,x,\delta)} \partial^- F(y) = \limsup_{y \to x} \partial^- F(y), \tag{3}$$

the *approximate Jacobian* of F at x.

Rademacher's Theorem asserts that F is differentiable a.e. on any neighborhood of x in which F is Lipschitz. Therefore $\partial_{aJ} F(x)$ is always a closed set and $\partial_{aJ} F(x) \neq \emptyset$ if F is Lipschitz near x. Like the approximate subdifferential, $\partial_{aJ} F$ is not stable under 'excluding sets of measure zero'. To be more precise, even when S is countable, adding a condition $x \notin S$ to the definition leads to a set that depends in general on S (see Section 4.4.1). Equations (1) and (2) show that when F is Fréchet differentiable, for fixed y^*, the approximate coderivative and contingent coderivative are completely characterized by

$$\partial_{aJ} F(x) = \{\lim JF(\tilde{x})\colon\ \tilde{x} \to x\} \quad \text{and} \quad \partial^- F(x) = \{JF(x)\}.$$

We conclude this section with Zahorski's Theorem [5] which is frequently used below.

PROPOSITION 2. *Let E be a set of type F_σ such that $d(E, x) = 1$ (the metric density of E at x) for all $x \in E$. Then there exists an approximately continuous function f such that $0 < f(x) \leqslant 1$ for all $x \in E$ and $f(x) = 0$ for all $x \notin E$. The function f is also upper semi-continuous.*

3. Real-Valued Lipschitz Functions

3.1. THE GRADIENT RANGES OF BUMP FUNCTIONS

In this section, we shall construct continuously Gateaux differentiable bump functions with prescribed gradient images. We use $R(\nabla f)$ to denote the gradient image of f on R^n, $\mathrm{bd}(C^0)$ the boundary of C^0, and $\mathrm{supp}(f)$ the supporting set of f. First, we make three definitions.

DEFINITION 9. Let $f\colon R^n \to R$ be convex and x and d be fixed in R^n. We call

$$f'(x, d) := \lim_{t \downarrow 0} \frac{f(x + td) - f(x)}{t} = \inf \left\{ \frac{f(x + td) - f(x)}{t}\colon\ t > 0 \right\},$$

the *directional derivative* of f at x in the direction d.

DEFINITION 10. A *convex body* in R^n is a bounded convex subset C such that $\mathrm{int}(C) \neq \emptyset$.

DEFINITION 11. Let C be a closed convex set containing the origin. The function ν_C defined by $\nu_C(x) := \inf\{\lambda > 0\colon\ x \in \lambda C\}$ is called the *gauge* of C. As usual, we set $\nu_C(x) := +\infty$ if $x \in \lambda C$ for no $\lambda > 0$.

It is well known that if $0 \in \mathrm{int}(C)$ then ν_C is a finite-valued nonnegative closed sublinear function. Our constructions used later are based on the bump functions built from gauges. The following classical results can be found in [8] and [13].

LEMMA 1. *Let S be a strictly convex nonempty closed set. Then for each $d \neq 0$ the face $F_S(d) := \{s \in S\colon\ \langle s, d\rangle = \sigma_S(d)\}$ is at most a singleton where $\sigma_S(d) := \sup\{\langle y, d\rangle\colon\ y \in S\}$ is the support function of S.*

LEMMA 2. *Let C be a closed convex set containing the origin. Its gauge ν_C is the support function of a closed convex set containing the origin, namely*

$$C^0 := \{s \in R^n\colon\ \langle s, d\rangle \leqslant 1 \quad \text{for all } d \in C\},$$

which is the polar of C.

LEMMA 3. *Let $f\colon\ R^n \to R$ be convex. For all x and d in R^n, we have*

$$F_{\partial_c f(x)}(d) = \partial_c[f'(x, \cdot)](d).$$

LEMMA 4. *Let C be a closed convex body with $0 \in \mathrm{int}(C)$. Then C^0 is a closed convex body and $0 \in \mathrm{int}(C^0)$.*

Let C be a convex body and $0 \in \mathrm{int}(C)$. Define $f(x) := \frac{3\sqrt{3}}{8}[(1 - \nu_C^2(x))^+]^2$. That is,

$$f(x) = \begin{cases} \frac{3\sqrt{3}}{8}(1 - \nu_C^2(x))^2, & \text{if } x \in C,\\ 0, & \text{if } x \notin C. \end{cases}$$

THEOREM 1. *Let f be defined as above. If C^0 is strictly convex then f is continuously Gateaux differentiable and $R(\nabla f) = C^0$.*

Proof. We consider two cases:

Case 1. Let $x \neq 0$. By Lemma 1, Lemma 2 and Lemma 3 we have:

$$\partial_c \nu_C(x) = \partial_c \sigma_{C^0}(x) = F_{C^0}(x).$$

By Theorem 2.3.9 [7, p. 42]

$$\begin{aligned} \partial_c f(x) &= \frac{3\sqrt{3}}{2}(1 - \nu_C^2(x))^+ \nu_C(x) \partial_c \nu_C(x) \\ &= \frac{3\sqrt{3}}{2}(1 - \nu_C^2(x))^+ \nu_C(x) \nabla \nu_C(x). \end{aligned}$$

Case 2. Let $x = 0$. Since $0 \in \text{int}(C)$, by Lemma 4 there exists $M > 0$ such that for every $y \in C^0$ we have $\|y\| \leqslant M$. Then

$$\begin{aligned}\lim_{\|x\|\to 0} \frac{\nu_C^2(x)}{\|x\|} &= \lim_{\|x\|\to 0} \nu_C(x)\frac{\nu_C(x)}{\|x\|} \\ &= \lim_{\|x\|\to 0} \nu_C(x)\frac{\sigma_{C^0}(x)}{\|x\|} \\ &\leqslant \lim_{\|x\|\to 0} \nu_C(x)M = 0,\end{aligned}$$

which implies $\nabla f(0) = \{0\}$. Hence

$$\begin{aligned}R(\nabla f) &= \bigcup_{x\in R^n} \frac{3\sqrt{3}}{2}(1-\nu_C^2(x))^+\nu_C(x)\nabla\nu_C(x) \\ &= \bigcup_{x\in C} \frac{3\sqrt{3}}{2}(1-\nu_C^2(x))^+\nu_C(x)\nabla\nu_C(x) \\ &= \bigcup_{1\geqslant\sigma\geqslant 0} \frac{3\sqrt{3}}{2}\sigma(1-\sigma^2)\text{bd}(C^0) = C^0,\end{aligned}$$

where $(3\sqrt{3}/2)\sigma(1-\sigma^2)$ is maximized at $1/\sqrt{3}$ with value 1. It follows that

$$R(\nabla f) = \bigcup_{\substack{x\in C\\ \nu_C(x)\leqslant \frac{1}{\sqrt{3}}}} \nabla f(x) = C^0. \qquad \square$$

THEOREM 2. *Every strictly convex closed body containing* 0 *in its interior is the gradient range of a continuous Gateaux differentiable bump function.*

Proof. Let C be any strictly convex closed body with $0 \in \text{int}(C)$. Define ν_{C^0} and f respectively by

$$\nu_{C^0}(x) := \inf\{t > 0\colon\ x \in tC^0\},$$

$$f(x) := \frac{3\sqrt{3}}{8}[(1-\nu_{C^0}^2(x))^+]^2.$$

Then by Theorem 1, f is strictly Gateaux differentiable and $R(\nabla f) = C$. By Lemma 4, f is a bump function. $\square$

THEOREM 3. *Let C_i be strictly convex closed bodies with $0 \in \text{int}(C_i)$ for $i \in I$ (a finite set). Then there is a continuous Gateaux differentiable bump function f such that $R(\nabla f) = \bigcup_{i\in I} C_i$.*

Proof. Let f_i be a bump function with $R(\nabla f_i) = C_i$ and with $\mathrm{supp}(f_i) \subset C_i^0$. Define g_i by

$$g_i(x) := \varepsilon_i f_i\left(\frac{x - x_i}{\varepsilon_i}\right).$$

Since C_i^0 is bounded and $0 \in \mathrm{int}(C_i^0)$, we can choose x_i and ε_i appropriately such that $\mathrm{supp}(g_i) \cap \mathrm{supp}(g_j) = \emptyset$ if $i \neq j$. Set $A_i := \mathrm{supp}(g_i)$ and define

$$g(x) := \begin{cases} g_i(x), & \text{if } x \in A_i, \\ 0, & \text{otherwise.} \end{cases}$$

Then g is a continuously Gateaux differentiable bump function and $R(\nabla g) = \bigcup_{i\in I} C_i$. □

EXAMPLE 1. As is well known, gauges and support functions of elliptic sets merit more detailed study. Given a symmetric positive definite operator Q, define

$$R^n \ni x \mapsto f(x) := \sqrt{\langle Qx, x\rangle}.$$

Then f is the gauge function of the sublevel-set $E_Q := \{x\colon\ f(x) \leqslant 1\}$. To see this, we write

$$\begin{aligned} f(x) &= \inf\{\lambda > 0\colon\ \langle Qx, x\rangle \leqslant \lambda^2\} \\ &= \inf\left\{\lambda > 0\colon\ \left\langle Q\frac{x}{\lambda}, \frac{x}{\lambda}\right\rangle \leqslant 1\right\} \\ &= \inf\left\{\lambda > 0\colon\ \frac{x}{\lambda} \in E_Q\right\}. \end{aligned}$$

Consider the polar of E_Q:

$$E_Q^0 := \{y\colon\ \langle y, x\rangle \leqslant 1 \text{ for all } x \text{ satisfying } \langle Qx, x\rangle \leqslant 1\}.$$

Letting $Q^{1/2}$ be the square root of Q, the change of variable $p = Q^{1/2}x$ gives

$$\begin{aligned} E_Q^0 &= \{y\colon\ \langle p, Q^{-1/2}y\rangle \leqslant 1, \|p\|^2 \leqslant 1\} \\ &= \{y\colon\ \|Q^{-1/2}y\| \leqslant 1\}, \end{aligned}$$

which is the dual ball of E_Q. By Lemma 2 the support function $\sigma_{E_Q^0}$ is exactly f. Let $g(x) := 3\sqrt{3}/8[(1 - f(x)^2)^+]^2$. Then

$$R(\nabla g) = \{y\colon\ \langle y, Q^{-1}y\rangle \leqslant 1\},$$

which is an elliptic set. By choosing different symmetric positive definite operators Q we can get different elliptic sets as the gradient ranges of bump functions. In particular, when $Q = I_n$ we get $E_Q = E_Q^0$ and this is the only norm on R^n having this property.

Remark 1. It is the shape of the range of the gradient of a C^1 bump function that determines the images of the subdifferentials which are constructed later. Theorem 3 shows we can intersperse different bumps to get interesting images.

3.2. THE CONSTRUCTION OF ALMOST EVERYWHERE NONCONVEX APPROXIMATE SUBDIFFERENTIALS

In this section, following [12] we shall use bump functions given in Section 3.1 to construct Lipschitz functions defined on R^2 whose approximate subdifferentials are not convex on sets with large measure. Note that for a Lipschitz function f we have $\{x\colon \partial_a f(x) = \partial_c f(x)\} = \{x\colon \partial_a f(x) \text{ is convex}\}$. We use h' to denote the gradient map of h, $h'(B)$ the gradient image of h on B, $\operatorname{dist}(p, h'(B))$ the distance of p to $h'(B)$.

THEOREM 4. *Let $\varepsilon > 0$. Then there exists a Lipschitz function defined on R^2 such that $0 < \mu(\{x\colon \partial_a f(x) \text{ is convex}\}) < \varepsilon$.*

We omit the proof of this theorem since it is exactly the same as the one given in [12]. The only difference is that we can use any bump function with nonconvex gradient image given in Section 3.1. We now sharpen Theorem 4 by showing:

THEOREM 5. *There is a Lipschitz function $f\colon R^2 \to R$ such that*

$$\mu(\{x\colon \partial_a f(x) \text{ is convex}\}) = 0.$$

Proof. Step 1. Construct sequences of closed balls $\{B_{in}\}$ such that

(i) For fixed i, the B_{in}'s are pairwise disjoint;
(ii) $\bigcup_{n=1}^{\infty} B_{(i+1)n}$ is dense and contained in $S_i := \bigcup_{n=1}^{\infty} B_{in}^0$;
(iii) $\mu(\bigcup_{n=1}^{\infty} B_{in}) \to 0$ as $i \to \infty$;
(iv) $\bigcup_{n=1}^{\infty} B_{0n}$ is dense in R^2,

where B_{in}^0 denotes the interior of B_{in} for every i and n. For fixed i on each $B_{in} := B(z_n^i, r_n^i)$ we define

$$f_i(x) := \begin{cases} r_n^i h((x - z_n^i)/r_n^i), & \text{if } x \in B_{in}^0 \\ 0, & \text{if } x \notin S_i, \end{cases}$$

where h is any bump function with $\operatorname{supp}(h) \subset B$ (the closed unit ball in R^2) and $h'(B)$ nonconvex.

Since $h'(B)$ is nonconvex we take $p \in \operatorname{conv}(h'(B))$ such that $\operatorname{dist}(p, h'(B)) = d > 0$. Let $M := \sup\{\|x^*\|\colon x^* \in h'(B)\}$ and $0 < k < \min(d/4M, 1/2)$. Set

$$f(x) := \sum_{n=0}^{\infty} k^n f_n(x) \quad \text{and} \quad F_m(x) := \sum_{n=m}^{\infty} k^n f_n(x).$$

We prove that $\{x\colon \partial_a f(x) \text{ is convex}\} \subset \bigcap_{n=0}^{\infty} S_n$.

Step 2. If $x \notin S_m$ then $\partial^- F_m(x) = \{(0,0)\}$. In fact, $F_m(x) = 0$ and, therefore, x is a local minimum of F_m and $F_m^-(x;v) \geqslant 0$ for all v. Fixing a direction v, there are two possibilities: Either there exists $t_n \downarrow 0$ such that $x + t_n v \notin S_m$ or there exists an $\varepsilon > 0$ such that $x + tv \in S_m$ for all $t \in (0,\varepsilon)$. In the first case $F_m^-(x;v) = 0$. In the second case, there exists an n such that $x + tv \in B_{mn}$ for all $t \in (0,\varepsilon)$ because the sets B_{mn} are pairwise disjoint and closed. As B_{mn} is closed we obtain $x \in \partial B_{mn}$. Then by (ii) $x \notin \bigcup_{n=1}^{\infty} B_{in}$ for any $i \geqslant m+1$. Therefore, there must be $t_\nu \downarrow 0$ such that $x + t_\nu v \notin S_{m+1}$. Thus

$$
\begin{aligned}
F_m^-(x;v) &= \liminf_{t\to 0^+} \frac{F_m(x+tv) - F_m(x)}{t} \\
&\leqslant \liminf_{\nu\to\infty} \frac{F_m(x+t_\nu v) - F_m(x)}{t_\nu} \\
&= k^m \liminf_{\nu\to\infty} \frac{f_m(x+t_\nu v) - f_m(x)}{t_\nu} \\
&= k^m h'((x - z_n^m)/r_n^m)v = 0.
\end{aligned}
$$

Therefore $F_m^-(x;v) = 0$ for all v, which is to say $\partial^- F_m(x) = \{(0,0)\}$. For any $x \in R^2$ and any positive integer n we have

$$\partial_a F_m(x) \subset \partial^- F_m(B_{mn}) \cup \{(0,0)\} = \partial^- F_m(B_{mn}).$$

If $x \notin S_m$, and $N(x,\delta)$ is any neighborhood of x then $N(x,\delta)$ contains a ball B_{mn}. Therefore $\partial^- F_m(B_{mn}) \subset \partial_a F_m(x)$. So for every $x \notin S_m$ we have $\partial_a F_m(x) = \partial^- F_m(B_{mn})$.

Step 3. If $x \notin S_m$ then $\partial_a F_m(x)$ is nonconvex. Since all F_m have the same structure we prove this only for $m = 0$. Let $x \in R^2 \backslash S_0$. By Step 2, we deduce that $\partial_a f(x) = \partial^- f(B_{0n})$. We only need to show $\partial^- f(B_{0n})$ is not convex. Indeed, for any neighborhood U of x there is B_{0n} such that $B_{0n} \subset U$. For any such B_{0n}, by the definition of f_0, there exists r_1^n, $r_2^n \in B_{0n}$ such that $p \in [f_0'(r_1^n), f_0'(r_2^n)]$ and $\operatorname{dist}(p, f_0'(B_{0n})) = d$. For any such B_{0n} and $y \in B_{0n}$, $x^* \in \partial^- F_1(y)$, we have $\|x^*\| \leqslant 2kM < d/2$. Since f_0 is continuously differentiable on B_{0n} we have

$$\partial^- f(y) = f_0'(y) + \partial^- F_1(y) \quad \text{for all } y \in B_{0n}.$$

Thus $\operatorname{dist}(p, \partial^- f(B_{0n})) > d/2$. On the other hand,

$$\partial^- f(r_i^n) = f_0'(r_i^n) + \partial^- F_1(r_i^n) \quad \text{for } i = 1,2.$$

This implies that there exists a q such that

$$\|q - p\| < d/2 \quad \text{and} \quad q \in \operatorname{conv}(\partial^- f(B_{0n})).$$

Therefore $q \notin \partial^- f(B_{0n})$. This shows $\partial^- f(B_{0n})$ is not convex.

To conclude the proof, it suffices to consider any $x \in S_N \setminus S_{N+1}$ for some N. We have

$$\partial_a f(x) = \sum_{n=0}^{N} k^n f_n'(x) + \partial_a F_{N+1}(x).$$

Since $\partial_a F_{N+1}(x)$ is not convex neither is $\partial_a f(x)$. Hence $\partial_a f(x)$ is not convex for any $x \notin \bigcap_{n=0}^{\infty} S_n$. □

We note that even though $\{x\colon\ \partial_a f(x)$ is convex$\}$ is a Lebesgue-null set, it is still big in category, that is a residual set.

4. Vector-Valued Lipschitz Functions

In this section we use vector-valued bump functions with prescribed gradient images to construct differentiable vector-valued Lipschitz functions with non-convex approximate Jacobians and vector-valued Lipschitz functions with disconnected approximate Jacobians.

4.1. A SMOOTH BUMP FUNCTION

In this section we generalize Volterra's real-valued example [5, p. 45] to obtain a vector-valued differentiable bump function with prescribed gradient image. To do this, we take any open interval $(a, b) \subset (0, 1)$ and define $f := (f_1, f_2)$ as follows:

$$f_1(x) := \begin{cases} 0, & \text{if } x = a, \\ (x-a)^2 \sin \frac{1}{(x-a)}, & \text{if } a < x \leqslant c, \\ (c-a)^2 \sin \frac{1}{(c-a)}, & \text{if } c < x \leqslant b+a-c, \\ (x-b)^2 \sin \frac{1}{(b-x)}, & \text{if } b+a-c < x < b, \\ 0, & \text{if } x = b, \end{cases}$$

$$f_2(x) := \begin{cases} 0, & \text{if } x = a, \\ (x-a)^2 \cos \frac{1}{(x-a)}, & \text{if } a < x \leqslant d, \\ (d-a)^2 \cos \frac{1}{(d-a)}, & \text{if } d < x \leqslant b+a-d, \\ (x-b)^2 \cos \frac{1}{(b-x)}, & \text{if } b+a-d < x < b, \\ 0, & \text{if } x = b, \end{cases}$$

where $c < d < (a+b)/2$ with

$$2(c-a) \sin \frac{1}{c-a} - \cos \frac{1}{c-a} = 0$$

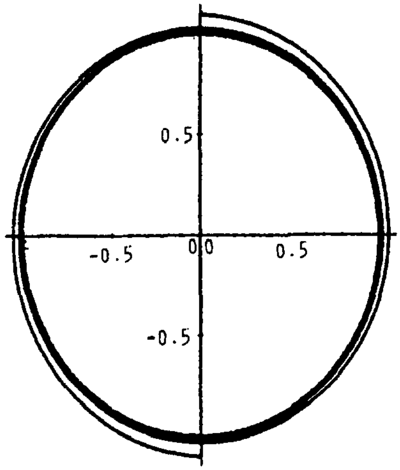

Figure 1.

and

$$2(d-a)\cos\frac{1}{d-a}+\sin\frac{1}{d-a}=0.$$

Then f_1 and f_2 are differentiable on $[a,b]$. When $a<x<c$ we have

$$f_1'(x):=2(x-a)\sin\frac{1}{x-a}-\cos\frac{1}{x-a},$$

$$f_2'(x):=2(x-a)\cos\frac{1}{x-a}+\sin\frac{1}{x-a},$$

with $(f_1'(x))^2+(f_2'(x))^2=1+4(x-a)^2$. When $b+a-c<x<b$ we have

$$f_1'(x):=-2(b-x)\sin\frac{1}{b-x}+\cos\frac{1}{b-x},$$

$$f_2'(x):=-2(b-x)\cos\frac{1}{b-x}-\sin\frac{1}{b-x},$$

with $(f_1'(x))^2+(f_2'(x))^2=4(b-x)^2+1$. More descriptively, $f'([a,b])$ looks like Figure 1. In Figure 1 we have set $a=0$, $b=1$, $c=0.2339300429$, and $d=0.4067166529$. It is the gradient image of this bump function that leads us, in the sequel, to construct vector-valued Fréchet differentiable functions with prescribed Mordukhovich–Ioffe coderivatives and Clarke generalized Jacobians.

4.2. THE CONSTRUCTION OF NONCONVEX APPROXIMATE JACOBIANS

4.2.1. *Nonconvexity Except on a Set of Small Measure*

Let $f\colon R\to R$ be an everywhere differentiable function. The Darboux property shows that the approximate subdifferential and the Clarke subdifferential are the same. However this is no longer true for a differentiable function F: $R \to R^2$. We can use the singularity of f given in Section 4.1 to construct a Fréchet differentiable vector-valued function whose approximate Jacobian, Dini Jacobian, and Clarke Jacobian are different on a set with large measure, which in turn implies that the approximate coderivative and contingent coderivative are different on a set with large measure.

To make these comments precise, we take a Cantor set $C \subset [0,1]$ with $\mu(C) > 0$. Then $[0,1] \setminus C$ is the union of a sequence of disjoint open intervals (a_n, b_n), $n = 1, 2, \ldots$. For each n define a bump function $\widetilde{F}_n$: $[a_n, b_n] \to R^2$ (like f in Section 4.1) by setting

$$\widetilde{F}_n(x) := \begin{cases} (f_{1\,n}(x), f_{2\,n}(x)), & \text{if } a_n < x < b_n, \\ (0,0), & \text{otherwise,} \end{cases}$$

where, for $i = 1, 2$, $f_{i\,n} := f_i$ with a and b replaced by a_n and b_n respectively. Define F: $[0,1] \to R^2$ by

$$F(x) := \begin{cases} \widetilde{F}_n(x), & \text{if } x \in (a_n, b_n), \\ (0,0), & \text{otherwise.} \end{cases}$$

We shall first show that F is Fréchet differentiable at every $x \in [0,1]$ and $F'(x) = (0,0)$ for every $x \in C$. Set $F := (F_1, F_2)$. It suffices to consider F_1. Given $\varepsilon > 0$ there is $\delta > 0$ such that for all n, if $b_n - a_n < \delta$, we have

$$\left|\frac{f_{1\,n}(x)}{x - a_n}\right| < \varepsilon \quad \text{and} \quad \left|\frac{f_{1\,n}(x)}{x - b_n}\right| < \varepsilon, \tag{4}$$

for all $x \in (a_n, b_n)$. For fixed $c \in C$ and $x \neq c$, consider

$$H(x) := \left|\frac{F_1(x) - F_1(c)}{x - c}\right|.$$

Suppose first that $0 \leqslant c < 1$ and let $\lambda := \limsup_{x \to c^+} H(x)$. If $c = a_n$ for some n, then clearly $\lambda = 0$. Otherwise there is a decreasing sequence of points x_ν in $(c, 1)$ converging to c such that $H(x_\nu) \to \lambda$. If infinitely many of the points x_ν are in C, then $\lambda = 0$. If not, then for all sufficiently large ν, $x_\nu \notin C$ and, consequently, $x_\nu \in (a_{n_\nu}, b_{n_\nu})$, where $a_{n_\nu} \to c$ and $b_{n_\nu} \to c$ as $\nu \to \infty$. In this case, for all sufficiently large ν we have $b_{n_\nu} - a_{n_\nu} < \delta$ and thus, by (4)

$$H(x_\nu) = \left|\frac{F_1(x_\nu)}{x_\nu - c}\right| \leqslant \left|\frac{f_{1n_\nu}(x_\nu)}{x_\nu - a_{n_\nu}}\right| < \varepsilon.$$

It follows that $0 \leqslant \lambda \leqslant \varepsilon$ and, hence, that $\lambda = 0$. This shows that, for $0 \leqslant c < 1$, $H(x) \to 0$ as $x \to c^+$, and likewise we get, for $0 < c \leqslant 1$, that $H(x) \to 0$ as $x \to c^-$. Therefore F_1 is differentiable at all points of C and similarly so also is F_2. Thus F is differentiable on $[0,1]$ with $F'(c) = (0,0)$ for all $c \in C$.

It may now be demonstrated what the approximate Jacobian of F looks like at each $c \in C$. Observe that if $c \in C$, then from the manner in which C is constructed, in each neighborhood of c there exists an interval (a_n, b_n) of the complement of C. Thus

$$\limsup_{n\to\infty} F'([a_n, b_n]) = \partial_{aJ}F(c).$$

Set $U := \{(x,y)\colon\ x^2 + y^2 = 1\} \cup \{(0,y)\colon\ -1 \leqslant y \leqslant 1\}$. Note that all cluster points of $F'([a_n, b_n])$ as $n \to \infty$ must either lie outside of the unit circle or in U. Since $U \subset F'([a_n, b_n])$ for all n, we have

$$U \subset \partial_{aJ}F(c) \quad \text{for all } c \in C, \tag{5}$$

which implies that $\partial_{aJ}F(c)$ is not convex for each $c \in C$. Furthermore, $\partial^- F(x) = \{JF(x)\} = \{(0,0)\} \neq \partial_{aJ}F(x)$ for every $x \in C$, and thus the Dini Jacobian and the approximate Jacobian of F are different on C. Hence, the contingent coderivative and approximate coderivative of F are different on C for every nonzero direction $y^* \in R^2$. Now for every $x \in C$ the generalized Clarke Jacobian set of F is

$$\operatorname{conv}\Big(\limsup_{n\to\infty} F'([a_n, b_n])\Big),$$

which implies $\operatorname{conv}(U) \subset \partial_{cJ}F(x)$ if $x \in C$. We formulate this discussion as a theorem.

THEOREM 6. *For any positive $\varepsilon < 1$ there exists a Fréchet differentiable and Lipschitz vector-valued function on $[0,1]$ such that*

$$0 < \mu(\{x \in [0,1]\colon\ \partial_{aJ}F(x) = \partial_{cJ}F(x)\}) < \varepsilon,$$

and

$$0 < \mu(\{x \in [0,1]\colon\ \partial^- F(x) = \partial_{aJ}F(x)\}) < \varepsilon.$$

Take an F_σ set $A \subset C$ with $\mu(A) = \mu(C) > 0$ and $d(A,x) = 1$ for all $x \in A$. By Proposition 2, we can find an approximately continuous $b\colon\ [0,1] \to R$ with $0 < b(x) \leqslant 1$ if $x \in A$ and $b(x) = 0$ if $x \notin A$. Moreover, b is also upper semi-continuous. Define $h(x) := \int_0^x b(s)\,\mathrm{d}s$. Then $h'(x) = b(x)$ for every $x \in [0,1]$. Now let $H\colon\ R \to R^2$ be defined by $H(x) := (0, h(x))$. It follows that $\partial_{aJ}(H + F)(x) = \partial_{aJ}F(x)$ for every $x \in [0,1]$. Furthermore, let $\overline{H}\colon\ R \to R^2$ be defined by $\overline{H}(x) := (h(x), 0)$. It follows that $\partial_{aJ}(\overline{H} + F)(x) \neq \partial_{aJ}F(x)$ for each $x \in A$, but $\partial_{cJ}F(x) = \partial_{cJ}(\overline{H} + F)(x)$ for each $x \in [0,1]$. We summarize these results as two corollaries.

COROLLARY 1. *Uncountably many Fréchet differentiable vector-valued Lipschitz functions, differing by more than constants, can share the same approximate Jacobians and thus the same approximate coderivatives. The same is true for the Clarke generalized Jacobians.*

COROLLARY 2. *For any given positive $\varepsilon < 1$ there exist two vector-valued Fréchet differentiable functions F, F_ε defined on $[0,1]$ such that*

$$\partial_{cJ}F(x) = \partial_{cJ}F_\varepsilon(x) \quad \textit{for all } x \in [0,1],$$

and

$$0 < \mu(\{x \in [0,1]\colon\ \partial_{aJ}F(x) = \partial_{aJ}F_\varepsilon(x)\}) < \varepsilon.$$

Note that our construction works, of course, for each nowhere dense closed set, whether or not it is perfect and whether or not it has positive measure.

4.2.2. *Almost Everywhere Nonconvexity*

In this section we sharpen Theorem 6 by showing that $\partial_{aJ}F$ can be nonconvex almost everywhere for vector-valued differentiable Lipschitz functions. To show this, let $A := \bigcap_{n=0}^{\infty} U_n$ be a dense G_δ set in $[0,1]$ with $U_{n+1} \subset U_n$. Then $U_n^c := [0,1] \setminus U_n$ is nowhere dense and closed. By Proposition 8 [19, p. 42] we can write $U_n := \bigcup_{l=1}^{\infty}(a_l^n, b_l^n)$ where $(a_i^n, b_i^n) \cap (a_j^n, b_j^n) = \emptyset$ if $i \neq j$. Moreover, we require that the U_n's satisfy: $\bigcup_{i=1}^{\infty}[a_i^n, b_i^n] \subset U_{n-1}$ for every $n \geqslant 1$. With each U_n, we associate a vector-valued Fréchet differentiable $F_n\colon R \to R^2$ in exactly the same way as in Section 4.2.1. For each n, we know $\|F_n'(x)\| \leqslant \sqrt{5}$ for every $x \in [0,1]$ and by (5) we see that $U \subset \partial_{aJ}F_n(x)$ if $x \in U_n^c$. Let $p := (1/2, 0)$. Then $\operatorname{dist}(p, U) = 1/2$ and $p \in \operatorname{conv}(F_n'([a_l^n, b_l^n]))$ for every n and l. Set $0 < k < 1/8\sqrt{5}$ and define

$$H(x) := \sum_{n=0}^{\infty} k^n F_n(x),$$

and

$$H_m(x) := \sum_{n=m}^{\infty} k^n F_n(x).$$

Then H and H_m are Fréchet differentiable on $[0,1]$. In particular, for each m we have $H_m'(x) = (0,0)$ for every $x \in U_m^c$ and

$$\partial_{aJ}H_m(x) = \limsup_{l\to\infty} H_m'([a_l^m, b_l^m]) \cup \{(0,0)\} \quad \text{for every } x \in U_m^c.$$

We claim that $\partial_{aJ}H_m(x)$ is nonconvex if $x \in U_m^c$. It suffices to show this only for $m = 0$ since the other cases are similar. We have

$$H'(x) = F_0'(x) + \sum_{n=1}^{\infty} k^n F_n'(x) \quad \text{for each } x \in [a_l^0, b_l^0].$$

Based on the construction of F_0, we have $\text{dist}(p, F_0'([a_l^0, b_l^0])) = 1/2$ which implies

$$\text{dist}(p, H'([a_l^0, b_l^0])) \geqslant \text{dist}(p, F_0'([a_l^0, b_l^0])) - 2k\sqrt{5} > \tfrac{1}{4} \quad \text{for each } l.$$

Then

$$\text{dist}(p, \partial_{aJ} H(x)) \geqslant \tfrac{1}{4} \quad \text{for } x \in U_0^c.$$

On the other hand, since $p = \lambda^l F_0'(r_1^l) + (1 - \lambda^l) F_0'(r_2^l)$ for some r_1^l, $r_2^l \in [a_l^0, b_l^0]$ with $\lambda^l \in (0,1)$ and $H'(r_i^l) = F_0'(r_i^l) + H_1'(r_i^l)$ for $i = 1, 2$, we derive that there is q_l such that

$$\|p - q_l\| < 2k\sqrt{5} \quad \text{and} \quad q_l \in \text{conv}(H'([a_l^0, b_l^0])).$$

Since $\{q_l\}$ is bounded, by taking a cluster point q of this sequence, we get

$$\|q - p\| \leqslant 2k\sqrt{5} < \tfrac{1}{4} \quad \text{and} \quad q \in \text{conv}(\partial_{aJ} H(x)).$$

Therefore $q \notin \partial_{aJ} H(x)$. This shows that $\partial_{aJ} H(x)$ is not convex if $x \in U_0^c$.

To conclude the proof, we need to consider any $x \in U_N \setminus U_{N+1}$ for some N. In this case we have

$$\partial_{aJ} H(x) = \sum_{n=0}^{N} k^n F_n'(x) + \partial_{aJ} H_{N+1}(x).$$

Since $\partial_{aJ} H_{N+1}(x)$ is not convex neither is $\partial_{aJ} H(x)$. To summarize

$$\{x \in [0,1]\colon\ \partial_{aJ} H(x) = \text{conv}(\partial_{aJ} H(x))\} \subset A.$$

Letting $\mu(A) = 0$, we get

THEOREM 7. *There exists a Fréchet differentiable vector-valued Lipschitz function H: $R \to R^2$ such that $\mu(\{x \in [0,1]\colon\ \partial_{aJ} H(x)$ is convex$\}) = 0$.*

Note that Baire's Theorem shows H' is generically continuous. Thus $\partial_{aJ} H$ is single-valued generically and $\partial_{aJ} H$ and $\partial_{cJ} H$ agree generically. Our construction also shows that the Dini Jacobian and the approximate Jacobian can be different almost everywhere which in turn implies that the contingent coderivative and the approximate coderivative can be different almost everywhere for every nonzero direction $y^* \in R^2$.

4.3. A SMOOTH BUMP WITH DISCONNECTED GRADIENT IMAGE

In Section 4.1 we constructed a differentiable vector-valued bump whose gradient image is connected. In this section we shall present a nonnegative vector-valued (i.e. each coordinate function is nonnegative valued) differentiable bump whose gradient image is disconnected. Note that for any topological spaces X and Y if $f\colon X \to Y$ is continuous then A is connected implies $f(A)$ is connected. This necessitates the gradient map to be discontinuous. Define $F\colon R \to R^2$ by

$$F(x) := \begin{cases} (f_1(x), f_2(x)), & \text{if } 0 < x < 1, \\ (0,0), & \text{else,} \end{cases}$$

where

$$f_1(x) := x^2(1-x)^2 \sin \frac{1-2x}{x(1-x)},$$

and

$$f_2(x) := x^2(1-x)^2 \cos \frac{1-2x}{x(1-x)}.$$

Then F is differentiable for all x, has support $[0,1]$, and $F'(0) = F'(1) = F(0) = F(1) = (0,0)$. Next we have, for $0 < x < 1$,

$$\begin{aligned} f_1'(x) &:= 2x(1-x)(1-2x) \sin \frac{1-2x}{x(1-x)} - \\ &\quad - (x^2 + (1-x)^2) \cos \frac{1-2x}{x(1-x)}, \\ f_2'(x) &:= 2x(1-x)(1-2x) \cos \frac{1-2x}{x(1-x)} + \\ &\quad + (x^2 + (1-x)^2) \sin \frac{1-2x}{x(1-x)}, \end{aligned}$$

and hence

$$f_1'(x)^2 + f_2'(x)^2 = 4x^2(1-x)^2(1-2x)^2 + ((1-x)^2 + x^2)^2 \geqslant \tfrac{1}{4}.$$

Since $F'(x) = (0,0)$ when $x \in R \setminus (0,1)$, it follows that the range of F' is not connected (see Figure 2).

We proceed to make the bump nonnegative. Define $\tilde{F}(x) := (\tilde{f}_1(x), \tilde{f}_2(x))$ by setting $\tilde{f}_1(x) := f_1(x) + 1$ and $\tilde{f}_2(x) := f_2(x) + 1$ in the interval $[0,1]$ and then constructing $\tilde{f}_1$ and $\tilde{f}_2$ outside this interval in such a way that they are nonnegative, differentiable, zero outside $(-a, a)$ for some sufficiently large a, and have

$$\sqrt{\tilde{f}_1{}'(x)^2 + \tilde{f}_2{}'(x)^2} < \frac{1}{4} \qquad \text{for } x \in R \setminus (0,1).$$

For example, to define $\tilde{f}_1$ on the interval $(1, +\infty)$ we take

$$a > 1 + 8\sqrt{2}, \quad b := \frac{2}{(a-1)^2}$$

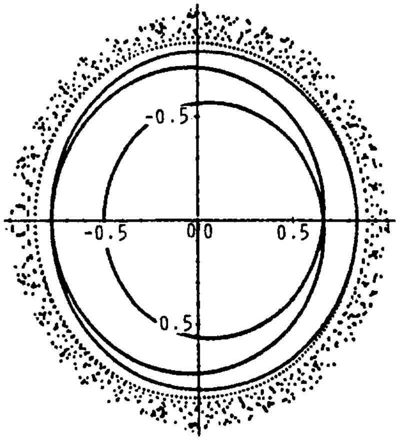

Figure 2.

and set

$$\tilde{f}_1(x) := \begin{cases} 1 - b(x-1)^2, & \text{if } 1 < x < \frac{1+a}{2}, \\ b(a-x)^2, & \text{if } \frac{1+a}{2} \leqslant x \leqslant a, \\ 0, & \text{if } x > a, \end{cases} \tag{6}$$

whence, for $x > 1$,

$$\tilde{f}_1{}'(x)^2 \leqslant b^2(a-1)^2 = \frac{4}{(a-1)^2} < \frac{1}{32}.$$

One can do the same process to get $\tilde{f}_2{}'(x)^2 \leqslant 1/32$ as $x > 1$. Similarly for the end point 0. This yields a differentiable, nonnegative $\widetilde{F}$ with support in $[-a, a]$. Moreover, $\widetilde{F}'([-a, a]) = A_1 \cup A_2$, where

$$A_1 := \left\{(\tilde{f}_1{}'(x), \tilde{f}_2{}'(x)) \colon \sqrt{\tilde{f}_1{}'(x)^2 + \tilde{f}_2{}'(x)^2} < \tfrac{1}{4}\right\},$$

$$A_2 := \left\{(\tilde{f}_1{}'(x), \tilde{f}_2{}'(x)) \colon \sqrt{\tilde{f}_1{}'(x)^2 + \tilde{f}_2{}'(x)^2} \geqslant \tfrac{1}{2}\right\},$$

which are disconnected. Hence

THEOREM 8. *There is a nonnegative, differentiable function F: $R \to R^2$ with compact support whose derivative has disconnected range.*

Remark 2. Let $G(t) := (G_1(t), G_2(t)) = \frac{1}{2a}\widetilde{F}((2t-1)a)$. Then G is a nonnegative bump with support in $[0, 1]$ and $G'([0, 1]) = \widetilde{F}'([-a, a])$.

Remark 3. Our example implies the following general fact: Let $a_1 < a_2$. Define F: $R \to R^2$ by setting $F(x) := (f_1(x), f_2(x))$ with

$$f_1(x) := \begin{cases} (x-a_1)^2(x-a_2)^2 \sin\left(\frac{1}{x-a_1} + \frac{1}{x-a_2}\right), & \text{if } a_1 < x < a_2, \\ 0, & \text{otherwise,} \end{cases}$$

$$f_2(x) := \begin{cases} (x-a_1)^2(x-a_2)^2 \cos\left(\frac{1}{x-a_1} + \frac{1}{x-a_2}\right), & \text{if } a_1 < x < a_2, \\ 0, & \text{otherwise.} \end{cases}$$

Then F is differentiable for all x, has support $[a_1, a_2]$, and for $a_1 < x < a_2$ we have

$$f_1'(x)^2 + f_2'(x)^2 \geqslant \frac{(a_1 - a_2)^4}{4},$$

which implies that the range of F' is disconnected.

Remark 4. Recently Malý [14] showed that the gradient map of every Fréchet differentiable real-valued function on a Banach space maps every closed convex set with nonempty interior into a connected set in the dual space. Our example shows that this is not true for vector-valued Fréchet differentiable bump functions.

4.4. THE CONSTRUCTION OF DISCONNECTED APPROXIMATE JACOBIANS

4.4.1. *Disconnectedness Except on a Set of Small Measure*

In this section we construct a vector-valued Lipschitz function whose approximate Jacobian has disconnected image on a set with positive measure. Let $C \subset [0,1]$ be a Cantor set with $\mu(C) > 0$. Then $U := [0,1] \setminus C = \bigcup_{n=1}^{\infty}(a_n, b_n)$ where $[a_i, b_i] \cap [a_j, b_j] = \emptyset$ if $i \neq j$. On each (a_n, b_n), we construct a function F_n like G given in Section 4.3. More precisely, we define F_n: $R \to R^2$ as follows:

$$F_n(x) := \begin{cases} (b_n - a_n) \cdot G\left(\frac{x-a_n}{b_n-a_n}\right), & \text{if } a_n < x < b_n, \\ (0,0), & \text{otherwise.} \end{cases}$$

Define $F := (f_1, f_2)$ by setting:

$$F(x) := \begin{cases} F_n(x), & \text{if } x \in U, \\ (0,0), & \text{otherwise.} \end{cases}$$

If $x \in U$ then $x \in (a_n, b_n)$ for some n, so F is differentiable around x, with

$$\partial^- F(x) = G'\left(\frac{x-a_n}{b_n-a_n}\right).$$

If $x \notin U$, then $f_1(x) = 0$ so x is a local minimum. By the construction $f_1^-(x; v) \geqslant 0$ for any v. Fixing a direction v, there are two possibilities: Either there is a decreasing sequence of positive numbers $t_n \to 0$ such that $x + t_n v \notin U$ or there is an $\varepsilon > 0$ such that $x + tv \in U$ for all $t \in (0, \varepsilon)$. In the first case we have

$$f_1^-(x, v) \leqslant \liminf_{n\to\infty} \frac{f_1(x + t_n v) - f_1(x)}{t_n} = 0.$$

Since we already know that $f_1^-(x; v) \geqslant 0$, we have $f_1^-(x; v) = 0$. In the second case, noting that it is impossible to express $[0, \varepsilon]$ as a union of disjoint closed intervals of positive length less than ε (see [22, p. 112]), there must be some n such that $x + tv \in [a_n, b_n]$ for all $t \in (0, \varepsilon)$. Since $x \notin U$, we know either $x = a_n$ or $x = b_n$. Suppose $x = a_n$. Using the fact that $x + tv \in [a_n, b_n]$ for small positive t we get $f_1^-(x; v) = G_1'(0) \cdot v = 0$. Therefore in each case when $x \notin U$ we get $f_1^-(x; v) = 0$ for any v, so $\partial^- f_1(x) = 0$. Similarly $\partial^- f_2(x) = 0$ if $x \notin U$. For every $x \in [0, 1]$ we have

$$\partial_{aJ} F(x) \subset G'([0, 1]) \cup \{(0, 0)\}.$$

If $x \notin U$, and $N(x, \delta)$ is any neighborhood of x then there is some n such that $[a_n, b_n] \subset N(x, \delta)$. Therefore $G'([0, 1]) \subset \partial_{aJ} F(x)$. So for any $x \notin U$ we have $\partial_{aJ} F(x) = G'([0, 1])$, which is disconnected. From the construction, we know that $\partial_a^\infty f_1(x) = \partial_a^\infty f_2(x) = 0$ for all $x \in [0, 1]$, so f_1 and f_2 are Lipschitz. Now consider the approximate coderivative of F. Since f_1 and f_2 are nonnegative and f_1 and f_2 have the same structure for fixed $y^* := (y_1^*, y_2^*) \in (R^2)^+$, the positive orthant, we have

$$(\langle y^*, F\rangle)^-(x; v) = y_1^* f_1^-(x; v) + y_2^* f_2^-(x; v) \quad \text{for every } v \in R.$$

Therefore

$$\partial_a(\langle y^*, F\rangle)(x) = \langle y^*, \partial_{aJ} F(x)\rangle.$$

That is, for fixed x the approximate coderivative of F in the direction $y^* \in (R^2)^+$ is determined by $\partial_{aJ} F$. Moreover, since the generalized Clarke Jacobian is not sensitive to null sets we have $\partial_{cJ} F(x) = \text{clconv}(G'([0, 1]))$ if $x \notin U$. Letting $1 > \mu(C) > 1 - \varepsilon$, we get

THEOREM 9. *Let $0 < \varepsilon < 1$. Then there exists a vector-valued Lipschitz function F: $[0, 1] \to R^2$ such that*

$$0 < \mu(\{x \in [0, 1]\colon\ \partial_{aJ} F(x) \text{ is connected}\}) < \varepsilon.$$

By (6) we have

$$\left\{\left(\frac{4(x - a)}{(a - 1)^2}, \frac{4(x - a)}{(a - 1)^2}\right)\colon\ \frac{1 + a}{2} \leqslant x \leqslant a\right\} \subset G'([0, 1]).$$

Take an F_σ set $A \subset C$ with $\mu(A) = \mu(C) > 0$ and $d(A, x) = 1$ for all $x \in A$. By Proposition 2, we can find an approximately continuous b: $[0,1] \to R$ with $0 < b(x) \leqslant 1$ if $x \in A$ and $b(x) = 0$ if $x \notin A$. Define $h(x) := \int_0^x b(s)\,ds$. Then $h'(x) = b(x)$ for every $x \in [0,1]$. Now define H: $R \to R^2$ by $H(x) := (-h(x), -h(x))$. It follows that for sufficiently small $\varepsilon > 0$ we have

$$\partial_{aJ}(F + \varepsilon H)(x) = \partial_{aJ} F(x) \quad \text{for all } x \in [0,1].$$

Observe that $\partial_{aJ}(F + \varepsilon H)(x)$ is still disconnected if $x \in C$. It is therefore all the more surprising, that the following corollary holds.

COROLLARY 3. *Uncountably many vector-valued Lipschitz functions, differing by more than constants, can share the same approximate Jacobian map whose values are disconnected except on a set of small measure.*

4.4.2. *Almost Everywhere Disconnectedness*

In this section we sharpen Theorem 9 by showing that for any $y^* \in (R^2)^+$ the map ∂_{aJ} associated with the approximate coderivative can be disconnected almost everywhere.

THEOREM 10. *There is a vector-valued Lipschitz function H: $[0,1] \to R^2$ such that*

$$\mu(\{x \in [0,1]\colon\ \partial_{aJ} H(x) \text{ is connected}\}) = 0.$$

Proof. Construct sequences of closed subintervals $\{[a_{n\,i}, b_{n\,i}]\}$ in $[0,1]$ such that

(i) For fixed n, $[a_{n\,i}, b_{n\,i}] \cap [a_{n\,j}, b_{n\,j}] = \emptyset$ if $i \neq j$;
(ii) $\bigcup_{i=1}^{\infty} [a_{n\,i}, b_{n\,i}]$ is dense and contained in $U_{n-1} := \bigcup_{i=1}^{\infty} (a_{n-1\,i}, b_{n-1\,i})$ for every $n \geqslant 1$;
(iii) $\mu(U_n) \to 0$ as $n \to \infty$;
(iv) $\bigcup_{i=1}^{\infty} (a_{0\,i}, b_{0\,i})$ is dense in $[0,1]$.

Now for fixed n on each $[a_{n\,i}, b_{n\,i}]$ we define

$$f_{n\,i}(x) := \begin{cases} (b_{n\,i} - a_{n\,i}) \cdot G\left(\frac{x - a_{n\,i}}{b_{n\,i} - a_{n\,i}}\right), & \text{if } a_{n\,i} < x < b_{n\,i}, \\ 0, & \text{otherwise,} \end{cases}$$

and

$$F_n(x) := \begin{cases} f_{n\,i}(x), & \text{if } x \in (a_{n\,i}, b_{n\,i}), \\ 0, & \text{otherwise.} \end{cases}$$

Let

$$M := \sup\{\|x^*\|:\ x^* \in G'([0,1])\} \quad \text{and} \quad 0 < k < \min\{\tfrac{1}{2}, \tfrac{1}{16M}\}.$$

Set

$$H(x) := \sum_{n=0}^{\infty} k^n F_n(x) \quad \text{and} \quad H_m(x) := \sum_{n=m}^{\infty} k^n F_n(x).$$

Step 1. If $x \notin U_m$ we have $\partial_{aJ} H_m(x) = \partial^- H_m([a_{m\,i}, b_{m\,i}])$. To show this, let $H_m(x) := (H_m^1(x), H_m^2(x))$. If $x \notin U_m$ then $\partial^- H_m(x) = \{(0,0)\}$. In fact, $H_m^1(x) = 0$ and, therefore, x is a local minimum of H_m^1 and $(H_m^1)^-(x; v) \geqslant 0$ for all v. Fixing a direction v, there are two possibilities: Either there exists a sequence of $t_n \downarrow 0$ such that $x + t_n v \notin U_m$ or there exists an $\varepsilon > 0$ such that $x + tv \in U_m$ for all $t \in (0, \varepsilon)$. In the first case $(H_m^1)^-(x; v) = 0$. In the second case, there exists an i such that $x + tv \in [a_{m\,i}, b_{m\,i}]$ for all $t \in (0, \varepsilon)$ and we obtain $x = a_{m\,i}$ or $x = b_{m\,i}$. Suppose $x = a_{m\,i}$. Then $x \notin U_n$ for $n \geqslant m + 1$. Therefore, there must be $t_\nu \downarrow 0$ such that $x + t_\nu v \notin U_{m+1}$. Thus

$$\begin{aligned}
(H_m^1)^-(x; v) &= \liminf_{t \to 0^+} \frac{H_m^1(x + tv) - H_m^1(x)}{t} \\
&\leqslant \liminf_{\nu \to \infty} \frac{H_m^1(x + t_\nu v) - H_m^1(x)}{t_\nu} \\
&= k^m \liminf_{\nu \to \infty} \frac{F_m^1(x + t_\nu v) - F_m^1(x)}{t_\nu} \\
&= k^m G_1'(0) \cdot v = 0.
\end{aligned}$$

The same is true when $x = b_{m\,i}$. Therefore, $(H_m^1)^-(x; v) = 0$ for all v, which is to say $\partial^- H_m^1(x) = \{0\}$. Similarly, $\partial^- H_m^2(x) = \{0\}$ if $x \notin U_m$. For every $x \in [0,1]$ and every positive integer i we have

$$\partial_{aJ} H_m(x) \subset \partial^- H_m([a_{m\,i}, b_{m\,i}]) \cup \{(0,0)\} = \partial^- H_m([a_{m\,i}, b_{m\,i}]).$$

If $x \notin U_m$ and $N(x, \delta)$ is any neighborhood of x, then $N(x, \delta)$ contains $[a_{m\,i}, b_{m\,i}]$ for some i. Therefore $\partial^- H_m([a_{m\,i}, b_{m\,i}]) \subset \partial_{aJ} H_m(x)$.

Step 2. If $x \notin U_m$, then $\partial_{aJ} H_m(x)$ is disconnected. Since all H_m have the same structure we prove this only for $m = 0$. By Step 1, we deduce that $\partial_{aJ} H(x) = \partial^- H([a_{0\,i}, b_{0\,i}])$. For every $y \in (a_{0\,i}, b_{0\,i})$ and $x^* \in \partial^- H_1(y)$, we have $\|x^*\| \leqslant 2kM$. Since F_0 is differentiable on $[a_{0\,i}, b_{0\,i}]$ we have

$$\partial^- H(y) = F_0'(y) + \partial^- H_1(y) \quad \text{for all } y \in [a_{0\,i}, b_{0\,i}].$$

Since $F_0'([a_{0\,i}, b_{0\,i}]) = G'([0,1]) = A_1 \cup A_2$, where

$$A_1 := \left\{(\tilde{f}_1{}'(t), \tilde{f}_2{}'(t)):\ \sqrt{\tilde{f}_1{}'(t)^2 + \tilde{f}_2{}'(t)^2} < \tfrac{1}{4}\right\},$$

$$A_2 := \left\{(\tilde{f}_1{}'(t), \tilde{f}_2{}'(t)):\ \sqrt{\tilde{f}_1{}'(t)^2 + \tilde{f}_2{}'(t)^2} \geqslant \tfrac{1}{2}\right\},$$

we obtain $\partial^- H([a_{0\,i}, b_{0\,i}]) \subset B_1 \cup B_2$ with $\partial^- H([a_{0\,i}, b_{0\,i}]) \cap B_k \neq \emptyset$ for $k = 1, 2$, where

$$B_1 := \Big(A_1 + \partial^- H_1([a_{0\,i}, b_{0\,i}])\Big),$$
$$B_2 := \Big(A_2 + \partial^- H_1([a_{0\,i}, b_{0\,i}])\Big).$$

It suffices to show that B_1 and B_2 are disconnected. Indeed, for every $(x^*, y^*) \in B_1$ we have $(x^*, y^*) = (x^*_{A_1}, y^*_{A_1}) + (x^*_{H_1}, y^*_{H_1})$ with $(x^*_{A_1}, y^*_{A_1}) \in A_1$ and $(x^*_{H_1}, y^*_{H_1}) \in \partial^- H_1([a_{0\,i}, b_{0\,i}])$. This implies $\|(x^*, y^*)\| \leqslant \frac{1}{4} + 2kM$. On the other hand, for every $(x^*, y^*) \in B_2$ we have $(x^*, y^*) = (x^*_{A_2}, y^*_{A_2}) + (x^*_{H_1}, y^*_{H_1})$ with $(x^*_{A_2}, y^*_{A_2}) \in A_2$ and $(x^*_{H_1}, y^*_{H_1}) \in \partial^- H_1([a_{0\,i}, b_{0\,i}])$. This implies $\|(x^*, y^*)\| \geqslant \frac{1}{2} - 2kM$. Therefore $\mathrm{cl}(B_1) \cap \mathrm{cl}(B_2) = \emptyset$ because $k < 1/16M$.

To conclude the proof, it suffices to consider any $x \in U_N \backslash U_{N+1}$ for some N. Note that for $0 \leqslant n \leqslant N$ each F_n is continuously differentiable on $U_N \setminus U_{N+1}$ except for a countable set. It follows that

$$\partial_{aJ} H(x) = \sum_{n=0}^{N} k^n F_n'(x) + \partial_{aJ} H_{N+1}(x),$$

holds on $U_N \backslash U_{N+1}$ except for a countable set. Since $\partial_{aJ} H_{N+1}(x)$ is not connected neither is $\partial_{aJ} H(x)$. Hence, $\partial_{aJ} H(x)$ is not connected for any $x \notin \bigcap_{n=0}^{\infty} U_n$ except for a countable set. □

References

1. Aubin, J. P. and Frankowska, H.: *Set-Valued Analysis*, Birkhäuser, Boston, 1990.
2. Borwein, J. M.: Minimal CUSCOS and subgradients of Lipschitz functions, in: *Fixed Point Theory and Its Applications*, Pitman Research Notes 252, Longman, Harlow, 1991, pp. 57–81.
3. Borwein J. M. and Fitzpatrick, S. P.: Characterization of Clarke subdifferentials among one-dimensional multifunctions, CECM Preprint 94–006 (1994).
4. Borwein, D., Borwein, J. M. and Wang, X.: Approximate subgradients and coderivatives in R^n, CECM Preprint 96–058 (1996).
5. Bruckner, A. M.: *Differentiation of Real Functions*, Lecture Notes in Math., edited by A. Dold and B. Eckmann, Springer-Verlag, New York, 1978.
6. Burrill, C. W. and Knudsen, J. R.: *Real Variables*, Holt, Rinehart and Wintson, New York, 1969.
7. Clarke, F. H.: *Optimization and Nonsmooth Analysis*, Wiley Interscience, New York, 1983.
8. Hiriart-Urruty, J. and Lemarechal, C.: *Convex Analysis and Minimization Algorithms I*, Springer-Verlag, Berlin, Heidelberg, 1993.
9. Ioffe, A. D.: Approximate subdifferentials and applications 3: The metric theory, *Mathematika* **36**(71) (1989), 1–38.
10. Ioffe, A. D.: Approximate subdifferentials and applications I: The finite dimensional theory, *Trans. Amer. Math. Soc.* **281** (1984), 390–416.
11. Ioffe, A. D.: Calculus of Dini subdifferentials of functions and contingent coderivatives of set-valued maps, *Nonlinear Anal. Theory, Methods Appl.* **8** (1984), 517–539.
12. Katriel, G.: Are the approximate and the Clarke subgradients generically equal?, *J. Math. Anal. Appl.* **193** (1995), 588–593.

13. Lay, S. R.: *Convex Sets and Their Applications*, Wiley, New York, 1982.
14. Malý, J.: Darboux property of gradients, *Real Analysis Exchange* (to appear).
15. Mordukhovich, B. S.: Maximum principle in problems of time optimal control with nonsmooth constraints, *J. Appl. Math. Mech.* **40** (1976), 960–969.
16. Mordukhovich, B. S.: Metric approximations and necessary optimality conditions for general classes of nonsmooth extremal problems, *Soviet Math. Dokl.* **22** (1980), 526–530.
17. Mordukhovich, B. S.: *Approximation Methods in Problems of Optimization and Control*, Nauka, Moscow, 1988.
18. Mordukhovich, B. S.: Generalized differential calculus for nonsmooth and set-valued mappings, *J. Math. Anal. Appl.* **183** (1994), 250–288.
19. Royden, H. L.: *Real Analysis*, Macmillan, New York, 1988.
20. Rockafellar, R. T.: Proximal subgradients, marginal values, and augmented lagrangians in nonconvex optimization, *Math. Oper. Res.* **6**(3) (1981), 424–436.
21. Rockafellar, R. T.: Dualization of subgradient conditions for optimality, Gruppo di Ottimizzazione e Ricerca Operativa 3.182 (620), Gennaio, 1992.
22. Stromberg, K. R.: *An Introduction to Classical Real Analysis*, Wadsworth Internat. Math. Series, 1981.
23. Warga, J.: Fat homeomorphisms and unbounded derivate container, *J. Math. Anal. Appl.* **81** (1981), 545–560.