# Spaceability of Clarke-saturated functions

Aris Daniilidis

University of Chile

BANACH SPACES AND OPTIMIZATION
(for the 60 years of Robert Deville)

## CRITICALITY (SMOOTH CASE)

- $f : \mathbb{R}^n \to \mathbb{R}$ differentiable function

**Definition (Critical point)**

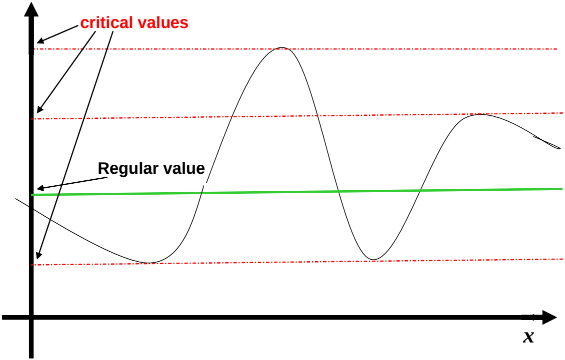
A point $x_0$ is called *critical*, if $df(x_0) \neq 0$.

- $S$ := set of critical points
- $f(S)$ := set of critical values

# MORSE-SARD THEOREM

**Theorem (Morse theorem)**

*If $f : \mathbb{R}^n \to \mathbb{R}$ is $C^k$ with $k \geq n$, then $m(f(S)) = 0$.*



**Theorem (Sard theorem)**

*For every $C^k$-smooth function $F : \mathbb{R}^n \to \mathbb{R}^m$ we have $m(F(S)) = 0$, provided $k \geq n - m + 1$.*

## Theorem (Sard theorem)

*For every $C^k$-smooth function $F : \mathbb{R}^n \to \mathbb{R}^m$ we have $m(F(S)) = 0$, provided $k \geq n - m + 1$.*

APPLICATIONS (DIFFERENTIAL GEOMETRY)

- Immersed submanifolds $\mathcal{N} \subset \mathcal{M}$ of positive co-dimension have zero measure in $\mathcal{M}$.
- (*Whitney embedding*) Every $C^\infty$ manifold of dimension $d$ can be embedded in $\mathbb{R}^{2d+1}$.

## APPLICATIONS (OPTIMIZATION)

- Constraint qualification condition (*genericity result*)

$$\begin{cases} \min \quad f(x) \\ h_i(x) = r_i \\ i \in \{1, \ldots, m\} \end{cases} \quad \longleftrightarrow \quad \begin{array}{c} H : \mathbb{R}^n \to \mathbb{R}^m \\ H = (h_1, \ldots, h_m) \end{array}$$

## APPLICATIONS (OPTIMIZATION)

- Constraint qualification condition (*genericity result*)

$$\left\{ \begin{array}{c} \min \quad f(x) \\ h_i(x) = r_i \\ i \in \{1, \dots, m\} \end{array} \right. \qquad \longleftrightarrow \qquad \begin{array}{c} H : \mathbb{R}^n \to \mathbb{R}^m \\ H = (h_1, \dots, h_m) \end{array}$$

The set of data $r = (r_1, \dots, r_m) \in \mathbb{R}^m$ for which the constraints do not satisfy the *qualification condition* at the solution $x_*$ is of measure zero.

## SHARPNESS OF THE THEOREM

**Theorem (Bates, 1993)**

*A $C^{n-1}$-smooth function $f : \mathbb{R}^n \to \mathbb{R}$ satisfies $m(f(S)) = 0$ if (and only if) $d^{(n-1)}f$ is locally Lipschitz.*

– Fails if $d^{(n-1)}f$ is $\beta$-Hölder with $\beta < 1$.

## SHARPNESS OF THE THEOREM

**Theorem (Bates, 1993)**

*A $C^{n-1}$-smooth function $f : \mathbb{R}^n \to \mathbb{R}$ satisfies $m(f(S)) = 0$ if (and only if) $d^{(n-1)}f$ is locally Lipschitz.*

– Fails if $d^{(n-1)}f$ is $\beta$-Hölder with $\beta < 1$.

## FAILURE IN INFINITE DIMENSIONS

**Example (Kupka, 1965)**

There exists a $C^\infty$ function $f : \mathcal{H} \to \mathbb{R}$ ($\mathcal{H}$ Hilbert) with $m(f(S)) > 0$.

## CLARKE SUBDIFFERENTIAL

- $f : \mathbb{R}^n \to \mathbb{R}$ (nonsmooth) Lipschitz

---

$D_f$ := differentiability points (dense by Rademacher)

$$\partial f(x_0) = \text{conv}\left\{ \lim_{x_n \to x_0} \nabla f(x_n) : \{x_n\}_n \subset D_f \right\}$$

## CLARKE SUBDIFFERENTIAL

- $f : \mathbb{R}^n \to \mathbb{R}$ (nonsmooth) Lipschitz

---

$D_f$ := differentiability points (dense by Rademacher)

$$\partial f(x_0) = \text{conv}\left\{ \lim_{x_n \to x_0} \nabla f(x_n) : \{x_n\}_n \subset D_f \right\}$$

## NONSMOOTH CRITICALITY

**Definition (Clarke critical point)**

A point $x_0$ is called *critical*, if $0 \in \partial f(x_0)$.

- If $f$ is $C^1$ then (it is locally Lipschitz and)

$$\partial f(x) = \{\nabla f(x)\}, \quad \text{for all } x \in \mathbb{R}^n$$

- No hope to treat the general nonsmooth case if $n > 1$ !

- If $f$ is $C^1$ then (it is locally Lipschitz and)

$$\partial f(x) = \{\nabla f(x)\}, \quad \text{for all } x \in \mathbb{R}^n$$

- No hope to treat the general nonsmooth case if $n > 1$ !

**Fact**

*If the Morse-Sard theorem fails for a smooth function, its critical set contains non-rectifiable components (Assouad embeddings).*

- If $f$ is $C^1$ then (it is locally Lipschitz and)

$$\partial f(x) = \{\nabla f(x)\}, \quad \text{for all } x \in \mathbb{R}^n$$

- No hope to treat the general nonsmooth case if $n > 1$ !

**Fact**

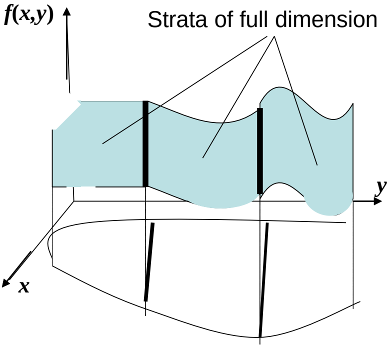
*If the Morse-Sard theorem fails for a smooth function, its critical set contains non-rectifiable components (Assouad embeddings).*

**Fact**

*The set of Clarke critical points usually has a nice structure (...but the nonsmooth chain rule fails.)*

## POSITIVE RESULTS

Assume that the graph of a Lipschitz function is a (finite disjoint) union of $C^k$ manifolds (strata), with $k \geq n$.



## CONTROL VIA STRUCTURE

Let $\mathcal{M}$ be a stratum of $\text{Graph}(f)$ and let $f_{\mathcal{M}}$ be the function (restriction of $f$) with $\text{Graph}(f_{\mathcal{M}}) = \mathcal{M}$.

**Fact (classical Morse theorem)**

*$f_{\mathcal{M}}$ satisfies the Morse-Sard theorem with respect to its (Riemann) gradient $\nabla f_{\mathcal{M}}$.*

## CONTROL VIA STRUCTURE

Let $\mathcal{M}$ be a stratum of Graph$(f)$ and let $f_{\mathcal{M}}$ be the function (restriction of $f$) with Graph$(f_{\mathcal{M}}) = \mathcal{M}$.

**Fact (classical Morse theorem)**

*$f_{\mathcal{M}}$ satisfies the Morse-Sard theorem with respect to its (Riemann) gradient $\nabla f_{\mathcal{M}}$.*

In particular, if $\mathcal{U}$ is a stratum of full dimension,

$$\partial f(x) = \{\nabla f(x)\} = \{\nabla f_{\mathcal{U}}(x)\}$$

- Nonsmoothness is localized at strata of lower dimension.

## CONTROL VIA STRUCTURE

Let $\mathcal{M}$ be a stratum of Graph($f$) and let $f_{\mathcal{M}}$ be the function (restriction of $f$) with Graph($f_{\mathcal{M}}$) $= \mathcal{M}$.

**Fact (classical Morse theorem)**

*$f_{\mathcal{M}}$ satisfies the Morse-Sard theorem with respect to its (Riemann) gradient $\nabla f_{\mathcal{M}}$.*

In particular, if $\mathcal{U}$ is a stratum of full dimension,

$$\partial f(x) = \{\nabla f(x)\} = \{\nabla f_{\mathcal{U}}(x)\}$$

- Nonsmoothness is localized at strata of lower dimension.
- Associate $\partial f(x)$ with the Riemann gradient $\nabla f_{\mathcal{M}}(x)$ ?

## CONTROL VIA STRUCTURE

**Fact**

*The graph of a semialgebraic (resp. tame) function stratifies into strata that fit together in a regular way (Whitney-(a) property).*

**Theorem (Projection formula)**

*If $f : \mathbb{R}^n \to \mathbb{R}$ is Lipschitz, Whitney stratifiable function, then*

$$\mathrm{Proj}_{T_{\mathcal{M}}(x)} [\partial f(x)] = \{\nabla f_{\mathcal{M}}(x)\}.$$

$$\left.\begin{array}{c} 0 \in \partial f(x) \\ x \in \mathcal{M} \end{array}\right\} \Longrightarrow \nabla f_{\mathcal{M}}(x) = 0$$

## CONTROL VIA STRUCTURE

**Corollary (Bolte, Daniilidis, Lewis, Shiota, 2007)**

*– If $f$ is $C^k$-Whitney stratifiable, $k \geq n$, the set of its Clarke critical values has measure zero.*

*– If $f$ is tame, then this set is locally finite.*

Extensions for multifunctions (Ioffe, 2009)

## CONTROL VIA CARDINALITY

Consider *continuous* selections $f : \mathbb{R}^n \to \mathbb{R}$ of the form

$$f(x) \in \{F(x,t) : t \in T\}, \quad \text{for all } x \in \mathbb{R}^n$$

where $F : \mathbb{R}^n \times T \to \mathbb{R}$

- $T$ compact countable
- $x \mapsto F(x,t)$ $C^k$-smooth ($k \geq n$)
- $F$, $\nabla_x F$ jointly continuous.

## CONTROL VIA CARDINALITY

Consider *continuous* selections $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ of the form

$$f(x) \in \{F(x,t) : t \in T\}, \quad \text{for all } x \in \mathbb{R}^n$$

where $F : \mathbb{R}^n \times T \longrightarrow \mathbb{R}$

- $T$ compact countable
- $x \mapsto F(x,t)$ $C^k$-smooth ($k \geq n$)
- $F$, $\nabla_x F$ jointly continuous.

**Theorem (Barbet, Dambrine, Daniilidis, 2013)**

*$f$ is Lipschitz and its set of Clarke critical values is null.*

**Fact**

*The proof uses the Morse–Sard theorem, but it also recovers it.*

## APPLICATION I: SEMI-INFINITE PROGRAMMING

Let $g_t : \mathbb{R}^n \to \mathbb{R}$ be $C^k$ functions, for $t \in T$ ($T$ countable compact) and assume:

$$(x,t) \longmapsto (g_t(x), \nabla g_t(x)) \quad \text{continuous.}$$

## Fact

*The proof uses the Morse–Sard theorem, but it also recovers it.*

### APPLICATION I: SEMI-INFINITE PROGRAMMING

Let $g_t : \mathbb{R}^n \to \mathbb{R}$ be $C^k$ functions, for $t \in T$ ($T$ countable compact) and assume:

$$(x,t) \longmapsto (g_t(x), \nabla g_t(x)) \quad \text{continuous.}$$

**Semi-infinite problem** (depending on $r \in \mathbb{R}$):

$$(\mathcal{P}_r) \qquad \min_{g_t(x) \leq r} \quad u(x)$$

where $u : \mathbb{R}^d \to \mathbb{R}$ is *regular* Lipschitz.

- Necessary optimality conditions for $(\mathcal{P}_r)$.

For *a.a.* $r \in \mathbb{R}$, and for every solution $\bar{x}$ of $(\mathcal{P}_r)$ there exist $\lambda_1, \ldots, \lambda_d \geq 0$ and $\{t_1, \ldots, t_d\} \subset T(\bar{x})$ such that

$$0 \in \partial u(\bar{x}) + \sum_{i=1}^{d} \lambda_i \nabla g_{t_i}(\bar{x})$$

- Necessary optimality conditions for $(\mathcal{P}_r)$.

For *a.a.* $r \in \mathbb{R}$, and for every solution $\bar{x}$ of $(\mathcal{P}_r)$ there exist $\lambda_1, \ldots, \lambda_d \geq 0$ and $\{t_1, \ldots, t_d\} \subset T(\bar{x})$ such that

$$0 \in \partial u(\bar{x}) + \sum_{i=1}^{d} \lambda_i \nabla g_{t_i}(\bar{x})$$

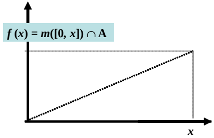
**Question**

*How often the Clarke subdifferential provides useful information ?*

## STRONG FAILURE OF SARD (LEBOURG'S EXAMPLE)

Let $A$ be a measurable subset of $\mathbb{R}$ that splits the family of intervals, that is, for every nontrivial interval $I \subset \mathbb{R}$ we have:

$$0 < m(A \cap I) < m(I)$$




$$f(x) = m(A \cap [0,x]) = \int_0^x \chi_A(t)dt$$

For all $x \in \mathbb{R}$ we have $\partial f(x) = [0,1]$ (every point is Clarke critical), but $f$ is strictly increasing (!)

Let $f : \mathcal{U} \subset \mathbb{R}^n \to \mathbb{R}$ be a Lipschitz continuous function. We denote:

$$||f||_{Lip} := \sup_{\substack{x,y \in \mathcal{U} \\ x \neq y}} \left\{ \frac{|f(x) - f(y)|}{||x - y||} \right\} \qquad \text{(Lipschitz constant)}$$

It follows easily $\partial f(x) \subset B_*(0, ||f||_{Lip})$, for all $x \in \mathcal{U}$.

**Definition (Clarke saturation)**

A Lipschitz function is called Clarke-saturated, if

$$\partial f(x) = B_*(0, ||f||_{Lip}), \qquad \text{for all } x \in \mathcal{U}$$

A slight modification of the Lebourg's example provide an example of a Clarke saturated function.

**Example (Clarke saturated function)**

Let $A$ be a measurable subset of $\mathbb{R}$ that splits the family of intervals. Then the function

$$f(x) = \int_0^x \left[\chi_A(t) - \chi_{\mathbb{R}\setminus A}(t)\right] dt$$

is Lipschitz ($||f||_{Lip} = 1$) and Clarke saturated.

- How often we come across Clarke-saturated functions ?

## GENERIC PATHOLOGY

**Theorem (X. Wang, PhD 1998)**

*A generic Lipschitz function is Clarke saturated.*

Series of papers of *J. Borwein, W. Moors, X. Wang* (1996–2006).

- Analogy with the genericity of continuous functions that are nowhere differentiable (Weierstrass example)

## GENERIC PATHOLOGY IN DETAILS

Fix a nonempty open set $\mathcal{U}$ and $k > 0$. Set

$$\mathrm{Lip}_k(\mathcal{U}) := \{f : \mathcal{U} \to \mathbb{R} : \quad ||f||_{Lip} \leq k\}$$

and $\rho$ the metric of uniform convergence on bounded sets.

- $(\mathrm{Lip}_k(\mathcal{U}), \rho)$ is a complete metric space

**Theorem**

*The set of Clarke saturated k-Lipschitz functions:*

$$\{f \in \mathrm{Lip}_k(\mathcal{U}) \mid \partial^\circ f \equiv \overline{B}_*(0,k)\}$$

*is generic in* $(\mathrm{Lip}_k(\mathcal{U}), \rho)$ .

## REVISITING ...

- The result is a consequence of the Baire category theorem (It is not constructive).

## REVISITING ...

- The result is a consequence of the Baire category theorem (It is not constructive).

- The result is in $\mathrm{Lip}_k(\mathcal{U})$ (not in $\mathrm{Lip}(\mathcal{U})$)
  – Lack of linear structure

## REVISITING ...

- The result is a consequence of the Baire category theorem (It is not constructive).
- The result is in $\mathrm{Lip}_k(\mathcal{U})$ (not in $\mathrm{Lip}(\mathcal{U})$)
  – Lack of linear structure
- Uniform convergence is not a natural topology for the class of Lipschitz functions

Fix $x_0 \in \mathcal{U}$ and set

$$\mathrm{Lip}_{x_0}(\mathcal{U}) = \{f \in \mathrm{Lip}_{x_0}(\mathcal{U}) : f(x_0) = 0\}.$$

Then:

- $\mathrm{Lip}_{x_0}(\mathcal{U})$ a Banach space under the norm $||\cdot||_{\mathrm{Lip}}$. (Pivot dual space for the free space of $\mathcal{U}$.)

**Fact**

*The set of Clarke saturated Lipschitz functions cannot be generic in* $(\mathrm{Lip}_{x_0}(\mathcal{U}), ||\cdot||_{Lip})$

Fix $x_0 \in \mathcal{U}$ and set

$$\mathrm{Lip}_{x_0}(\mathcal{U}) = \{f \in \mathrm{Lip}_{x_0}(\mathcal{U}) : f(x_0) = 0\}.$$

Then:

- $\mathrm{Lip}_{x_0}(\mathcal{U})$ a Banach space under the norm $||\cdot||_{\mathrm{Lip}}$. (Pivot dual space for the free space of $\mathcal{U}$.)

**Fact**

*The set of Clarke saturated Lipschitz functions cannot be generic in $(\mathrm{Lip}_{x_0}(\mathcal{U}), ||\cdot||_{Lip})$*

- How to measure the size of Clarke saturated functions in $(\mathrm{Lip}_{x_0}(\mathcal{U}), ||\cdot||_{Lip})$ ?

**Definition (V. I. Gurariy)**

A subset $M$ of functions on $\mathbb{R}$ is called **lineable** (resp. **spaceable**) if $M \cup \{0\}$ contains an infinite dimensional (resp. a closed infinite dimensional) subspace.

## Definition (V. I. Gurariy)

A subset $M$ of functions on $\mathbb{R}$ is called **lineable** (resp. **spaceable**) if $M \cup \{0\}$ contains an infinite dimensional (resp. a closed infinite dimensional) subspace.

## Fact

– *The set of continuous nowhere differentiable functions in* $\mathbb{R}$ *is lineable.*
– *The set of continuous nowhere differentiable functions in* $C[0,1]$ *is spaceable.*

Let $\mathcal{U} \neq \varnothing$ be an open convex subset of $(\mathbb{R}^n, ||\cdot||_1)$.

## Theorem (Daniilidis–Flores, 2019)

(i) *(lineability) The vector space* $\mathrm{Lip}(\mathcal{U})$ *contains a linear subspace of Clarke-saturated functions of uncountable dimension.*

(ii) *(spaceability) The Banach space* $(\mathrm{Lip}_{x_0}(\mathcal{U}), ||\cdot||_{\mathrm{Lip}})$ *contains a subspace of Clarke-saturated functions isometric to* $\ell^\infty(\mathbb{N})$.

## SCKETCHING THE PROOF $(n = 1)$

- Case $n = 1$

### Lemma (countable splitting partition)

*There exists a countable partition $\{A_k\}_{k\in\mathbb{N}}$ of $\mathbb{R}$ each of which splits the family of intervals.*

We set:

$$g_k(x) = \chi_{A_{2k+1}}(x) - \chi_{A_{2k}}(x), \quad \text{and} \quad f_k(x) = \int_{x_0}^{x} g_k(t)dt$$

## SCKETCHING THE PROOF $(n = 1)$

**Fact**

*Every linear combination of the functions $\{f_k\}_{k\in\mathbb{N}}$ is Clarke-saturated.*

## SCKETCHING THE PROOF $(n = 1)$

### Fact

*Every linear combination of the functions $\{f_k\}_{k\in\mathbb{N}}$ is Clarke-saturated.*

### Fact

*We can build an isometry between $c_0(\mathbb{N})$ and the vector space generated by $\{f_k\}_{k\in\mathbb{N}}$.*

## SCKETCHING THE PROOF $(n = 1)$

**Fact**

*Every linear combination of the functions $\{f_k\}_{k\in\mathbb{N}}$ is Clarke-saturated.*

**Fact**

*We can build an isometry between $c_0(\mathbb{N})$ and the vector space generated by $\{f_k\}_{k\in\mathbb{N}}$.*

**Fact**

*We can extend to an isometry between $\ell^\infty(\mathbb{N})$ and a subspace of Clarke-saturated functions.*

## Theorem (G. Flores, Master Thesis, 2016 –(*))

*The linear operator* $\mathrm{D} : \mathrm{Lip}_{x_0}(\mathcal{U}) \to L^\infty(\mathcal{U}; \ell_n^\infty)$ *defined by*

$$\mathrm{D}f = Df \quad a.e.$$

*defines an isometry between* $\mathrm{Lip}_{x_0}(\mathcal{U})$ *and*

$$\{g \in L^\infty(\mathcal{U}; \ell_d^\infty) \,:\, \partial_i g_j = \partial_j g_i \ \forall i,j\}.$$

*that is,*

$$\int_{\mathcal{U}} g_j(x) \frac{\partial \varphi}{\partial x_i}(x) dx = \int_{\mathcal{U}} g_i(x) \frac{\partial \varphi}{\partial x_j} dx, \quad \text{for every } \varphi \in C_0^\infty(\mathcal{U}).$$

(*) Proved independently in Cúth, Kalenda, Kaplický, Matematica, 2017)

## SCKETCHING THE PROOF (GENERAL CASE)

- General case : $(\mathbb{R}^n, ||\cdot||_1) = \ell_n^1$

We set

$$\begin{cases} G^k : \mathcal{U} \to (\mathbb{R}^n, ||\cdot||_\infty) = (\mathbb{R}^n, ||\cdot||_1)^* \\ G^k(x) := (\chi_{A_{2k+1}}(x_1) - \chi_{A_{2k}}(x_1), \ldots, \chi_{A_{2k+1}}(x_d) - \chi_{A_{2k}}(x_d)). \end{cases}$$

that is,

$$\langle G^k(x), e_i \rangle = g_k(\langle x, e_i \rangle),$$

**Fact**

*There exist* $\{f_k\}_{k\geq 0} \subseteq \mathrm{Lip}_{x_0}(\mathcal{U})$ *such that* $f_k = \mathrm{D}^{-1}(G^k)$, *for all* $k \geq 1$.

- The rest of the proof follows the 1-dimensional case.

- If we consider $(\mathbb{R}^n, ||\cdot||_2)$, then $\partial f(x)$ does not cover the whole dual ball $B_*(0, ||f||_{Lip})$; Still, it does contain a ball around 0.

- The property "every point is Clarke critical" is a spaceable property in $\mathrm{Lip}_{x_0}(\mathcal{U})$, where $\mathcal{U} \subset \mathbb{R}^n$.

## REFERENCES

- J. Borwein, X. Wang, Lipschitz functions with maximal subdifferentials are generic, *Proc. Amer. Math. Soc.* **128** (2000), 3221–3229.
- J. Bolte, A. Daniilidis, A. Lewis, M. Shiota, Clarke subgradients of stratifiable functions, *SIAM J. Optim.* **18** (2007), 556–572
- L. Barbet, A. Daniilidis, M. Dambrine, The Morse-Sard Theorem for Clarke critical values, *Advances in Math.* **242** (2013), 217–227.
- A. Daniilidis, G. Flores, Linear structure of functions with maximal Clarke subdifferential, *SIAM J. Optim.* **29** (2019), 511–521.