TECHNISCHE UNIVERSITÄT WIEN

Variational Analysis, Dynamics and Operations Research

SWM VADOR

Steepest geometric descent for regularized quasiconvex functions

A. Daniilidis, D. Salas

Research Report 2023-07

Oktober 2023

ISSN 2521-313X

Variational Analysis, Dynamics and Operations Research
Institute of Statistics and Mathematical Methods in Economics
TU Wien

Research Unit VADOR
Wiedner Hauptstraße 8 / E105-04
1040 Vienna, Austria
E-mail: vador@tuwien.ac.at

Steepest geometric descent for regularized quasiconvex functions

Aris Daniilidis, David Salas

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 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.

Key words. Steepest descent curves, quasiconvex functions, max-convolution, sweeping process.

AMS 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 smooth function $f : \mathbb{R}^d \to \mathbb{R}$ we call steepest descent curve the solution of the gradient flow equation

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

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 of steepest descent curves can still be established for convex functions. Indeed, if $f : \mathbb{R}^d \to \mathbb{R}$ 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., [2]). Similarly, existence of solutions for several gradient descent and proximal methods are often based on convexity (see, e.g., [18]).

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 [10]. 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: semi-algebraic functions [11, 14], geodesically convex functions in metric spaces [1] and smooth functions on Riemannian manifolds (see, e.g., [23]). However, existence of steepest descent curves is in general hard to verify, even for Lipschitz functions in $\mathbb{R}^d$. Due to this obstruction, the authors in [8, 15] consider the more general notion of trajectories of a convex foliation (terminology introduced in [9]) and establish existence of such orbits (see [8, Theorem 2.6] *e.g*). In case the foliation is given by the sublevel set of a quasiconvex function, the above orbits are call orbits of geometric descent. Their connection with steepest descent orbits has been explored in [11, 14]: these curves fail to be steepest descent curves in general, but instead correspond to what the authors in [11] 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 extended-valued quasi-convex functions which are locally Lipszchitz on their domain. 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 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., [20] and the references therein). Indeed, we develop a technique which allows to provide, for any quasiconvex function $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ that is lower semicontinous and locally Lipschitz on its domain, and any $\varepsilon > 0$, 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 [8, 11]. 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 Section 2 we fix our terminology and quote some preliminary results. In Section 3, under adequate assymptions 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 Section 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 $\mathrm{int}(A)$, $\overline{A}$, $\mathrm{bd}(A)$ and A° its interior, closure, boundary and (negative) polar set, respectively.

For a function $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ 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 $\operatorname{dom} f$, that is, $\operatorname{dom} f = \{x \in \mathbb{R}^d \mid f(x) < +\infty\}$.

A function $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ 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. In what follows, we study functions belonging to the following class:

$$\mathcal{Q} = \{f : \mathbb{R}^d \to \mathbb{R}_+ \cup \{+\infty\} \mid f \text{ is lsc quasiconvex and locally Lipschitz on} \operatorname{dom} f\}. \tag{2.3}$$

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

$$|\nabla f|(x) := \begin{cases} \limsup\limits_{y \to x} \dfrac{(f(x) - f(y))^+}{\|x - y\|}, & \text{if } x \in \operatorname{dom} f \\ +\infty, & \text{otherwise,} \end{cases} \tag{2.4}$$

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., [11]). Thus, we consider the limiting slope $\overline{|\nabla f|}$ as the lower semicontinuous closure of $|\nabla f|$, that is,

$$\overline{|\nabla f|}(x) = \liminf_{y \to x} |\nabla f|(y). \tag{2.5}$$

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

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

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

is open, proj_S is well-defined on $U_{\rho(\cdot)}(S)$ and d_S^2 is of class $\mathcal{C}^1$ 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., [22]) 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, Frechet, Limiting, Clarke, and the Polar of the Bouligand tangent cone). 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, for $S \subset \mathbb{R}^d$ prox-regular and $x \in S$, we define the *(Clarke) tangent cone* and the *(Clarke) normal cone* of S at x as

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

where Liminf is the inferior limit of sets in the sense of Painlevé-Kuratowski (see, e.g., [22]).

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

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 $r > 0$ (i.e., $K(t)$ 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., [21]). 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 three statements are classic in the literature, while the fourth one follows from [11, Theorem 3.4 and Claim 3.6].

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

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

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

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

(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

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

(v) If u is differentiable at t and f is differentiable at $u(t)$ then $-\dot{u}(t) \in \mathbb{R}_+\{\nabla f(u(t))\}$.

Last but not least, following [12], given a Lipschitz function $f : \mathbb{R}^d \to \mathbb{R}^d$, we denote by $Df(x)$ its derivative at a point $x \in \mathbb{R}^d$, and define the *Jacobian of* f as

$$Jf(x) = |\det(Df(x))|, \tag{2.10}$$

at each point $x \in \mathbb{R}^d$ where the derivative $Df(x)$ exists. 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 $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ which is locally Lipschitz-continuous on its domain. We further consider functions that satisfy the following regularity hypotheses:

(H1) For every $\alpha \in (\inf f, \sup f)$, the sublevel set $S(\alpha) := [f \le \alpha]$ is compact and has nonempty interior.

(H2) For every $x \in \operatorname{dom} f \setminus \operatorname{argmin} f$, the slope of f is bounded away from zero near x, that is, there exists $\delta, \ell > 0$ such that

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

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

Notice that hypothesis (H1) yields that $\operatorname{argmin} f$ is nonempty while (H2) together with continuity of f on its domain yield that for every $x \in \operatorname{int}(\operatorname{dom} f)$ and every $\delta > 0$ such that $B(x,\delta) \subset \operatorname{dom} f$, it holds:

$$\operatorname{bd}[f \le f(x)] \bigcap B(x,\delta) \;=\; [f = f(x)] \bigcap B(x,\delta). \tag{3.1}$$

Indeed, continuity of f directly entails the inclusion. If the reverse inclusion does not hold, then there would exists $z \in \mathbb{R}^d$ and $\varepsilon > 0$ small enough such that $B(z,\varepsilon) \subset [f = f(x)] \cap B(x,\delta)$. This would yield that $|\nabla f|(z) = 0$, which would be a contradiction.

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

Example 3.1. Let $S : [0,2] \rightrightarrows \mathbb{R}^2$ be a convex valued function defined by

$$S(t) := \begin{cases} \operatorname{co}\Big(\big((2-t)\mathbb{B}_2\big) \bigcup \big(\overbrace{(0,3-2t) + (1-t)\mathbb{B}_2}^{\bar{B}\left((0,3-2t),1-t\right)}\big)\Big), & \text{if } t \in [0,1), \\ (2-t)\,\mathbb{B}_2, & \text{if } t \in [1,2]. \end{cases}$$

We set $f(x) = \inf\{t : x \in S(t)\}$. Then the function f is lower semicontinuous, quasiconvex and locally Lipschity on its domain. Moreover, for every $t \in [0,2]$, the set $[f = t] \equiv \operatorname{bd} S(t)$ is a smooth manifold, whose the internal curvature is $(1-t)$ for $t \in [0,1)$, and $2-t$ for $t \in [1,2]$. $\Diamond$

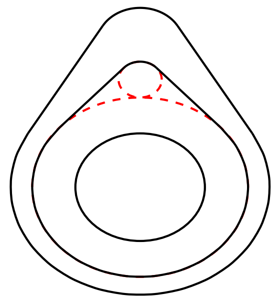

Figure 1: Boundary of $S(t)$ for $t = 1/2$, $t = 3/4$ and $t = 5/4$.

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

Proof. Let $S = [f \leq \alpha]$ and $U = \mathbb{R}^d \setminus \mathrm{int}(S)$. Since U is prox-regular, using [19, Theorem 6.42] (and noting that prox-regularity entails regularity in the sense of [19, Definition 6.4]), we get that

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

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

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

Combining the above equations we deduce that $T_{\mathrm{bd}\, S}(x)$ is a vector space. Using [22, Proposition 2.113.(a5)], we get $\mathbb{R}^d \setminus T_S(x) \subset T_U(x)$ and consequently $\mathrm{bd}\, T_S(x) \subset T_{\mathrm{bd}\, S}(x)$. We conclude that $T_{\mathrm{bd}\, S}(x)$ is of codimension 1. Therefore, for every $x \in \mathrm{bd}\, S$, 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, $N(\mathrm{bd}\, S, x) = \mathbb{R}\hat{n}(x)$. Since S is convex, the mapping $\hat{n} : \mathrm{bd}\, S \to \mathbb{S}_d$ is continuous. Furthermore, since U is r-prox-regular for some $r > 0$, then the set $U_\varepsilon = \{z \ : \ d_S(x) \geq \varepsilon\}$ must be $(r + \varepsilon)$-prox-regular. By noting that

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

we deduce that $\hat{n} : \mathrm{bd}\, S \to \mathbb{S}_d$ is also Lipschitz-continuous and the proof is complete. $\square$

Remark 3.3. An alternative proof of the above lemma can be derived using the enhanced Baillon-Haddad theorem of [17], 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 $U = \mathbb{R}^d \setminus \mathrm{int}(S)$.

Example 3.4. Smoothness of $\mathrm{bd}[f \leq \alpha]$ does not entail that this set coincides with the corresponding level set $[f = \alpha]$. Discrepancies may appear due to the cutting effect of $\mathrm{dom}\, f$ as illustrates the following example: Set $D = \mathrm{co}(\mathbb{B}_2 \cup ((3,0)+\mathbb{B}_2)) \subset \mathbb{R}^2$, and define $f : \mathbb{R}^2 \to \{+\infty\}$ 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 $[f \leq t] = \mathrm{co}(\mathbb{B}_2 \cup ((t,0)+\mathbb{B}_2))$, for $t \in [0,3]$, $[f \leq t] = \mathrm{dom}\, f = D$ if $t \geq 3$, and $[f \leq t] = \emptyset$ if $t < 0$ (see Figure 2 below).

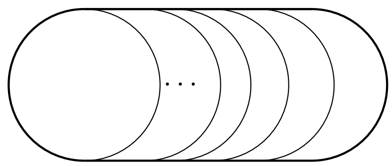

Figure 2: The sublevel sets have smooth boundaries but these latter do not coincide with the level sets.

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 $(x,y) \in \mathrm{dom}\, f \setminus \mathrm{argmin}\, f = D \setminus \mathbb{B}_2$. The first and third hypotheses follows from the construction. $\Diamond$

In what follows, for $\alpha_1 < \alpha_2$ we set $T := \alpha_2 - \alpha_1 > 0$ and consider the *annulus* set

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

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

$$\begin{cases} \widehat{\mathcal{S}} : [0,T] \rightrightarrows \mathbb{R}^d \\ \widehat{\mathcal{S}}(t) := S(\alpha_2 - t) \equiv [f \leq \alpha_2 - t] \end{cases} \quad \text{and} \quad \begin{cases} \widehat{\mathcal{U}} : [-T,0] \rightrightarrows \mathbb{R}^d \\ \widehat{\mathcal{U}}(\tau) := U(\alpha_2 + \tau) \equiv \mathbb{R}^d \setminus \mathrm{int}([f \leq \alpha_2 + \tau]). \end{cases} \tag{3.3}$$

Lemma 3.5. *The moving set $\widehat{\mathcal{S}}$ (resp. $\widehat{\mathcal{U}}$) is Lipschitz continuous provided (H1) holds (resp., (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

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

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,

$$\begin{aligned} d_H(\widehat{\mathcal{S}}(t),\widehat{\mathcal{S}}(s)) &= \sup_{x\in\widehat{\mathcal{S}}(s)} d(x,\widehat{\mathcal{S}}(t)) \\ &= \sup_{x\in\widehat{\mathcal{S}}(s)} d(x,[f\le \alpha_1 - t]) \\ &\le \sup_{x\in\widehat{\mathcal{S}}(s)} \frac{1}{\ell}(f(x)-\alpha_1+t)^+ = \frac{1}{\ell}|t-s|. \end{aligned}$$

Thus, $\widehat{\mathcal{S}}$ is a $\frac{1}{\ell}$-Lipschitz set-valued map. Now, fix $x \in \widehat{\mathcal{U}}(-t)\setminus\widehat{\mathcal{U}}(-s) = \widehat{\mathcal{U}}(-t)\cap\widehat{\mathcal{S}}(s)$ and let $\hat{n}$ the exterior unit vector of $\widehat{\mathcal{S}}(t)$ at $\mathrm{proj}_{\widehat{\mathcal{S}}(t)}(x)$. Then, since f is coercive, there exists $\zeta > 0$ such that $\mathrm{proj}_{\widehat{\mathcal{S}}(t)}(x)+\zeta\hat{n} \in \mathrm{bd}\,\widehat{\mathcal{S}}(s) = \mathrm{bd}\,\widehat{\mathcal{U}}(-s)$. Clearly

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

With this in mind, we can write

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

Thus, $\widehat{\mathcal{U}}$ 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 sets $\widehat{\mathcal{S}} : [0,T] \rightrightarrows \mathbb{R}^d$ and $\widehat{\mathcal{U}} : [-T,0] \rightrightarrows \mathbb{R}^d$, that is,

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

We now set

$$M = \mathrm{bd}([f\le\alpha_2]) \qquad \text{and} \qquad \mathcal{R} = [f\le\alpha_2]\setminus \mathrm{int}([f\le\alpha_1]),$$

and consider the mapping

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

where $u(\cdot,m)$ is the unique solution of the first sweeping process differential inclusion of (3.4) with initial condition $x_0 = m$.

Proposition 3.6 (Inversion of the sweeping flow)**.** *Let $m \in M \bigcap \mathrm{int}(\mathrm{dom}\, f)$. Under (H1)–(H3), the mapping $u : [0,T]\times M \to \mathcal{R}$ is one-to-one and bi-Lipschitz near m.*

Proof. Let us first show that u is (globally) Lipschitz. Let us denote by K the (common) Lipschitz constant of the moving sets $\widehat{\mathcal{S}}(\cdot)$ and $\widehat{\mathcal{U}}(\cdot)$, given by Lemma 3.5 and let (t_1,m_1) and (t_2,m_2) in $[0,T]\times M$. Then

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

The last inequality stems from the fact that the solutions of a Lipschitz sweeping process is Lipschitz with the same constant (see [21]) and that for two different solutions, the convex values of $\widehat{\mathcal{S}}$ yield that the distance between these is decreasing with respect to the time (see [16]).

Let further $\delta > 0$ be such that $B(m, \delta) \subset \operatorname{int}(\operatorname{dom} f)$ and set $\Gamma = M \cap B(m, \delta)$. Since f is continuous on its domain, we deduce that $\Gamma = [f = \alpha_2] \cap B(m, \delta)$. Moreover, by continuity of u, we get that there exists $\varepsilon > 0$ such that

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

Evoking Proposition 2.1 we deduce that for every $(t, m) \in (0, \varepsilon] \times \Gamma$, $f(u(t, m)) = \alpha_2 - t$, that is, u takes values in $\mathcal{R}$.

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

$$\begin{cases} \dot{v}(t) \in -N(\widehat{\mathcal{U}}(t); v(t)) \qquad t \in [-\bar{t}, 0], \\ v(-\bar{t}) = \bar{u}. \end{cases} \tag{3.5}$$

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

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

Thus, $(t_1, m_1) = (t_2, m_2)$, 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

$$\begin{cases} \dot{v}_i(t) \in -N(v_i(t), \widehat{\mathcal{U}}(t)), \qquad t \in [-t_2, 0], \\ v_i(-t_2) = \bar{u}_i, \end{cases} \qquad \text{for } i = 1, 2.$$

Thanks to Lemma 3.5 (see, e.g., [21]) and Proposition 2.1, the above differential inclusions have unique solutions, $v_1, v_2 : [-t_2, 0] \to \mathbb{R}^d$. It is not hard to realize 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 us consider $\mathcal{P}_k := \{s_0, \dots, s_k\}$ a uniform partition of $[-t_2, 0]$ and for $i \in \{1, 2\}$, let $v_{i,k}$ be the polygonal curve emanating from u_i associated to $\mathcal{P}_k$ defined by

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

Since for some $r > 0$ the sets $\widehat{\mathcal{U}}(t)$ are r-uniformly prox-regular for all $t \in [-t_2, 0]$, taking k sufficiently large such that the width of the partition $|\mathcal{P}_k|$ is less than $\frac{1}{r}$, the curves $v_{1,k}(\cdot)$ and $v_{2,k}(\cdot)$ are both well-defined. Furthermore, for each $j \in \{1, \dots, k\}$

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

Since $\mathrm{proj}(\cdot, \widehat{\mathcal{U}}(s_j))$ is Lipschitz with constant $(1-\kappa)^{-1}$ over the set $\widehat{\mathcal{U}}(s_j) + K|sj - s_{j-1}|\mathbb{B}_d$ (see, e.g., [6]) we deduce:

$$\begin{aligned}
\|v_{1,k}(s_j) - v_{2,k}(s_j)\| &\le \left(1 - r^{-1} \cdot K|s_j - s_{j-1}|\right)^{-1} \|v_{1,k}(s_{j-1}) - v_{2,k}(s_{j-1})\| \\
&\le \left(1 - \frac{r^{-1} \cdot KT}{k}\right)^{-1} \|v_{1,k}(s_{j-1}) - v_{2,k}(s_{j-1})\|,
\end{aligned}$$

Thus,

$$\|v_{1,k}(0) - v_{2,k}(0)\| \le \left(1 - \frac{r^{-1} \cdot KT}{k}\right)^{-k} \|u_1 - u_2\|.$$

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

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

Note that

$$|t_1 - t_2| = |f(u_1) - f(u_2)| \le L\|u_1 - u_2\|,$$

where L is the Lipschitz constant of f in $\mathcal{R}$. We conclude that

$$\|(t_1, m_1) - (t_2, m_2)\| \le \left(L + e^{r^{-1}KT}\right) \|u(t_1, m_1) - u(t_2, m_2)\|.$$

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

The above proposition shows that the mapping $u : [0, T] \times M \to \mathcal{R}$ locally induces a foliation of the annulus $\mathcal{R} = [\alpha_1 \le f \le \alpha_2]$, near every point $m \in M \cap \mathrm{int}(\mathrm{dom}\, f)$. The problem appears at points in $\mathrm{bd}(\mathrm{dom}\, f)$ that belong to the boundary of more than one sublevel set, where the mapping u loses injectivity. However, if $\mathrm{dom}\, f = \mathbb{R}^d$, the neighborhood $[0, \varepsilon] \times \Gamma$ can be taken to be $[0, T] \times M$ and $u : [0, T] \times M \to \mathcal{R}$ induces a (complete) foliation of the whole annulus $\mathcal{R}$. This is illustrated in Figure 3.

For every $m \in M \bigcap \mathrm{int}(\mathrm{dom}\, f)$ the set $\Gamma = M \cap B(m, \delta)$ is a $\mathcal{C}^{1,1}$-submanifold of codimension 1 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 $\mathcal{R} = S(0) \setminus \mathrm{int}(S(T))$ with the usual Lebesgue measure $\mathcal{L}_d$ of $\mathbb{R}^d$.

Theorem 3.7 (control of null sets)**.** *Let $\mathcal{N} \subset \mathbb{R}^d$ be a null measure set, and assume (H1)–(H3) hold. Let $\bar{m} \in M \cap \mathrm{int}(\mathrm{dom}\, f)$ and let $\delta > 0$ such that $\Gamma = M \bigcap B(\bar{m}, \delta) \subset \mathrm{int}(\mathrm{dom}\, f)$. Then, there exist $\varepsilon > 0$ and a subset $A \subset \Gamma$ of full measure (i.e. $\mathcal{H}^{d-1}(A) = \mathcal{H}^{d-1}(\Gamma)$) such that for every $m \in A$*

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

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

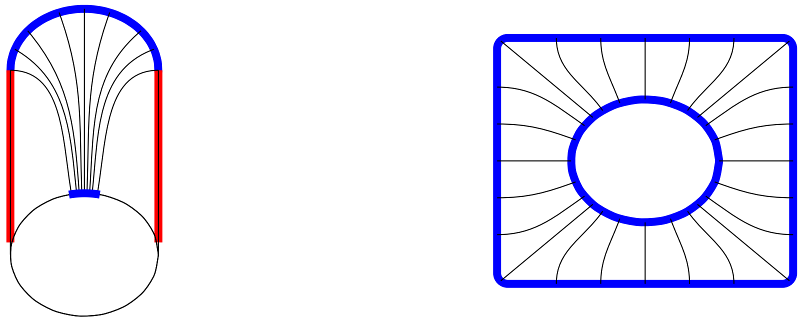

Figure 3: Illustration of foliation induced by $u : [0,T] \times M \to \mathcal{R}$. *Left*: In blue, the part of M, corresponding to $M \bigcap \mathrm{int}(\mathrm{dom}\, f)$, that is injectively transported; In red, the part of M, corresponding to $M \bigcap \mathrm{bd}(\mathrm{dom}\, f)$, where u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

Proof. Let $\varepsilon > 0$ such that $[0,\varepsilon] \times \Gamma$ is the neighborhood that appears in the proof of Proposition 3.6. 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 $u : [0,\varepsilon] \times \Gamma \to \mathcal{R}$ is measurable, and therefore so it 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., [12, 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 $Ju(t,\gamma) = |\det(Du(t,\gamma))|$ is well-defined μ–a.e. in $[0,\varepsilon] \times \Gamma$. Thus, we can apply the co-area formula (see, e.g., [12, Theorem 3.10]) to write

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

where the last equality comes from the fact that u is a bijection between $[0,\varepsilon] \times \Gamma$ and $\mathcal{O}$ and consequently, $\mathcal{H}^0 \left(u^{-1}(\mathcal{N}) \bigcap u^{-1}(x) \right) = \mathbb{1}_{\mathcal{N}}(x)$, 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,

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

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

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

The proof is complete. $\square$

Combining Proposition 3.6 with Theorem 3.7, we will show that the mapping $u : [0,T] \times M \to \mathcal{R}$ induces steepest descent curves almost everywhere. Indeed, for the spacial case $\mathrm{dom}\, f = \mathbb{R}^d$ 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 3.7) 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 $t \in \_m$ 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 $t \in [0,T]$, which implies that $u(\cdot, m)$ is a steepest descent curve. The general case, stated in the next theorem, follows the same idea, with a localization argument.

Theorem 3.8 (Existence of steepest descent curves)**.** *Let $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be a quasiconvex, locally Lipschitz and coercive function, satisfying hypotheses (H1), (H2) and (H3). Then, for almost all $x \in \operatorname{int}(\operatorname{dom} f)$, f admits a steepest descent curve emanating from x.*

Proof. Let $\widetilde{\mathcal{N}}$ be the set of all $x \in \operatorname{int}(\operatorname{dom} f)$ for which f does not admit a steepest descent curve emanating from x and let $\mathcal{N}$ be the set of all $x \in \operatorname{int}(\operatorname{dom} f)$ for which f is not differentiable at x. By Radamacher's theorem $\mathcal{L}_d(\mathcal{N}) = 0$. We shall show that the set $\widetilde{\mathcal{N}}$ is also null. To this end, fix $\alpha \in (\min f, \sup f)$ and set

$$\begin{aligned} &K_{\alpha,n} = [f \leq \alpha] \setminus \operatorname{int}([f \leq \min f + \tfrac{1}{n}]); \text{ and} \\ &\widetilde{\mathcal{N}}_{\alpha,n} = \widetilde{\mathcal{N}} \cap K_{\alpha,n} \end{aligned}$$

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

$$\int_{-\infty}^{\infty} \mathcal{H}^{d-1}(\widetilde{\mathcal{N}}_{\alpha,n} \cap h_{\alpha,n}^{-1}(t))dt = \int_{\widetilde{\mathcal{N}}_{\alpha,n}} Jh_{\alpha,n}(x)dx = \int_{\widetilde{\mathcal{N}}_{\alpha,n}} Jf(x)dx.$$

Since $K_{\alpha,n}$ is compact and does not contain $\operatorname{argmin} f$, hypothesis (H2) ensures that there exists $c \in \mathbb{R}$ such that $Jf(x) \geq c$ for almost all $x \in K_{\alpha,n}$. Thus,

$$\begin{aligned} \mathcal{L}_d(\widetilde{\mathcal{N}}_{\alpha,n}) \ &\leq \ \frac{1}{c} \int_{-\infty}^{\infty} \mathcal{H}^{d-1}(\widetilde{\mathcal{N}}_{\alpha,n} \cap h_{\alpha,n}^{-1}(t))\, dt \\ &= \ \frac{1}{c} \int_{\min f + \frac{1}{n}}^{\alpha} \mathcal{H}^{d-1}(\widetilde{\mathcal{N}}_{\alpha,n} \cap f^{-1}(t))\, dt. \end{aligned}$$

Let us fix $t \in [\min f + 1/n, \alpha]$, set $M = f^{-1}(t)$ and choose $m \in f^{-1}(t) \bigcap \operatorname{int}(\operatorname{dom} f)$. Let $[0,\varepsilon] \times \Gamma_m$ and $A_m \subset \Gamma_m$ be the neighborhood of m and the full measure subset of Γ_m given by Theorem 3.7 for 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 $u(t,\gamma)$. Then, using Proposition 2.1 we can apply chain rule todeduce 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 $\widetilde{\mathcal{N}}_{\alpha,n} \cap A_m = \emptyset$. Since $f^{-1}(t) \cap \operatorname{int}(\operatorname{dom} f)$ is σ-compact, it can be covered by countably many sets $\{\Gamma_{m_k} \ : \ k \in \mathbb{N}\}$, yielding

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

Since the latter conclusion holds for every $t \in [\min f + 1/n, \alpha]$, we deduce that $\mathcal{L}_d(\widetilde{\mathcal{N}}_{\alpha,n}) = 0$. Taking $n \to \infty$ and $\alpha \nearrow \sup f$, we deduce that

$$\mathcal{L}_d(\widetilde{\mathcal{N}} \setminus \operatorname{argmin} f) = 0.$$

The assertion follows observing that, by definition of steepest descent curves, the set $\widetilde{\mathcal{N}} \bigcap \operatorname{argmin} f$ is empty. $\square$

4 A regularization scheme for quasiconvex functions

In order to apply the results of Section 3, we present a regularization scheme based on the max-convolution operator (see, e.g., [20]). For two functions $f, g : \mathbb{R}^d \to \overline{\mathbb{R}}$ 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 consider a quasiconvex function $f \in \mathcal{Q}$ satisfying $\inf f = 0$. 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

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

The main property we are going to use for f_ε is that $[f_\varepsilon \leq \alpha] = [f \leq \alpha] + \varepsilon\mathbb{B}$, for every $\alpha \in \mathbb{R}$. If $\inf f > -\infty$, we can easily adapt the definition of f_ε by considering $f_\varepsilon = \inf f + (f - \inf f) \diamond I_{\varepsilon\mathbb{B}}$, which still preserves the formulae of the sublevel sets. 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. *Let $f \in \mathcal{Q}$ and $\varepsilon > 0$. We define the regularized function $f_\varepsilon : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ given by*

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

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

In what follows, we focus our attention on functions $f \in \mathcal{Q}$ with $\min f = 0$. The general case will be treated later on in 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 $f \in \mathcal{Q}$ with $\min f = 0$, 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* $\operatorname{bd}[f_\varepsilon \leq \alpha]$ *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) f_ε *is locally Lipschitz-continuous on its domain* $\operatorname{dom} f_\varepsilon = \operatorname{dom} f + \varepsilon\mathbb{B}$.

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

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

Let us now choose $x_0 \in \operatorname{int}(\operatorname{dom} f)$ such that $\overline{|\nabla f|}(x_0) > \ell$ (recall definition in (2.5)). Pick $\delta > 0$ sufficiently small such that

$$|\nabla f|(y) > \ell, \text{ for every } y \in \bar{B}(x_0, \delta) \subset \operatorname{dom} f \tag{4.4}$$

and consider the function

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

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

Lemma 4.3. *If* (4.4) *holds, the limiting slope* $\overline{|\nabla h|}$ *is strictly positive on* $\operatorname{dom} h \setminus \operatorname{argmin} h$.

Proof. Set $C := \bar{B}(x_0, \delta)$. Choose $z \in \operatorname{dom} h \setminus \operatorname{argmin} h$ 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,

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

Thus, applying [13, 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

$$\begin{aligned} |\nabla h|(z) &= \limsup_{\beta \nearrow \alpha} \frac{\alpha - \beta}{d(z, [h \leq \beta])} \\ &\geq \limsup_{\beta \nearrow \alpha} \frac{\delta - d(x_0, [f \leq \beta])}{\delta + d(x_0, [f \leq \beta])} \frac{\alpha - \beta}{d(z, [f \leq \beta])} \\ &= \frac{\delta - d(x_0, [f \leq \alpha])}{\delta + d(x_0, [f \leq \alpha])} |\nabla f|(z) \geq \underbrace{\frac{\delta - d(x_0, [f \leq \alpha])}{\delta + d(x_0, [f \leq \alpha])} \ell}_{:= \phi(\alpha)}. \end{aligned}$$

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

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

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

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 (iv).
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-lower critical points, in the sense of (4.5) below. Let us set:

$$\mathcal{U}_\varepsilon := \{x \in \operatorname{dom} f_\varepsilon : \overline{|\nabla f|}(\operatorname{proj}(x; [f \leq f_\varepsilon(x)])) > 0\} \tag{4.5}$$

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 $x \in \operatorname{dom} f_\varepsilon$.*

Proof. Let $x \in \operatorname{dom} f_\varepsilon$ satisfying (4.5), and let $z = \operatorname{proj}(x, [f \leq f_\varepsilon(x)])$. Then, there exist $\delta, \ell > 0$ such that for all $z' \in B(z, 2\delta) \cap \operatorname{dom} f$, $|\nabla f|(z') > \ell$.

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 [3, Example 4.1] and Lemma 3.5, we deduce that the mapping $y \mapsto \operatorname{proj}(y, [h \leq g_\varepsilon(y)])$ is continuous over the set $[g_\varepsilon > \min h]$. This yields that there exists $\eta > 0$ such that

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

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

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

Thus, g_ε and h_ε coincide in $B(x, \eta)$. Now, applying Lemma 4.4 and Theorem 3.8, we deduce that h_ε, and so g_ε admits steepest descent curves emanating from almost every point in

$$B(x, \eta) \bigcap \operatorname{int}(\operatorname{dom} h_\varepsilon) = B(x, \eta) \bigcap \operatorname{int}(\operatorname{dom} g_\varepsilon).$$

The proof is complete. □

(*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 [11]: 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. 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.

Acknowledgements. The first author acknowledges support from the Austrian Science Fund (FWF, P-36344-N). The second author was partially funded by ANID-Chile, through the grant Fondecyt 11220586, and FB210005 BASAL funds for centers of excellence.

References

[1] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. *Gradient flows in metric spaces and in the space of probability measures*. Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, second edition, 2008.

[2] Hedy Attouch, Giuseppe Buttazzo, and Gérard Michaille. *Variational analysis in Sobolev and BV spaces*, volume 17 of *MOS-SIAM Series on Optimization*. Society for Industrial and Applied Mathematics (SIAM); Mathematical Optimization Society, Philadelphia, PA, second edition, 2014.

[3] Hédy Attouch and Roger J.-B. Wets. Quantitative stability of variational systems. II. A framework for nonlinear conditioning. *SIAM J. Optim.*, 3(2):359–381, 1993.

[4] D. Aussel. New developments in quasiconvex optimization. In *Fixed point theory, variational analysis, and optimization*, pages 171–205. CRC Press, Boca Raton, FL, 2014.

[5] Dominique Azé and Jean-Noël Corvellec. Characterizations of error bounds for lower semicontinuous functions on metric spaces. *ESAIM Control Optim. Calc. Var.*, 10(3):409–425, 2004.

[6] Giovanni Colombo and Lionel Thibault. Prox-regular sets and applications. In *Handbook of nonconvex analysis and applications*, pages 99–182. Int. Press, Somerville, MA, 2010.

[7] Marc-Olivier Czarnecki and Lionel Thibault. Sublevel representations of epi-Lipschitz sets and other properties. *Math. Program.*, 168(1-2):555–569, 2018.

[8] A. Daniilidis, D. Drusvyatskiy, and A. S. Lewis. Orbits of geometric descent. *Canad. Math. Bull.*, 58(1):44–50, 2015.

[9] Aris Daniilidis, Olivier Ley, and Stéphane Sabourau. Asymptotic behaviour of self-contracted planar curves and gradient orbits of convex functions. *J. Math. Pures Appl. (9)*, 94(2):183–199, 2010.

[10] Ennio De Giorgi, Antonio Marino, and Mario Tosques. 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.

[11] D. Drusvyatskiy, A. D. Ioffe, and A. S. Lewis. Curves of descent. *SIAM J. Control Optim.*, 53(1):114–138, 2015.

[12] Lawrence C. Evans and Ronald F. Gariepy. *Measure theory and fine properties of functions.* Textbooks in Mathematics. CRC Press, Boca Raton, FL, revised edition, 2015.

[13] Armin Hoffmann. The distance to the intersection of two convex sets expressed by the distances to each of them. *Math. Nachr.*, 157:81–98, 1992.

[14] Alexander D. Ioffe. *Variational analysis of regular mappings.* Springer Monographs in Mathematics. Springer, Cham, 2017.

[15] Marco Longinetti, Paolo Manselli, and Adriana Venturi. On steepest descent curves for quasi convex families in $\mathbb{R}^n$. *Math. Nachr.*, 288(4):420–442, 2015.

[16] Jean-Jacques Moreau. Evolution problem associated with a moving convex set in a Hilbert space. *J. Differential Equations*, 26(3):347–374, 1977.

[17] Pedro Pérez-Aros and Emilio Vilches. An enhanced Baillon-Haddad theorem for convex functions defined on convex sets. *Appl. Math. Optim.*, 83(3):2241–2252, 2021.

[18] Juan Peypouquet and Sylvain Sorin. Evolution equations for maximal monotone operators: asymptotic analysis in continuous and discrete time. *J. Convex Anal.*, 17(3-4):1113–1163, 2010.

[19] R. Tyrrell Rockafellar and Roger J.-B. Wets. *Variational analysis*, volume 317 of *Grundlehren der mathematischen Wissenschaften.* Springer-Verlag, Berlin, 1998.

[20] A. Seeger and M. Volle. On a convolution operation obtained by adding level sets: classical and new results. *RAIRO Rech. Opér.*, 29(2):131–154, 1995.

[21] Lionel Thibault. Sweeping process with regular and nonregular sets. *J. Differential Equations*, 193(1):1–26, 2003.

[22] Lionel Thibault. *Unilateral Variational Analysis in Banach Spaces: (In 2 Parts)Part I: General TheoryPart II: Special Classes of Functions and Sets.* World Scientific Publishing Co., Inc., River Edge, NJ, 2023.

[23] Cédric Villani. *Optimal transport*, volume 338 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences].* Springer-Verlag, Berlin, 2009.

Aris DANIILIDIS

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

E-mail: aris.daniilidis@tuwien.ac.at
https://www.arisdaniilidis.at/

Research supported by the Austrian Science Fund grant FWF P-36344N.

David SALAS

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

E-mail: david.salas@uoh.cl
http://davidsalasvidela.cl

Research supported by the grants:

CMM FB210005 BASAL funds for centers of excellence (ANID-Chile)
FONDECYT 11220586 (Chile)