



Steepest Geometric Descent for Regularized Quasiconvex Functions

Aris Daniilidis¹ · David Salas²

Received: 1 October 2023 / Accepted: 7 August 2024
© The Author(s) 2024

Abstract

We establish existence of steepest descent curves emanating from almost every point of a regular locally Lipschitz quasiconvex functions, where regularity means that the sweeping process flow induced by the sublevel sets is reversible. We then use max-convolution to regularize general quasiconvex functions and obtain a result of the same nature in a more general setting.

Keywords Steepest descent curves · Quasiconvex functions · Max-convolution · Sweeping process

Mathematics Subject Classification Primary 26B25 · 37C10 · Secondary 34A60 · 49J52 · 49J53

1 Introduction

Steepest descent curves are at the core of the theory of variational analysis, differential equations and optimization. Given a C^1 -smooth function $f : \mathbb{R}^d \rightarrow \mathbb{R}$ we call steepest descent curve the solution of the gradient flow equation

$$\begin{cases} \dot{x}(t) = -\nabla f(x(t)), & \text{a.e. } t \in [0, T], \\ x(0) = x_0. \end{cases} \quad (1.1)$$

It is well-known that the above differential equation has a solution (as a direct application of the Picard-Lindelöf theorem). When the assumption of smoothness is missing, the existence

Communicated by Claudia Sagastizabal

✉ A. Daniilidis
aris.daniilidis@tuwien.ac.at

D. Salas
david.salas@uoh.cl

¹ Institut für Stochastik und Wirtschaftsmathematik, VADOR E105-04 TU Wien, Wiedner Hauptstraße 8, A-1040 Wien, Austria

² Instituto de Ciencias de la Ingeniería, Universidad de O'Higgins, Av. Libertador Bernardo O'Higgins 611, Rancagua, Chile

of steepest descent curves can still be established for convex functions. Indeed, if the function $f : \mathbb{R}^d \rightarrow \mathbb{R}$ is convex, its steepest descent curves are solutions of the subdifferential inclusion

$$\begin{cases} \dot{x}(t) \in -\partial f(x(t)), & \text{a.e. } t \in [0, T], \\ x(0) = x_0. \end{cases} \tag{1.2}$$

It is well-known that the above differential inclusion admits a (unique) solution (see, e.g., [3]). Similarly, existence of solutions for several gradient descent and proximal methods are often based on convexity (see, e.g., [20]).

In the setting of metric analysis, a steepest descent curve is a 1-Lipschitz curve verifying the metric equation

$$\begin{cases} (f \circ x)'(t) = -|\nabla f|(x(t)), & \text{a.e. } t \in [0, T], \\ x(0) = x_0, \end{cases} \tag{1.3}$$

where $|\nabla f|$ denotes the (metric) slope of f introduced in [12]. This last formulation coincides with (1.1) in the smooth case and with (1.2) in the convex case after performing the usual arc-length reparametrization. The metric gradient flow given by (1.3) has been studied in detail (we refer to [1] for a comprehensive exposition). Remarkable families of functions also admit steepest descent curves in the above cases: geodesically convex functions in metric spaces [1] and smooth functions on Riemannian manifolds (see, e.g., [25]). However, existence of steepest descent curves is in general hard to verify, even for Lipschitz functions in \mathbb{R}^d (see comments of [16, Sect. 9.3.5]). Due to this obstruction, the authors in [11, 17] consider the more general notion of trajectories of a convex foliation (terminology introduced in [10]) and establish existence of such orbits (see, e.g., [11, Theorem 2.6]). In case the foliation is given by the sublevel set of a quasiconvex function, the above orbits are called orbits of geometric descent. Their connection with steepest descent orbits has been explored in [13, 16]: these curves fail to be steepest descent curves in general, but instead correspond to what the authors in [13] called *curves of near-steepest descent*. One of the main difficulties is in the fact that the slope mapping $x \mapsto |\nabla f|(x)$ fails to be lower-semicontinuous, inducing a gap with respect to its closure $x \mapsto \overline{|\nabla f|}(x)$. Curves of near-steepest descent lie, in some sense, within this gap.

In this work, we are interested in steepest descent curves for the class of locally Lipschitz quasiconvex functions. This is another very important family in the context of optimization with amenable properties (see, e.g., [4]). Even though the desired existence result seems to fail for this class as well (it is not yet clear if this is the case or not), we have been able to provide a positive existence result for the class of *regular* quasiconvex functions, where regularity ensures that the sweeping process flow induced by the sublevel sets is reversible. Then, for the general case of locally Lipschitz quasiconvex functions, we consider a regularization scheme using the max-convolution operator (see, e.g., [22] and the references therein). Indeed, we define for any locally Lipschitz quasiconvex function $f : \mathbb{R}^d \rightarrow \mathbb{R}$ and $\varepsilon > 0$, a regularization function f_ε satisfying the following properties:

- (i). f_ε admits steepest descent curves for almost every initial data on its domain.
- (ii). Every critical point of f_ε is at distance at most ε of a critical point of f .

Our work borrows heavily from the geometric approach of [11, 13]. We look at curves of geometric descent. Then, under regularity assumptions we are able to reverting the (unilateral) sweeping process and deduce that almost every curve of geometric descent is in fact a steepest descent curve.

The rest of the paper is organized as follows: In Sect. 2 we fix our terminology and quote some preliminary results. In Sect. 3, under adequate assumptions of the quasiconvex function (ensuring the reversibility of its sweeping process flow), we directly relate steepest descent curves as solutions of the sweeping process induced by the sublevel sets. In Sect. 4, we extend the results of the preceding section to general quasiconvex function, by means of regularization and localization.

2 Preliminaries

Throughout this work, we consider the Euclidean space \mathbb{R}^d endowed with its usual inner product $\langle \cdot, \cdot \rangle$ and its induced norm $\| \cdot \|$. We denote by $B(x, r)$ (respectively, $\bar{B}(x, r)$) the open (respectively, closed) ball centered at x of radius $r > 0$ and by \mathbb{B}_d (respectively, \mathbb{S}_d) the unit closed ball (respectively, the unit sphere). For a set $A \subset \mathbb{R}^d$, we denote by $\text{int}(A)$, \bar{A} , $\text{bd}(A)$ and A° its interior, closure, boundary and (negative) polar set, respectively.

For a function $f : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$ and $\alpha \in \mathbb{R}$, we denote by $[f \leq \alpha]$, the α -sublevel set of f , that is,

$$[f \leq \alpha] = \{x \in \mathbb{R}^d \mid f(x) \leq \alpha\}. \tag{2.1}$$

Similarly, we define the strict α -sublevel set $[f < \alpha]$, and the corresponding sets $[f = \alpha]$, $[f > \alpha]$ and $[f \geq \alpha]$. We denote its (effective) domain by $\text{dom } f$, that is, $\text{dom } f = \{x \in \mathbb{R}^d \mid f(x) < +\infty\}$. A function $f : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$ is called coercive if for every $\alpha < \sup f$ the sublevel set $[f \leq \alpha]$ is compact.

A function $f : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$ is said to be *quasiconvex* if

$$\forall x, y \in \mathbb{R}^d, \forall t \in [0, 1], \quad f(tx + (1 - t)y) \leq \max\{f(x), f(y)\}. \tag{2.2}$$

It is well-known that f is quasiconvex if and only if every sublevel set $[f \leq \alpha]$ is convex. Recall that a function is lower semicontinuous (lsc, for short) if the sublevel sets are closed.

For a function $f : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$ we define the metric slope $|\nabla f|$ as

$$|\nabla f|(x) := \begin{cases} \limsup_{y \rightarrow x} \frac{(f(x) - f(y))^+}{\|x - y\|}, & \text{if } x \in \text{dom } f \\ +\infty, & \text{otherwise,} \end{cases} \tag{2.3}$$

where $a^+ = \max\{a, 0\}$. The metric slope enjoys several interesting properties (see, e.g., [1, 5]), but it is well-known that it might fail to be lower semicontinuous (see, e.g., [13]). Thus, we consider the limiting slope $\overline{|\nabla f|}$ as the lower semicontinuous closure of $|\nabla f|$, that is,

$$\overline{|\nabla f|}(x) = \liminf_{y \rightarrow_f x} |\nabla f|(y), \tag{2.4}$$

where $y \rightarrow_f x$ means $(y, f(y)) \rightarrow (x, f(x))$. A point $x \in \mathbb{R}^d$ is called *critical* for f if $\overline{|\nabla f|}(x) = 0$.

In what follows, we denote by either $d_S(x)$ or $d(x, S)$ the distance of $x \in \mathbb{R}^d$ to the set $S \subset \mathbb{R}^d$ and by $\text{Proj}_S(x)$ or $\text{Proj}(x; S)$ the set of nearest points from x in S . Whenever this set is a singleton, we call this unique nearest point as *metric projection* and we denote it by $\text{proj}_S(x)$.

A set S is said to be *prox-regular* if there exists a continuous function $\rho : S \rightarrow (0, +\infty]$ such that the enlargement of S given by

$$U_{\rho(\cdot)}(S) = \{u \in \mathbb{R}^d : \exists y \in \text{Proj}_S(u) \text{ with } d_S(u) < \rho(y)\} \tag{2.5}$$

is open and the projection proj_S is well-defined on $U_{\rho(\cdot)}(S)$ (equivalently, d_S^2 is C^1 on $U_{\rho(\cdot)}(S)$, see, e.g. [6, Prop. 4 and Prop. 11]). For $r > 0$, we say that S is *r-prox-regular* if the function $\rho(\cdot)$ can be taken as $\rho \equiv r$. Every convex set is $(+\infty)$ -prox-regular.

It is well-known (see, e.g., [24]) that for a prox-regular set, Bouligand and Clarke tangent cones coincide at every point (this is known as tangential regularity) and the same applies to the classical notions of normal cones (proximal, Fréchet, limiting, Clarke). Since we are going to work only with convex and prox-regular sets, the notions of tangent and normal cones are unambiguously defined, that is, if $S \subset \mathbb{R}^d$ is prox-regular and $x \in S$, we define the (Clarke) *tangent cone* and the (Clarke) *normal cone* of S at x by the formulae

$$T(S; x) := \text{Liminf}_{S \ni y \rightarrow x; t \downarrow 0} \frac{1}{t}(S - y) \quad \text{and} \quad N(S; x) := [T(S; x)]^\circ, \tag{2.6}$$

where Liminf is the inferior limit of sets in the sense of Painlevé-Kuratowski (see, e.g., [24]).

A set-valued map $M : A \rightrightarrows B$ is a mapping that assigns to each $a \in A$ a subset $M(a)$ of B . We denote the domain of M and the graph of M as the sets $\text{dom } M = \{a \in A : M(a) \neq \emptyset\}$ and $\text{gph } M = \{(a, b) : b \in M(a)\}$. In the particular case when $A = [0, T]$ and $B = \mathbb{R}^d$, we say that the set-valued map is a *moving set map* (also called *sweeping process map*, see [8]).

For two sets $A, B \subset \mathbb{R}^d$, the Hausdorff distance between A and B is given by

$$d_H(A, B) = \max \left\{ \sup_{a \in A} d_B(a), \sup_{b \in B} d_A(b) \right\} \in [0, +\infty]. \tag{2.7}$$

A set-valued map $M : A \subset \mathbb{R}^p \rightrightarrows \mathbb{R}^q$ is said to be Lipschitz-continuous if there exists $L > 0$ such that

$$\forall x, y \in A, \quad d_H(M(x), M(y)) \leq L \|x - y\|. \tag{2.8}$$

Let $K : [0, T] \rightrightarrows \mathbb{R}^d$ be a moving set with prox-regular values. We define the sweeping process differential inclusion of K as

$$\begin{cases} \dot{u}(t) \in -N(K(t); u(t)), & \text{a.e. } t \in [0, T] \\ u(0) = x_0 \in K(0), \end{cases} \tag{2.9}$$

It is well known that if K is Lipschitz continuous and uniformly r -prox-regular for some $r > 0$ (that is, $K(t)$ is r -prox-regular for every $t \in [0, T]$), the sweeping process admits a unique solution for every initial condition $x_0 \in K(0)$ (see, e.g., [23]). The following proposition surveys the main properties of such a solution in the case of the sublevel moving set $K : [0, T] \rightrightarrows \mathbb{R}^d$ given by $K(t) = [f \leq f(x_0) - t]$. The first statement is a classical fact in the sweeping process theory (see, e.g., [23]). The second and third statements follow directly from the fact that the process $K : [0, T] \rightrightarrows \mathbb{R}^d$ is parametrized with respect to the values of the function f , while the fourth statement follows from [13, Theorem 3.4 and Claim 3.6].

Proposition 2.1 *Let $f : \mathbb{R}^d \rightarrow \mathbb{R}$ be a locally Lipschitz function and let $\alpha \in \mathbb{R}$ and $T > 0$ be such that $\alpha - T > \inf f$. Let $K : [0, T] \rightrightarrows \mathbb{R}^d$ be the sublevel moving set starting from α , that is,*

$$K(t) = [f \leq \alpha - t], \quad \forall t \in [0, T].$$

If K is Lipschitz-continuous and uniformly r -prox-regular, then for every $x_0 \in K(0) \setminus \text{int}(K(T))$, the sweeping process (2.9) has a unique solution $u : [0, T] \rightarrow \mathbb{R}^d$, satisfying that

- (i) $u(\cdot)$ is Lipschitz-continuous on $[0, T]$.
- (ii) For each $t \in [0, \alpha - f(x_0)]$, $u(t) = x_0$ and

$$u(t) \in \text{bd } K(t), \quad \forall t \in [\alpha - f(x_0), T].$$

- (iii) $(f \circ u)'(t) = -1$ for a.e. $t \in [\alpha - f(x_0), T]$, and
- (iv) $u(\cdot)$ is a curve of near-maximal slope of f , in the sense that

$$\frac{1}{|\nabla f|(u(t))} \leq \|\dot{u}(t)\| \leq \frac{1}{|\nabla f|(u(t))}, \quad \text{for a.e. } t \in [\alpha - f(x_0), T].$$

Last but not least, following [14], we denote by $DF(x)$ the derivative of a Lipschitz function $F : \mathbb{R}^d \rightarrow \mathbb{R}^d$ at each point of differentiability $x \in \mathbb{R}^d$ and by

$$JF(x) = |\det(DF(x))|, \tag{2.10}$$

the Jacobian determinant of F . We finally denote by \mathcal{H}^m the m -dimensional Hausdorff measure.

3 Reversible Geometric Descent for Regular Quasiconvex Functions

This section is devoted to the study of geometrical curves of descent for a lower semicontinuous quasiconvex function $f : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$ which is *continuous on the interior of its domain*. We further consider functions that satisfy the following regularity hypotheses:

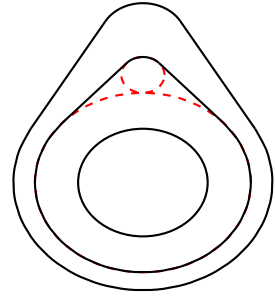
- (H1) f is coercive and $\text{int}[f \leq \alpha] \neq \emptyset$, for every $\alpha > \inf f$.
- (H2) The slope of f is bounded away from zero around every $x \in \text{dom } f \setminus \text{argmin } f$, that is, there exist $\delta, \ell > 0$ such that

$$|\nabla f|(y) > \ell, \quad \forall y \in B(x, \delta) \cap \text{dom } f.$$

- (H3) For every $\alpha \in (\inf f, \sup f)$, there exist $\eta, r > 0$ such that for every $\beta \in (\alpha - \eta, \alpha + \eta)$, the set $\mathbb{R}^d \setminus \text{int}([f \leq \beta])$ is r -prox-regular.

Remark 3.1 While (H1) is a common assumption, (H2)–(H3) are rather strong regularity conditions. Hypothesis (H2) is equivalent to saying that the only critical points of f are its minimizers. (Such functions are often called *pseudoconvex*, see e.g. [9]). Hypothesis (H3) entails the smoothness of the level sets (see Lemma 3.3 below). We shall show in Sect. 4 that every locally Lipschitz quasiconvex function can be regularized in a way that (H1)–(H3) are fulfilled.

Fig. 1 Boundary of $S(s)$ for $s = 0.75, s = 1.25$ and $s = 1.5$



Notice that hypothesis (H2) together with continuity of f on the interior of its domain yield that for every $x \in \text{int}(\text{dom } f)$ and $\delta > 0$ such that $B(x, \delta) \subset \text{dom } f$, setting $f(x) = r$ it holds:

$$\text{bd}[f \leq r] \cap B(x, \delta) = [f = r] \cap B(x, \delta). \tag{3.1}$$

Indeed, continuity of f entails directly the left-to-right inclusion. If the reverse inclusion does not hold, then there would exist $z \in \mathbb{R}^d$ and $\varepsilon > 0$ small enough such that

$$B(z, \varepsilon) \subset [f = f(x)] \cap B(x, \delta).$$

This would yield $|\nabla f|(z) = 0$ contradicting (H2).

Let us also mention that it is easy to construct a quasiconvex function $f : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$ whose sublevel sets have smooth boundaries, yet failing the reversibility hypothesis (H3).

Example 3.2 Let $S : [0, 2] \rightrightarrows \mathbb{R}^2$ be a convex valued function defined by

$$S(s) := \begin{cases} B(0, s), & \text{if } s \in [0, 1). \\ \text{co}\left(B(0, s) \cup B((0, 2s - 1), s - 1)\right), & \text{if } s \in [1, 2], \end{cases}$$

We set $f(x) = \inf\{s : x \in S(s)\}$. Then f is lower semicontinuous, quasiconvex and locally Lipschitz on the interior of its domain. Moreover, for every $s \in [0, 2]$, the set $[f = s] \equiv \text{bd } S(s)$ is a smooth manifold, whose minimal value of the internal curvature is $s - 1$ for $s \in [1, 2]$, and s for $s \in [0, 1]$ (see Fig. 1). ◇

Let $\alpha_1, \alpha_2 \in \mathbb{R}$ such that $\inf f < \alpha_1 < \alpha_2 < \sup f$. The goal of this section is to show that under (H1)–(H3), the function f admits steepest descent curves, which locally induce a foliation of the annulus $[\alpha_1 \leq f \leq \alpha_2]$.

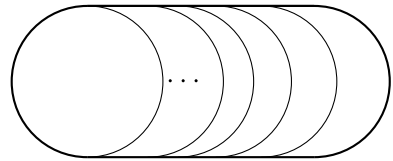
The first step is the following proposition that shows that prox-regularity of the boundary entails in fact smoothness of it.

Lemma 3.3 (Smoothness of the boundaries) *Under (H1)–(H3), for every $\alpha \in (\inf f, \sup f)$, the set $\mathcal{M} = \text{bd}([f \leq \alpha]) \cap \text{int}(\text{dom } f)$ is a $\mathcal{C}^{1,1}$ -submanifold.*

Proof Let $S = [f \leq \alpha]$ and $U = \mathbb{R}^d \setminus \text{int}(S)$. Since U is r -prox-regular, using [21, Theorem 6.42] (and noting that prox-regularity entails regularity in the sense of [21, Definition 6.4]), we get that

$$T_{\text{bd } S}(x) = T_S(x) \cap T_U(x) \quad \forall x \in \text{bd } S.$$

Fig. 2 Each sublevel set has a smooth boundary which does not coincide with the corresponding level set



Moreover, since S is convex with nonempty interior, we can apply [7, Proposition 2.3] to deduce that

$$T_S(x) = -T_U(x), \quad \forall x \in \text{bd } S.$$

Combining the above equations we deduce that $T_{\text{bd } S}(x)$ is a vector space. Using [24, Proposition 2.113 (a5)], we get $\mathbb{R}^d \setminus T_S(x) \subset T_U(x)$ and consequently $\text{bd } T_S(x) \subset T_{\text{bd } S}(x)$. We conclude that $T_{\text{bd } S}(x)$ is of codimension 1. Therefore, for every $x \in \text{bd } S$, there exists a unique unit vector $\hat{n}(x) \in \mathbb{S}_d$ such that

$$N(S, x) = -N(U, x) = \mathbb{R}_+ \hat{n}(x).$$

Thus, $N(\text{bd } S, x) = \mathbb{R} \hat{n}(x)$. Since S is convex, the mapping $\hat{n} : \text{bd } S \rightarrow \mathbb{S}_d$ is continuous. Furthermore, since U is r -prox-regular for some $r > 0$, then the set $U_{-\varepsilon} = \{z : d_S(z) \geq \varepsilon\}$ must be $(r + \varepsilon)$ -prox-regular. By noting that

$$\hat{n}(x) = \frac{1}{\varepsilon}(\text{proj}(x, U_{-\varepsilon}) - x), \quad \forall x \in \text{int}(\text{dom } f) \cap \text{bd } S,$$

we deduce that $\hat{n} : \text{int}(\text{dom } f) \cap \text{bd } S \rightarrow \mathbb{S}_d$ is also Lipschitz-continuous and the proof is complete. \square

Remark 3.4 An alternative proof of the above lemma can be derived using the enhanced Baillon-Haddad theorem of [19], by representing the convex set $[f \leq \alpha]$ as the epigraph of a convex function over an appropriate subspace of codimension 1. The above presentation aims at further describing the behavior of the tangent and normal cones of $S = [f \leq \alpha]$ and $U = \mathbb{R}^d \setminus \text{int}(S)$.

Example 3.5 Smoothness of $\text{bd}[f \leq \alpha]$ does not entail that this set coincides with the corresponding level set $[f = \alpha]$. Discrepancies may appear due to the cutting effect of the boundary of the domain as illustrated in the following example:

Set $D = \text{co}(\mathbb{B}_2 \cup ((3, 0) + \mathbb{B}_2)) \subset \mathbb{R}^2$ and define $f : \mathbb{R}^2 \rightarrow \{+\infty\}$ given by

$$f(x, y) = \begin{cases} \min\{t : (x, y) \in (t, 0) + \mathbb{B}_2\}, & \text{if } (x, y) \in D \\ +\infty, & \text{otherwise.} \end{cases} \tag{3.2}$$

The above function is quasiconvex and its sublevel sets are given by $[f \leq t] = \text{co}(\mathbb{B}_2 \cup ((t, 0) + \mathbb{B}_2))$, for $t \in [0, 3]$, $[f \leq t] = \text{dom } f = D$ if $t \geq 3$, and $[f \leq t] = \emptyset$ if $t < 0$ (see Fig. 2).

It is easy to see that f satisfies (H1), (H2) and (H3). The second one follows from the remark that for every angle $\theta \in [-\pi/2, \pi/2]$ one has that

$$f(\cos(\theta) + t, \sin(\theta)) = t, \quad \forall t \in [0, 3],$$

and consequently $|\nabla f|(x, y) \geq 1$ for all $(x, y) \in \text{dom } f \setminus \text{argmin } f = D \setminus \mathbb{B}_2$. The first and third hypotheses follows from the construction. \diamond

In what follows, let $\alpha_1, \alpha_2 \in \mathbb{R}$ be such that $\alpha_1 < \alpha_2$ and set

$$T := \alpha_2 - \alpha_1 > 0.$$

We consider the *annulus* set

$$\mathcal{R}(\alpha_1, \alpha_2) = [\alpha_1 \leq f \leq \alpha_2]$$

as well as the decreasing moving set maps $\widehat{\mathcal{S}}, \widehat{\mathcal{U}}$ given by

$$\left\{ \begin{array}{l} \widehat{\mathcal{S}} : [0, T] \rightrightarrows \mathbb{R}^d \\ \widehat{\mathcal{S}}(t) := [f \leq \alpha_2 - t] \end{array} \right. \quad \text{and} \quad \left\{ \begin{array}{l} \widehat{\mathcal{U}} : [-T, 0] \rightrightarrows \mathbb{R}^d \\ \widehat{\mathcal{U}}(\tau) := \underbrace{\mathbb{R}^d \setminus \text{int}([f \leq \alpha_2 + \tau])}_{[\text{int}(\widehat{\mathcal{S}}(-\tau))]^c} \end{array} \right. \quad (3.3)$$

The following lemma establishes the variational regularity of the sublevel sets of f , under hypotheses (H1)-(H3). Namely, we prove the Lipschitz continuity of the set-valued maps given in (3.3). Similar results involving metric regularity can be obtained as well using the metric slope (see, e.g., [5, 16]).

Lemma 3.6 *The moving set maps $\widehat{\mathcal{S}}$ and $\widehat{\mathcal{U}}$ are Lipschitz continuous provided (H1)-(H2) hold. Moreover, $\widehat{\mathcal{U}}$ is uniformly r -prox-regular for some $r > 0$, provided (H3) holds.*

Proof Thanks to (H2) for each $x \in \mathcal{R}(\alpha_1, \alpha_2)$, there exist $\delta_x > 0$ and $\ell_x > 0$ such that

$$|\nabla f|(z) \geq \ell_x, \quad \text{for all } z \in B(x, \delta_x).$$

Since $\mathcal{R}(\alpha_1, \alpha_2)$ is compact due to the coercivity of f , we deduce that there is $\ell > 0$ such that $|\nabla f|(x) \geq \ell$ for all $x \in \mathcal{R}(\alpha_1, \alpha_2)$. By [5, Theorem 2.1], for all $x \in \mathcal{R}(\alpha_1, \alpha_2)$ and $\alpha \in [\alpha_1, \alpha_2]$ we have:

$$d(x, [f \leq \alpha]) \leq \frac{1}{\ell}(f(x) - \alpha)^+.$$

Now, choose $s, t \in [0, T]$ and suppose that $s < t$. Then,

$$\begin{aligned} d_H(\widehat{\mathcal{S}}(t), \widehat{\mathcal{S}}(s)) &= \sup_{x \in \widehat{\mathcal{S}}(s)} d(x, \widehat{\mathcal{S}}(t)) = \sup_{x \in \widehat{\mathcal{S}}(s)} d(x, [f \leq \alpha_2 - t]) \\ &\leq \sup_{x \in \widehat{\mathcal{S}}(s)} \frac{1}{\ell}(f(x) - \alpha_2 + t)^+ = \frac{1}{\ell}|t - s|. \end{aligned}$$

Thus, $\widehat{\mathcal{S}}$ is a $(1/\ell)$ -Lipschitz set-valued map. Now take

$$x \in \widehat{\mathcal{U}}(-t) \setminus \widehat{\mathcal{U}}(-s) = [\text{int}(\widehat{\mathcal{S}}(t))]^c \setminus [\text{int}(\widehat{\mathcal{S}}(s))]^c = \widehat{\mathcal{U}}(-t) \cap \text{int}(\widehat{\mathcal{S}}(s))$$

(notice that by (H1) the above set is nonempty) and let \hat{n} the exterior unit vector of $\widehat{\mathcal{S}}(t)$ at $\text{proj}_{\widehat{\mathcal{S}}(t)}(x)$. Then, since f is coercive, there exists $\zeta > 0$ such that

$$\text{proj}_{\widehat{\mathcal{S}}(t)}(x) + \zeta \hat{n} \in \text{bd} \widehat{\mathcal{S}}(s) = \text{bd} \widehat{\mathcal{U}}(-s).$$

Clearly

$$d(x, \widehat{U}(-s)) \leq \zeta = d(\text{proj}_{\widehat{S}(t)}(x) + \zeta \widehat{n}, \widehat{S}(t)) \leq \sup_{y \in \widehat{S}(s)} d(y, \widehat{S}(t)).$$

With this in mind, we can write

$$d_H(\widehat{U}(-t), \widehat{U}(-s)) = \sup_{x \in \widehat{U}(-t) \cap \widehat{S}(s)} d(x, \widehat{U}(-s)) \leq \sup_{y \in \widehat{S}(s)} d(y, \widehat{S}(t)) \leq \frac{1}{\ell} |t - s|.$$

Thus, \widehat{U} is also $\frac{1}{\ell}$ -Lipschitz. The last assertion of the statement is straightforward. □

In what follows, we consider the sweeping process (2.9) for the moving set maps $\widehat{S} : [0, T] \rightrightarrows \mathbb{R}^d$ and $\widehat{U} : [-T, 0] \rightrightarrows \mathbb{R}^d$, that is,

$$\begin{cases} \dot{u}(t) \in -N(\widehat{S}(t); u(t)), & t \in [0, T] \\ u(0) = x_0 \in \widehat{S}(0) \end{cases} \quad \text{and} \quad \begin{cases} \dot{v}(\tau) \in -N(\widehat{U}(\tau); v(\tau)), & \tau \in [-T, 0] \\ v(-T) = y_0 \in \widehat{U}(-T) \end{cases} \quad (3.4)$$

We now set

$$M = \text{bd}([f \leq \alpha_2]) \quad \text{and} \quad \mathcal{R} = [f \leq \alpha_2] \setminus \text{int}([f \leq \alpha_1]), \quad (3.5)$$

and consider the mapping

$$\begin{aligned} u : [0, T] \times M &\rightarrow \mathcal{R} \\ (t, m) &\mapsto u(t, m), \end{aligned} \quad (3.6)$$

where $u(\cdot, m)$ is the unique solution of the first sweeping process differential inclusion of (3.4) with initial condition $x_0 = m$. In what follows, we endow the set $[0, T] \times M$ with the distance:

$$D((t_1, m_1), (t_2, m_2)) := |t_1 - t_2| + \|m_1 - m_2\|, \quad \forall t_1, t_2 \in [0, T], \forall m_1, m_2 \in M. \quad (3.7)$$

Proposition 3.7 (Inversion of the sweeping flow) *Let $m \in M \cap \text{int}(\text{dom } f)$. Suppose that f is Lipschitz-continuous around $m \in M$. Then, under (H1)–(H3), the mapping $u : [0, T] \times M \rightarrow \mathcal{R}$ is one-to-one and bi-Lipschitz around $(0, m)$.*

Proof Let us first show that u is (globally) Lipschitz. Let us denote by K the (common) Lipschitz constant of the moving set maps $\widehat{S}(\cdot)$ and $\widehat{U}(\cdot)$, given by Lemma 3.6. Notice that every trajectory of a K -Lipschitz sweeping process map is a Lipschitz curve with the same constant K (see e.g. [23]), that is, for every $m \in M$ and $t_1, t_2 \in [0, T]$ it holds

$$\|u(t_1, m) - u(t_2, m)\| \leq K |t_1 - t_2|.$$

Recall also (see [18]) that since \widehat{S} has convex values, the distance between two different trajectories is decreasing in time. We deduce directly that for $(t_1, m_1), (t_2, m_2) \in [0, T] \times M$ it holds:

$$\begin{aligned} \|u(t_1, m_1) - u(t_2, m_2)\| &\leq \|u(t_1, m_1) - u(t_2, m_1)\| + \|u(t_2, m_1) - u(t_2, m_2)\| \\ &\leq (\max\{1, K\}) D((t_1, m_1), (t_2, m_2)) \end{aligned}$$

Take $\delta > 0$ so that $B(m, \delta) \subset \text{int}(\text{dom } f)$. Recalling the definition of the set M from (3.5), since f is continuous on $B(m, \delta)$, we deduce that

$$\Gamma := M \cap B(m, \delta) = [f = \alpha_2] \cap B(m, \delta).$$

Moreover, by continuity of the mapping u (given in (3.6)), we get that there exists $\varepsilon > 0$ such that

$$u([0, \varepsilon] \times \Gamma) \subset \text{int}(\text{dom } f).$$

Shrinking δ and ε if necessary, we can assume that f is Lipschitz-continuous on $u([0, \varepsilon] \times \Gamma)$. Evoking Proposition 2.1 we deduce that for every $(t, m) \in (0, \varepsilon] \times \Gamma$, we have $f(u(t, m)) = \alpha_2 - t$, that is, the mapping u takes values in \mathcal{R} , given in (3.5). Moreover,

$$|t_1 - t_2| = |f(u_1) - f(u_2)| \leq L \|u_1 - u_2\|, \tag{3.8}$$

where L is the Lipschitz constant of f on the compact set $u([0, \varepsilon] \times \Gamma) \subset \text{int}(\text{dom } f)$.

Let us now show that $u(\cdot, \cdot)$ is one-to-one on $[0, \varepsilon] \times \Gamma$.

To this end, let $(t_1, m_1), (t_2, m_2) \in [0, \varepsilon] \times \Gamma$ be such that $u(t_1, m_1) = u(t_2, m_2)$. Since

$$f(m_1) = f(m_2) = \alpha_2,$$

we get that

$$\alpha_2 - t_1 = f(u(t_1, m_1)) = f(u(t_2, m_2)) = \alpha_2 - t_2,$$

which yields that $t_1 = t_2$. Let us denote by \bar{t} the common value of $t_1 = t_2$, and by \bar{u} the common value of $u(t_1, m_1) = u(t_2, m_2)$. Consider the differential inclusion

$$\begin{cases} \dot{v}(t) \in -N(\widehat{U}(t); v(t)), & t \in [-\bar{t}, 0], \\ v(-\bar{t}) = \bar{u}. \end{cases} \tag{3.9}$$

Hypotheses (H1)–(H3) ensure that $\widehat{U} : [-\bar{t}, 0] \rightrightarrows \mathbb{R}^d$ is uniformly prox-regular and Lipschitz continuous, entailing that the above differential inclusion has a unique solution $v : [-\bar{t}, 0] \rightarrow \mathbb{R}^d$ (c.f. Proposition 2.1). Noting that both $t \mapsto u(-t, m_1)$ and $t \mapsto u(-t, m_2)$ are solutions of the above differential inclusion, we deduce that

$$m_1 = u(0, m_1) = v(0) = u(0, m_2) = m_2,$$

and consequently $(t_1, m_1) = (t_2, m_2)$, proving that $u(\cdot, \cdot)$ is one-to-one.

Let us now show that the flow can be reversed, and that the expansion of the reversed flow can be controlled: to this end, let $(t_1, m_1), (t_2, m_2) \in [0, \varepsilon] \times \Gamma$, with $t_1 \leq t_2$, and set $\bar{u}_i = u(t_i, m_i)$, for $i \in \{1, 2\}$. Consider the differential inclusions, for $i \in \{1, 2\}$:

$$\begin{cases} \dot{v}_i(t) \in -N(v_i(t), \widehat{U}(t)), & t \in [-t_2, 0], \\ v_i(-t_2) = \bar{u}_i. \end{cases}$$

Thanks to Lemma 3.6 (see, e.g., [23]) and Proposition 2.1, the above differential inclusions have unique solutions, $v_1, v_2 : [-t_2, 0] \rightarrow \mathbb{R}^d$. It is not hard to see that $v_2(t) = u(-t, m_2)$ for

every $t \in [-t_2, 0]$ and that

$$v_1(t) = \begin{cases} \bar{u}_1, & \text{if } t \in [-t_2, -t_1] \\ u(-t, m_1), & \text{if } t \in [-t_1, 0]. \end{cases}$$

Let $\mathcal{P}_k := \{s_0, \dots, s_k\}$ be a uniform partition of the interval $[-t_2, 0]$ with width

$$|\mathcal{P}_k| := |s_j - s_{j-1}| = \frac{t_2}{k}, \quad \text{for all } j \in \{1, \dots, k\}.$$

Let $r > 0$ be such that the sets $\widehat{\mathcal{U}}(t)$ are r -uniformly prox-regular for all $t \in [-t_2, 0]$ and take k sufficiently large such that

$$\theta := \frac{K |\mathcal{P}_k|}{r} = \frac{K t_2}{r k} < 1. \tag{3.10}$$

Then for $i \in \{1, 2\}$, we define the polygonal curve $v_{i,k}$ emanating from u_i associated to \mathcal{P}_k as follows:

$$v_{i,k}(s_j) = \begin{cases} \bar{u}_i, & \text{if } j = 0 \\ \text{proj}(v_{i,k}(s_{j-1}); \widehat{\mathcal{U}}(s_j)), & \text{if } j \in \{1, \dots, k\}. \end{cases}$$

Notice that all projections are well-defined. In particular, for every $j \in \{1, \dots, k\}$

$$v_{1,k}(s_{j-1}), v_{2,k}(s_{j-1}) \in \widehat{\mathcal{U}}(s_j) + \underbrace{(K |s_j - s_{j-1}|)}_{\theta r} \mathbb{B}_d = \widehat{\mathcal{U}}(s_j) + \theta r \mathbb{B}_d.$$

Since the projection $\text{proj}(\cdot, \widehat{\mathcal{U}}(s_j))$ is Lipschitz with constant $(1 - \theta)^{-1}$ on the set

$$\widehat{\mathcal{U}}(s_j) + K |s_j - s_{j-1}| \mathbb{B}_d \equiv \widehat{\mathcal{U}}(s_j) + \theta r \mathbb{B}_d$$

(see, e.g., [6]) we obtain

$$\|v_{1,k}(s_j) - v_{2,k}(s_j)\| \leq (1 - \theta)^{-1} \|v_{1,k}(s_{j-1}) - v_{2,k}(s_{j-1})\|, \quad j \in \{1, \dots, k\}$$

and we deduce $\|v_{1,k}(0) - v_{2,k}(0)\| \leq (1 - \theta)^{-k} \|u_1 - u_2\|$.

Setting $\gamma := r^{-1} K t_2$ and recalling (3.10), we obtain

$$\|v_{1,k}(0) - v_{2,k}(0)\| \leq \left(1 - \frac{\gamma}{k}\right)^{-k} \|u_1 - u_2\|.$$

Since $v_{i,k}$ converges uniformly to v_i as $k \rightarrow \infty$ (see, e.g., [6]), we conclude that

$$\|m_1 - m_2\| = \|v_1(0) - v_2(0)\| \leq e^{r^{-1} K t_2} \|u_1 - u_2\|.$$

Recalling (3.8) we conclude that

$$D((t_1, m_1), (t_2, m_2)) \leq \left(L + e^{r^{-1} K T}\right) \|u(t_1, m_1) - u(t_2, m_2)\|.$$

This shows that u is bi-Lipschitz and the proof is complete. □

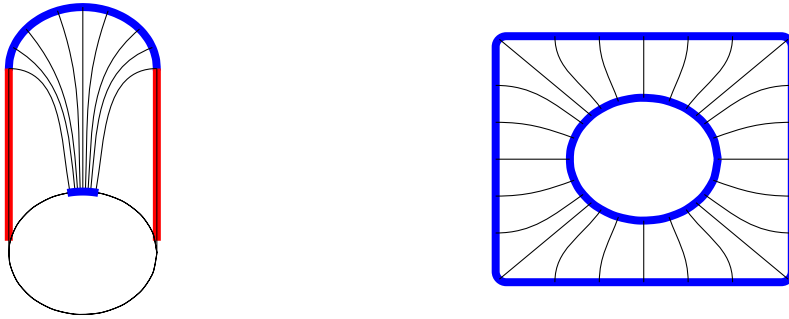


Fig. 3 Illustration of foliation induced by $u : [0, T] \times M \rightarrow \mathcal{R}$. *Left:* In blue, the part of M , corresponding to $M \cap \text{int}(\text{dom } f)$, which is injectively transported; In red, the part $M \cap \text{bd}(\text{dom } f)$, where u fails injectivity. *Right:* The case where $\text{dom } f = \mathbb{R}^d$ and u induces a complete foliation. (Color figure online)

The above proposition shows that the mapping $u : [0, T] \times M \rightarrow \mathcal{R}$ locally induces a foliation of the annulus $\mathcal{R} = [\alpha_1 \leq f \leq \alpha_2]$, near every point $m \in M \cap \text{int}(\text{dom } f)$ satisfying that f is Lipschitz-continuous on a neighborhood of m . The problem appears at points in $\text{bd}(\text{dom } f)$ that belong to the boundary of several sublevel sets, since at these points the mapping u loses injectivity. However, if $\text{dom } f = \mathbb{R}^d$ and f is locally Lipschitz, the neighborhood $[0, \varepsilon] \times \Gamma$ can be taken to be $[0, T] \times M$ and $u : [0, T] \times M \rightarrow \mathcal{R}$ induces a (complete) foliation of the whole annulus \mathcal{R} . This is illustrated in Fig. 3.

Recall that for every $m \in M \cap \text{int}(\text{dom } f)$ the set $\Gamma = M \cap B(m, \delta)$ is a $\mathcal{C}^{1,1}$ -submanifold of codimension 1 (see Lemma 3.3) and can naturally be endowed with its Hausdorff measure \mathcal{H}^{d-1} . Consequently, we can consider the measure $\mu := \mathcal{L}_1 \times \mathcal{H}^{d-1}$ over $[0, T] \times \Gamma$, where \mathcal{L}_1 is the Lebesgue measure over $[0, T]$. We further endow the set $\mathcal{R} = [f \leq 0] \setminus \text{int}[f \leq T]$ with the usual Lebesgue measure \mathcal{L}_d of \mathbb{R}^d .

Theorem 3.8 (control of null sets) *Let $\mathcal{N} \subset \mathbb{R}^d$ be a null measure set, and assume (H1)–(H3) hold. Let $\bar{m} \in M \cap \text{int}(\text{dom } f)$ such that f is Lipschitz-continuous near \bar{m} . Then, there exist $\varepsilon, \delta > 0$ and a subset $A \subset \Gamma := B(\bar{m}, \delta) \cap M$ of full measure (i.e. $\mathcal{H}^{d-1}(A) = \mathcal{H}^{d-1}(\Gamma)$) such that*

- (i) $\Gamma \subset \text{int}(\text{dom } f)$ and $u(\Gamma \times [0, \varepsilon]) \subset \text{int}(\text{dom } f)$.
- (ii) f is Lipschitz on $u(\Gamma \times [0, \varepsilon])$.
- (iii) for every $m \in A$, one has

$$\mathcal{L}_1(\{t \in [0, \varepsilon] : u(t, m) \in \mathcal{N}\}) = 0.$$

If $\text{dom } f = \mathbb{R}^d$ and f is locally Lipschitz, then $[0, \varepsilon] \times \Gamma$ can be taken to be $[0, T] \times M$.

Proof Let $\varepsilon > 0$ such that $[0, \varepsilon] \times \Gamma$ is the neighborhood that appears in the proof of Proposition 3.7 so that assertions (i), (ii) hold. Without loss of generality, let us assume that $\mathcal{N} \subset \mathcal{O} := u([0, \varepsilon] \times \Gamma)$. Note first that $\mu = \mathcal{L}_1 \times \mathcal{H}^{d-1}$ is a Borel measure over $[0, \varepsilon] \times \Gamma$. This yields that $u : [0, \varepsilon] \times \Gamma \rightarrow \mathcal{R}$ is measurable, and therefore so is $\mathbb{1}_{\mathcal{N}} \circ u$. Furthermore, since $\mathbb{1}_{\mathcal{N}} \circ u$ is integrable, we can apply Fubini’s theorem (see, e.g., [14, Theorem 1.22]) to get that the mapping $\gamma \in \Gamma \mapsto \int_0^\varepsilon \mathbb{1}_{\mathcal{N}}(u(t, \gamma)) dt$ is \mathcal{H}^{d-1} -measurable and that

$$\mu(u^{-1}(\mathcal{N})) \equiv \int_{u^{-1}(\mathcal{N})} 1 d\mu = \int_{[0, \varepsilon] \times \Gamma} \mathbb{1}_{\mathcal{N}}(u(t, \gamma)) d\mu(t, \gamma)$$

$$= \int_{\Gamma} \int_0^{\varepsilon} \mathbb{1}_{\mathcal{N}}(u(t, \gamma)) dt d\mathcal{H}^{d-1}(\gamma).$$

Since u is Lipschitz-continuous, its Jacobian determinant $Ju(t, \gamma) = |\det(Du(t, \gamma))|$ is well-defined μ -a.e. in $[0, \varepsilon] \times \Gamma$. Thus, we can apply the co-area formula (see, e.g., [14, Theorem 3.10]) to write

$$\int_{u^{-1}(\mathcal{N})} Ju(t, \gamma) d\mu(t, \gamma) = \int_{\mathcal{O}} \underbrace{\mathcal{H}^0(u^{-1}(\mathcal{N}) \cap u^{-1}(x))}_{\in \{0,1\}} dx = \int_{\mathcal{N}} dx = \mathcal{L}_d(\mathcal{N}) = 0,$$

where the last equality comes from the fact that u is a bijection between $[0, \varepsilon] \times \Gamma$ and \mathcal{O} and consequently, $\mathcal{H}^0(u^{-1}(\mathcal{N}) \cap u^{-1}(x)) = \mathbb{1}_{\mathcal{N}}(x)$, for all $x \in \mathcal{O}$. Finally, since u is bi-Lipschitz, there exists a constant $c > 0$ such that $Ju(t, \gamma) \geq c$ for μ -almost every $(t, \gamma) \in [0, \varepsilon] \times \Gamma$. Thus,

$$c \int_{\Gamma} \int_0^{\varepsilon} \mathbb{1}_{\mathcal{N}}(u(t, \gamma)) dt d\mathcal{H}^{d-1}(\gamma) = \int_{u^{-1}(\mathcal{N})} cd\mu \leq \int_{u^{-1}(\mathcal{N})} Ju(t, \gamma)\mu(t, \gamma) = \int_{\mathcal{N}} dx = 0.$$

Then, the mapping $\gamma \mapsto \int_0^{\varepsilon} \mathbb{1}_{\mathcal{N}}(u(t, \gamma)) dt$ is zero \mathcal{H}^{d-1} -almost everywhere in Γ , and so

$$\mathcal{L}_1(\{t \in [0, \varepsilon] : u(t, \gamma) \in \mathcal{N}\}) = 0, \quad \text{for } \mathcal{H}^{d-1}\text{-a.e. } \gamma \in \Gamma.$$

The proof is complete. □

Combining Proposition 3.7 with Theorem 3.8, we will show that, for locally Lipschitz functions, the mapping $u : [0, T] \times M \rightarrow \mathcal{R}$ induces steepest descent curves almost everywhere. Indeed, for the special case where $\text{dom } f = \mathbb{R}^d$ the argument goes as follows: let \mathcal{N} be the set of non-differentiability points of f . Then by Rademacher theorem $\mathcal{L}_d(\mathcal{N}) = 0$. For every $m \in A$ (the full measure set given by Theorem 3.8) the set

$$I_m = \{t \in [0, T] : u(\cdot, m) \text{ is differentiable at } t \text{ and } f \text{ is differentiable at } u(t, m)\}$$

must be of full measure. Applying chain rule at every point $t \in I_m$ we deduce:

$$-1 = (f \circ u(\cdot, m))'(t) = |\nabla f|(u(t, m))u'(t, m).$$

Thus, $u'(t, m) = -|\nabla f|(u(t, m))^{-1}$ for almost every $t \in [0, T]$, yielding that $u(\cdot, m)$ is a steepest descent curve. The following theorem deals with the general case. The proof follows the same idea together with a localization argument.

Theorem 3.9 (Existence of steepest descent curves) *Let $f : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous quasiconvex, which is locally Lipschitz on an open set $\mathcal{O} \subset \text{dom } f$. Assume further that f satisfies (H1)–(H3). Then, for almost every $x \in \mathcal{O}$, the function f admits a steepest descent curve emanating from x .*

Proof Without loss of generality, we may assume that \mathcal{O} is convex and $0 \in \mathcal{O}$. Let $\tilde{\mathcal{N}}$ be the set of all $x \in \mathcal{O}$ for which f does not admit a steepest descent curve emanating from x . (Notice that $\tilde{\mathcal{N}} \cap \text{argmin } f$ is trivially empty.)

Let further \mathcal{N} be the set of all $x \in \mathcal{O}$ for which f is not differentiable at x . By Rademacher’s theorem $\mathcal{L}_d(\mathcal{N}) = 0$. We shall show that the set $\tilde{\mathcal{N}}$ is also null.

To this end, for every $j \in \mathbb{N}$, let us define

$$B_j = \left(1 - \frac{1}{j}\right) \overline{\mathcal{O}}.$$

Clearly, $B_j \subset \mathcal{O}$ and $\mathcal{O} = \bigcup_{j \in \mathbb{N}} B_j$. Now, fix $\alpha \in (\inf_{\mathcal{O}} f, \sup_{\mathcal{O}} f)$ and set

$$K_{\alpha,n,j} = (B_j \cap [f \leq \alpha]) \setminus \text{int}([f \leq \inf_{\mathcal{O}} f + \frac{1}{n}]); \text{ and}$$

$$\tilde{\mathcal{N}}_{\alpha,n,j} = \tilde{\mathcal{N}} \cap K_{\alpha,n,j}$$

Notice that $K_{\alpha,n,j}$ is compact and does not intersect $\text{argmin} f$. Consequently, by hypothesis (H2) we deduce that there exists $c > 0$ such that $\|\nabla f(x)\| > c$ for every $x \in K_{\alpha,n,j} \setminus \mathcal{N}$.

Denote by $h_{\alpha,n,j} : \mathbb{R}^d \rightarrow \mathbb{R}$ any Lipschitz extension of f from $K_{\alpha,n,j}$ to \mathbb{R}^d . Applying the co-area formula, we deduce

$$\int_{-\infty}^{\infty} \mathcal{H}^{d-1}(\tilde{\mathcal{N}}_{\alpha,n,j} \cap h_{\alpha,n,j}^{-1}(t)) dt = \int_{\tilde{\mathcal{N}}_{\alpha,n,j}} \|\nabla h_{\alpha,n,j}(x)\| dx$$

$$= \int_{\tilde{\mathcal{N}}_{\alpha,n,j}} \|\nabla f(x)\| dx \geq c \mathcal{L}_d(\tilde{\mathcal{N}}_{\alpha,n,j}).$$

Therefore

$$\mathcal{L}_d(\tilde{\mathcal{N}}_{\alpha,n,j}) \leq \frac{1}{c} \int_{-\infty}^{\infty} \mathcal{H}^{d-1}(\tilde{\mathcal{N}}_{\alpha,n,j} \cap h_{\alpha,n,j}^{-1}(t)) dt = \frac{1}{c} \int_{\inf_{\mathcal{O}} f + \frac{1}{n}}^{\alpha} \mathcal{H}^{d-1}(\tilde{\mathcal{N}}_{\alpha,n,j} \cap f^{-1}(t)) dt.$$

Let us fix $t \in [\inf_{\mathcal{O}} f + 1/n, \alpha]$, set $M = f^{-1}(t)$ and choose $m \in f^{-1}(t) \cap \mathcal{O}$. There exist $\delta > 0$ and $\varepsilon > 0$ such that $\Gamma_m := M \cap B(m, \delta) \subset \text{int}(\text{dom} f)$ and $u([0, \varepsilon] \times \Gamma_m) \subset \mathcal{O}$. Let further $A_m \subset \Gamma_m$ be the full measure subset given by Theorem 3.8 for $\Gamma = \Gamma_m$ and the null measure set \mathcal{N} . Then, for every $\gamma \in A_m$ and almost every $t \in [0, \varepsilon]$, $u(\cdot, \gamma)$ is differentiable at t and f is differentiable at $u(t, \gamma)$.

Recall that the boundary of the set $S := [f \leq f(u(t, \gamma))]$ is a $C^{1,1}$ -manifold around $u(t, \gamma)$ for every $t \in [0, \varepsilon]$ (cf. Lemma 3.3). We deduce that $N(S, u(t, \gamma)) = \mathbb{R}_+ \{\nabla f(u(t, \gamma))\}$ whenever f is differentiable at $u(t, \gamma)$. Then, for almost all $t \in [0, \varepsilon]$, we have that $\frac{d}{dt} u(t, \gamma) \in \mathbb{R}_+ \{\nabla f(u(t, \gamma))\}$. Using this fact and Proposition 2.1, we can apply chain rule to deduce that for almost all $t \in [0, \varepsilon]$

$$1 = -(f \circ u(\cdot, \gamma))'(t) = \left\| \frac{d}{dt} u(t, \gamma) \right\| \|\nabla f(u(t, \gamma))\| = \left\| \frac{d}{dt} u(t, \gamma) \right\| |\nabla f|(u(t, \gamma)).$$

We conclude that for every $\gamma \in A_m$, $u(\cdot, \gamma)$ is a steepest descent curve of f emanating from γ , and so $\tilde{\mathcal{N}}_{\alpha,n,j} \cap A_m = \emptyset$. Since $M \cap \mathcal{O} = f^{-1}(t) \cap \mathcal{O}$ is σ -compact, it can be covered by countably many sets $\{\Gamma_{m_k} : k \in \mathbb{N}\}$, yielding

$$\mathcal{H}^{d-1}(\tilde{\mathcal{N}}_{\alpha,n,j} \cap f^{-1}(t)) \leq \mathcal{H}^{d-1}(f^{-1}(t) \setminus \bigcup A_{m_k}) = 0.$$

Since the latter conclusion holds for every $t \in [\inf_{\mathcal{O}} f + 1/n, \alpha]$, we deduce that $\mathcal{L}_d(\tilde{\mathcal{N}}_{\alpha,n,j}) = 0$. Taking $n \rightarrow \infty$, $j \rightarrow \infty$ and $\alpha \nearrow \sup_{\mathcal{O}} f$, we deduce that

$$\mathcal{L}_d(\tilde{\mathcal{N}} \setminus (\text{argmin}_{\mathcal{O}} f \cup \text{argmax}_{\mathcal{O}} f)) = 0.$$

Note that (H2) yields that all level sets of f (except possibly $\operatorname{argmin} f$) must have empty interior. Thus, $\operatorname{argmax}_{\mathcal{O}} f = [f = \sup_{\mathcal{O}} f] \cap \mathcal{O}$ has null measure. Moreover, if $\inf_{\mathcal{O}} f > \min f$, then $\operatorname{argmin}_{\mathcal{O}} f$ also has null measure, while if $\inf_{\mathcal{O}} f = \min f$, then $\tilde{\mathcal{N}} \cap \operatorname{argmin}_{\mathcal{O}} f = \emptyset$. Thus,

$$\begin{aligned} \mathcal{L}_d(\tilde{\mathcal{N}}) &= \mathcal{L}_d(\tilde{\mathcal{N}} \setminus (\operatorname{argmin}_{\mathcal{O}} f \cup \operatorname{argmax}_{\mathcal{O}} f)) + \mathcal{L}_d(\tilde{\mathcal{N}} \cap \operatorname{argmax}_{\mathcal{O}} f) \\ &\quad + \mathcal{L}_d(\tilde{\mathcal{N}} \cap \operatorname{argmin}_{\mathcal{O}} f) = 0. \end{aligned}$$

The proof is complete. □

4 Regularizing Locally Lipschitz Quasiconvex

In order to apply the results of Sect. 3, we present a regularization scheme based on the max-convolution operator (see, e.g., [22]). Given two functions $f, g : \mathbb{R}^d \rightarrow \overline{\mathbb{R}}$ the max-convolution (or sublevel-convolution) of f and g , denoted by $f \diamond g$, is defined as

$$(f \diamond g)(x) := \inf_{w \in \mathbb{R}^d} \max\{f(x - w), g(w)\}. \tag{4.1}$$

Notice that whenever the infimum of (4.1) is exact, we have

$$[f \diamond g \leq \alpha] = [f \leq \alpha] + [g \leq \alpha] \quad \text{and} \quad [f \diamond g < \alpha] = [f < \alpha] + [g < \alpha], \tag{4.2}$$

for every $\alpha \in \mathbb{R}$.

In what follows we simply denote by $\mathbb{B} \equiv \mathbb{B}_d$ the closed unit ball of \mathbb{R}^d . Let us assume $\inf f = 0$. Let $\varepsilon > 0$ and let us denote by $I_{\varepsilon\mathbb{B}}$ the indicator function of $\varepsilon\mathbb{B}$, that is,

$$I_{\varepsilon\mathbb{B}}(x) = \begin{cases} 0, & \text{if } x \in \varepsilon\mathbb{B} \\ +\infty, & \text{if } x \notin \varepsilon\mathbb{B}. \end{cases}$$

We focus on a particular max-convolution with $g \equiv I_{\varepsilon\mathbb{B}}$, namely, we study the function $f_\varepsilon = f \diamond I_{\varepsilon\mathbb{B}}$ defined by

$$f_\varepsilon(x) = (f \diamond I_{\varepsilon\mathbb{B}})(x) = \inf_{w \in \mathbb{B}} f(x - \varepsilon w), \quad \forall x \in \mathbb{R}^d. \tag{4.3}$$

If $\inf f > -\infty$, we can easily adapt the definition of f_ε by setting

$$f_\varepsilon = \inf f + (f - \inf f) \diamond I_{\varepsilon\mathbb{B}}.$$

In both cases, the formulae of the sublevel sets is preserved:

$$[f_\varepsilon \leq \alpha] = [f \leq \alpha] + \varepsilon\mathbb{B}, \quad \text{for every } \alpha \in \mathbb{R}.$$

However, if f is not bounded from below, the max-convolution loses that key property. Thus, for the general case, we consider the following definition.

Definition 4.1 For a lower semicontinuous function $f : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$ and $\varepsilon > 0$ we define the regularized function $f_\varepsilon : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$ as follows:

$$f_\varepsilon(x) = \inf_{w \in \mathbb{B}} f(x - \varepsilon w).$$

Notice that f_ε is the unique function satisfying $[f_\varepsilon \leq \alpha] = [f \leq \alpha] + \varepsilon\mathbb{B}$, for every $\alpha \in \mathbb{R}$.

In this section we will study the class

$$\mathcal{Q} = \{f : \mathbb{R}^d \rightarrow \mathbb{R} \mid f \text{ is quasicontex and locally Lipschitz}\}. \tag{4.4}$$

Let us first focus our attention on quasicontex functions satisfying $\min f = 0$. The general case will be treated in forthcoming Theorem 4.5. The next proposition surveys some relevant properties of the above max-convolution that we will use in the subsequent development.

Proposition 4.2 *Assume that $f : \mathbb{R}^d \rightarrow \mathbb{R} \cup \{+\infty\}$ is lower semicontinuous, quasicontex and $\min f = 0$. Let $\varepsilon > 0$ and consider the max-convolution $f_\varepsilon = f \diamond I_{\varepsilon\mathbb{B}}$. Then, the following properties hold:*

- (i) $[f_\varepsilon \leq \alpha] = [f \leq \alpha] + \varepsilon\mathbb{B}$, for every $\alpha \geq 0$; therefore $\text{bd}[f_\varepsilon \leq \alpha]$ is an ε -prox-regular $\mathcal{C}^{1,1}$ -submanifold.
- (ii) For $\varepsilon_1, \varepsilon_2 > 0$ such that $\varepsilon_1 + \varepsilon_2 = \varepsilon$, one has that

$$f_\varepsilon = f_{\varepsilon_1} \diamond I_{\varepsilon_2\mathbb{B}} = f_{\varepsilon_2} \diamond I_{\varepsilon_1\mathbb{B}}.$$

- (iii) For each $x \in \mathbb{R}^d$, $f_\varepsilon(x) = f(z)$, where $z = \text{proj}(x; [f \leq f_\varepsilon(x)])$. Moreover, one has that

$$|\nabla f_\varepsilon|(x) \geq |\nabla f|(z).$$

- (iv) If f is finite-valued and locally Lipschitz-continuous, then f_ε is also locally Lipschitz-continuous.

Proof Assertion (i) follows from the fact that the infimum in (4.3) is exact and by the $\mathcal{C}^{1,1}$ -smoothness of the distance function $d(\cdot, [f \leq \alpha])$ on the set $\mathbb{R}^n \setminus [f \leq \alpha]$, which is due to the convexity of $[f \leq \alpha]$: indeed, it suffices to notice that $\text{bd}([f \leq \alpha]) = [d(\cdot, [f \leq \alpha]) = \varepsilon]$.

To establish (ii), we consider the following straightforward computation:

$$\begin{aligned} (f_{\varepsilon_1} \diamond I_{\varepsilon_2\mathbb{B}})(x) &= \inf_{w_2 \in \mathbb{B}} f_{\varepsilon_1}(x - \varepsilon_2 w_2) \\ &= \inf_{w_2 \in \mathbb{B}} \inf_{w_1 \in \mathbb{B}} f(x - \varepsilon_1 w_1 - \varepsilon_2 w_2) = \inf_{w \in \mathbb{B}} f(x - \varepsilon w) = f_\varepsilon(x). \end{aligned}$$

The assertion follows by interchanging the roles of ε_1 and ε_2 .

Let us now deal with (iii). The first part of this assertion is straightforward. Indeed, since $z \in [f \leq f_\varepsilon(x)]$, one has that $f(z) \leq f_\varepsilon(x)$. On the other hand, since the infimum in (4.1) is exact, there exists $w \in \varepsilon\mathbb{B}$ such that $f_\varepsilon(x) = f(x - w)$. Thus, the distance of x to $[f \leq f_\varepsilon(x)]$ is smaller than ε . This yields that $\|x - z\| \leq \varepsilon$, and so $f_\varepsilon(x) \leq f(z)$.

Now, let $x \in \text{dom } f_\varepsilon$ and let $z = \text{proj}(x; [f \leq f_\varepsilon(x)])$. Set $w = x - z$ and notice (as mentioned before) that $\|w\| \leq \varepsilon$. Therefore, for every $y \in \text{dom } f$, $f_\varepsilon(y + w) \leq f(y)$. Let us also notice that

$$\{y + w : y \in [f \leq f(z)]\} \subset [f_\varepsilon \leq f_\varepsilon(x)].$$

It follows that

$$|\nabla f|(z) = \limsup_{[f \leq f(z)] \ni y \rightarrow z} \frac{f(z) - f(y)}{d(z, y)} = \limsup_{[f \leq f_\varepsilon(x)] \ni y \rightarrow z} \frac{f_\varepsilon(x) - f(y)}{d(x, y + w)}$$

$$\begin{aligned} &\leq \limsup_{[f \leq f(z)] \ni y \rightarrow z} \frac{f_\varepsilon(x) - f_\varepsilon(y + w)}{d(x, y + w)} \\ &\leq \limsup_{[f_\varepsilon \leq f_\varepsilon(x)] \ni y \rightarrow x} \frac{f_\varepsilon(x) - f_\varepsilon(y)}{d(x, y)} = |\nabla f_\varepsilon|(x). \end{aligned}$$

It remains to establish (iv). This follows from a minor adaptation of [22, Proposition 3.1]. Choose $x \in \mathbb{R}^d$ and $\delta > 0$, and take $B = \bar{B}(x, \delta + \varepsilon) = \bar{B}(0, \delta) + \varepsilon\mathbb{B}$. Since B is compact, the function f is Lipschitz continuous on B . Let us denote by $L_B > 0$ this Lipschitz constant and take $y, z \in B(x, \delta)$. Let $w \in \varepsilon\mathbb{B}$ be such that $f_\varepsilon(y) = f(y + w)$. Then, since $y + w, z + w \in B$ we deduce

$$f_\varepsilon(z) \leq f(z + w) \leq f(y + w) + L_B\|y - z\| = f_\varepsilon(y) + L_B\|y - z\|.$$

Interchanging the roles of y and z in the above development, we obtain

$$|f_\varepsilon(z) - f_\varepsilon(y)| \leq L_B\|y - z\|,$$

which yields that f_ε is Lipschitz continuous on $B(x, \delta)$. The proof is complete. □

Let us now choose $x_0 \in \mathbb{R}^d$ such that $|\overline{\nabla f}|(x_0) > \ell$ (recall definition in (2.4)). Pick $\delta > 0$ sufficiently small such that

$$|\nabla f|(y) > \ell, \text{ for every } y \in \bar{B}(x_0, \delta) \subset \text{dom } f \tag{4.5}$$

and consider the function

$$h := f + I_{\bar{B}(x_0, \delta)}. \tag{4.6}$$

Notice that since $\min f = 0$, one has that $\min h \geq 0$.

Lemma 4.3 *Assume that (4.5) holds and h is given by (4.6). Then the limiting slope $|\overline{\nabla h}|$ is strictly positive on $\text{dom } h \setminus \text{argmin } h$.*

Proof Set $C := \text{dom } h = \bar{B}(x_0, \delta)$. Choose $z \in \text{dom } h \setminus \text{argmin } h$ set $\alpha = f(z)$ and $\beta \in (\min h, \alpha)$. Note that the set $[\min h < f < \beta] \cap C$ has nonempty interior and consequently, by construction,

$$\begin{aligned} \sup_{y \in [f \leq \beta] \cap C} d(y, \mathbb{R}^d \setminus C) &\geq d(\text{proj}(x_0; [f \leq \beta]), \mathbb{R}^d \setminus C) \\ &= \delta - d(x_0, [f \leq \beta]) > 0. \end{aligned}$$

Thus, applying [15, Lemma 1], we deduce

$$d(z, [h \leq \beta]) = d(z, [f \leq \beta] \cap C) \leq \frac{\delta + d(x_0, [f \leq \beta])}{\delta - d(x_0, [f \leq \beta])} d(z, [f \leq \beta]).$$

Taking the limit $\beta \nearrow \alpha = f(z)$, we obtain

$$|\nabla h|(z) = \limsup_{\beta \nearrow \alpha} \frac{\alpha - \beta}{d(z, [h \leq \beta])} \geq \limsup_{\beta \nearrow \alpha} \frac{\delta - d(x_0, [f \leq \beta])}{\delta + d(x_0, [f \leq \beta])} \frac{\alpha - \beta}{d(z, [f \leq \beta])}$$

$$= \frac{\delta - d(x_0, [f \leq \alpha])}{\delta + d(x_0, [f \leq \alpha])} |\nabla f|(z) \geq \underbrace{\frac{\delta - d(x_0, [f \leq \alpha])}{\delta + d(x_0, [f \leq \alpha])}}_{:=\phi(\alpha)} \ell.$$

Since the function $z \mapsto \phi(h(z))$ is continuous and strictly positive on $\text{dom } h \setminus \text{argmin } h$ the assertion follows. □

Lemma 4.4 *For every $\varepsilon > 0$, the function $h_\varepsilon = h \diamond I_{\varepsilon\mathbb{B}}$ is quasiconvex, coercive, and satisfies (H1)–(H3).*

Proof Since f is quasiconvex, it is straightforward that h is quasiconvex. Moreover, by construction $\text{dom } h \equiv \bar{B}(x_0, \delta)$. Thus, Proposition 4.2 entails that h_ε is quasiconvex and coercive, where the last property follows from the fact that the domain $\text{dom } h_\varepsilon$ coincides with the compact ball $\bar{B}(x_0, \delta) + \varepsilon\mathbb{B}$.

Notice further that coercivity and Proposition 4.2 (i) yield that the function h_ε verifies (H1) and (H3), with $r = \varepsilon$, while (H2) follows by Lemma 4.3 and Proposition 4.2 (iii).

The proof is complete. □

We are now ready to establish the main result of this section, which provides steepest descent curves for almost every point of the regularized function f_ε (where $f \in \mathcal{Q}$ is not necessarily assumed to be bounded from below) stemming from non-critical points, in the sense of (4.7) below. Let us set:

$$\mathcal{U}_\varepsilon := \{x \in \mathbb{R}^d : \overline{|\nabla f|}(\text{proj}(x; [f \leq f_\varepsilon(x)])) > 0\} \tag{4.7}$$

Theorem 4.5 *Let $f \in \mathcal{Q}$ (not necessarily bounded from below) and $\varepsilon > 0$. The regularized function f_ε admits steepest descent curves emanating from almost every $x \in \mathcal{U}_\varepsilon$. In particular, if f verifies (H2), then f_ε admits steepest descent curves emanating from almost every $x \in \mathbb{R}^d$.*

Proof Let $x \in \mathcal{U}_\varepsilon$, and let $z = \text{proj}(x, [f \leq f_\varepsilon(x)])$. Then, there exist $\delta, \ell > 0$ such that for all $z' \in B(z, 2\delta)$, we have $|\nabla f|(z') > \ell$.

Take $r = \min_{\bar{B}(z, 2\delta)} f$ and note that f and $g = \max\{f, r\}$ coincide over $B(z, 2\delta)$, and so f_ε and g_ε coincide on a neighborhood of x . Thus, we can replace f by g and assume, without losing any generality, that $r = \min g = 0$.

Take $h = g + I_{\bar{B}(z, \delta)}$. Then, by Lemma 4.3, h verifies (H2). By invoking [2, Example 4.1] and Lemma 3.6, we deduce that the mapping $y \mapsto \text{proj}(y, [h \leq g_\varepsilon(y)])$ is continuous over the set $[g_\varepsilon > \min h]$. This yields that there exists $\eta > 0$ such that

$$\text{proj}(y, [h \leq g_\varepsilon(y)]) \subset z + \frac{\delta}{2}\mathbb{B}, \quad \forall y \in B(x, \eta).$$

In particular, for all $y \in B(x, \eta)$ one has that $\text{proj}(y, [g \leq g_\varepsilon(y)]) = \text{proj}(y, [h \leq g_\varepsilon(y)])$ and consequently

$$\begin{aligned} g_\varepsilon(y) &\geq h_\varepsilon(y) = \inf_{w \in y + \varepsilon\mathbb{B}} h(y - w) \\ &\geq \inf_{w \in y + \varepsilon\mathbb{B}} g(y - w) = g_\varepsilon(y). \end{aligned}$$

Thus, g_ε and h_ε coincide in $B(x, \eta)$ and since g_ε is finite, $B(x, \eta) \subset \text{dom } h_\varepsilon$. In particular, h_ε is locally Lipschitz on $B(x, \eta)$. Now, applying Lemma 4.4 and Theorem 3.9, we deduce

that h_ε , and so g_ε , admits steepest descent curves emanating from almost every point in $B(x, \eta)$. The proof is complete. \square

(Open questions) The question of characterizing locally Lipschitz functions that admit steepest descent curves on their domains is challenging. This is open even for the class of quasiconvex functions, where a sufficient condition was obtained in [13]: namely, this happens whenever the slope mapping $x \mapsto |\nabla f|(x)$ is lower semicontinuous, since in this case, every near-steepest descent curve is also a steepest descent curve. Concurrently, it is not known if Theorem 4.5 presented hereby is tight, or if the regularized quasiconvex functions f_ε admits steepest descent curves at every point. In addition, it is still unclear if Theorem 4.5 holds true for extended-valued quasiconvex functions that are locally Lipschitz on their domains. The main obstruction seems to be the max-convolution f_ε does not directly inherit the Lipschitz continuity near the boundary of the regularized domain $\text{dom } f_\varepsilon$. Last, but not least, we do not dispose a satisfactory characterization for the slope to be lower semicontinuous. In particular, it is not known if the slope of a regularized function $x \mapsto |\nabla f_\varepsilon|(x)$ is lower semicontinuous or not.

Funding Open access funding provided by TU Wien (TUW). This research was funded in whole or in part by the Austrian Science Fund (FWF) [DOI 10.55776/P36344N]. For open access purposes, the first author has applied a CC BY public copyright license to any author accepted manuscript version arising from this submission. The second author was partially funded by the grant FONDECYT 11220586 (ANID, Chile) and by the BASAL fund FB210005 (Centers of excellence, ANID, Chile).

Data availability declaration The manuscript has no associated data.

Declarations

Competing Interest declaration There are no competing interests related to this research.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Ambrosio, L., Gigli, N., Savaré, G.: Gradient Flows in Metric Spaces and in the Space of Probability Measures, 2nd edn. Lectures in Mathematics ETH Zürich. Birkhäuser, Basel (2008)
2. Attouch, H., Wets, R.: Quantitative stability of variational systems. II. A framework for nonlinear conditioning. *SIAM J. Optim.* **3**(2), 359–381 (1993)
3. Attouch, H., Buttazzo, G., Michaille, G.: Variational Analysis in Sobolev and BV Spaces, 2nd edn. MOS-SIAM Series on Optimization, vol. 17. Mathematical Optimization Society, Philadelphia (2014)
4. Aussel, D.: New developments in quasiconvex optimization. In: Fixed Point Theory, Variational Analysis, and Optimization, pp. 171–205. CRC Press, Boca Raton (2014)
5. Azé, D., Corvellec, J.-N.: Characterizations of error bounds for lower semicontinuous functions on metric spaces. *ESAIM Control Optim. Calc. Var.* **10**(3), 409–425 (2004)
6. Colombo, G., Thibault, L.: Prox-regular sets and applications. In: Handbook of Nonconvex Analysis and Applications, pp. 99–182. Int. Press, Somerville (2010)
7. Czarnecki, M.-O., Thibault, L.: Sublevel representations of epi-Lipschitz sets and other properties. *Math. Program.* **168**(1–2), 555–569 (2018)

8. Daniilidis, A., Drusvyatskiy, D.: Sweeping by a tame process. *Ann. Inst. Fourier (Grenoble)* **67**(5), 2201–2223 (2017)
9. Daniilidis, A., Hadjisavvas, N.: On the subdifferentials of quasiconvex and pseudoconvex functions and cyclic monotonicity. *J. Math. Anal. Appl.* **237**(1), 30–42 (1999)
10. Daniilidis, A., Ley, O., Sabourau, S.: Asymptotic behaviour of self-contracted planar curves and gradient orbits of convex functions. *J. Math. Pures Appl.* (9) **94**(2), 183–199 (2010)
11. Daniilidis, A., Drusvyatskiy, D., Lewis, A.S.: Orbits of geometric descent. *Can. Math. Bull.* **58**(1), 44–50 (2015)
12. De Giorgi, E., Marino, A., Tosques, M.: Problems of evolution in metric spaces and maximal decreasing curve. *Atti Accad. Naz. Lincei, Rend. Cl. Sci. Fis. Mat. Nat.* (8) **68**(3), 180–187 (1980)
13. Drusvyatskiy, D., Ioffe, A.D., Lewis, A.S.: Curves of descent. *SIAM J. Control Optim.* **53**(1), 114–138 (2015)
14. Evans, L., Gariepy, R.: *Measure Theory and Fine Properties of Functions*. Textbooks in Mathematics. CRC Press, Boca Raton (2015). Revised edition
15. Hoffmann, A.: The distance to the intersection of two convex sets expressed by the distances to each of them. *Math. Nachr.* **157**, 81–98 (1992)
16. Ioffe, A.: *Variational Analysis of Regular Mappings*. Springer Monographs in Mathematics. Springer, Cham (2017)
17. Longinetti, M., Manselli, P., Venturi, A.: On steepest descent curves for quasi convex families in \mathbb{R}^n . *Math. Nachr.* **288**(4), 420–442 (2015)
18. Moreau, J.-J.: Evolution problem associated with a moving convex set in a Hilbert space. *J. Differ. Equ.* **26**(3), 347–374 (1977)
19. Pérez-Aros, P., Vilches, E.: An enhanced Baillon-Haddad theorem for convex functions defined on convex sets. *Appl. Math. Optim.* **83**(3), 2241–2252 (2021)
20. Peypouquet, J., Sorin, S.: Evolution equations for maximal monotone operators: asymptotic analysis in continuous and discrete time. *J. Convex Anal.* **17**(3–4), 1113–1163 (2010)
21. Rockafellar, R.T., Wets, R.: *Variational Analysis*. Grundlehren der mathematischen Wissenschaften, vol. 317. Springer, Berlin (1998)
22. Seeger, A., Volle, M.: On a convolution operation obtained by adding level sets: classical and new results. *RAIRO. Rech. Opér.* **29**(2), 131–154 (1995)
23. Thibault, L.: Sweeping process with regular and nonregular sets. *J. Differ. Equ.* **193**(1), 1–26 (2003)
24. Thibault, L.: *Unilateral Variational Analysis in Banach Spaces: (Part I: General Theory. Part II: Special Classes of Functions and Sets)*. World Scientific Publishing Co., Inc., River Edge (2023)
25. Villani, C.: *Optimal Transport*. Grundlehren der mathematischen Wissenschaften, vol. 338. Springer, Berlin (2009)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.