Propagation of singularities of Moreau envelopes and distance functions in a Hilbert space

Thomas Strömberg[1]

Abstract

The Moreau envelopes of nonconvex functions on a real Hilbert space have points of nondifferentiability called singularities. For the first time in an infinite-dimensional space, the propagation of singularities along intrinsic characteristic curves $\boldsymbol{x}(t)$ is studied. The properties of the intrinsic characteristics are investigated in some detail. It is proved that singularities propagate along intrinsic characteristics not only locally but globally in t. Actually, this paper distinguishes between "singularities" and "strong singularities," two concepts that agree for distance functions but disagree for Moreau envelopes, in general. Along intrinsic characteristics, propagating singularities instantaneously transform into strong singularities. Concerning the distance function to a closed nonempty set E, it is demonstrated that the singular set and $\complement E$ are homotopy equivalent provided E satisfies a certain condition which is weaker than the boundedness of $\complement E$. The notion of intrinsic characteristics was introduced by Cannarsa and Cheng (Calc Var Part Differ Equ 56, 2017).

1 Introduction

This paper examines in a real Hilbert space H the motion of the singularities of the *Moreau envelopes* [35, 36] of a lower semicontinuous function $f : H \to \mathbb{R} \cup \{\infty\}$, i.e.,

$$f_t(x) = \inf_{y \in H} \left(f(y) + \frac{1}{2t} \|x - y\|^2 \right), \quad t > 0, \ x \in H. \tag{1}$$

✉ Thomas Strömberg
thomas.stromberg@ltu.se

1 Department of Engineering Sciences and Mathematics, Luleå University of Technology, 971 87 Luleå, Sweden

In the absence of convexity of f, we prove that singularities, in a sense made precise below, propagate along intrinsic characteristic curves, a notion introduced for Hamilton–Jacobi equations in a finite-dimensional context by Cannarsa and Cheng in [13]. The accompanying set of minimizers

$$P_{tf}(x) = \arg\min_{y \in H} \left(tf(y) + \frac{1}{2}\|x - y\|^2 \right)$$

defines the *proximal mapping* $P_{tf} \colon H \rightrightarrows H$ of index t. Assuming

$$\inf_{x \in H} \left(f(x) + \alpha\|x\|^2 \right) \in \mathbb{R} \quad \text{for all} \quad \alpha > 0, \tag{2}$$

f_t is a semiconcave (see Lemma 2 below) real-valued function for all $t \in (0, \infty)$ such that $\lim_{t \downarrow 0} f_t(x) = f(x)$ for all $x \in H$. The distance function d_E appears when f is the indicator function $\mathfrak{I}_E$ of a closed nonempty subset E of H (i.e., $\mathfrak{I}_E(x) = 0$ if $x \in E$ while $\mathfrak{I}_E(x) = \infty$ otherwise). Indeed, then

$$f_t(x) = \frac{1}{2t}d_E^2(x) \quad \text{and} \quad P_{tf} = P_E,$$

where the distance function and the associated metric projection to E send each $x \in H$ to

$$d_E(x) = \inf_{y \in E} \|x - y\| \quad \text{and}$$
$$P_E(x) = \arg\min_{y \in E} \|x - y\| = \{y \in E : \|x - y\| = d_E(x)\},$$

respectively.

In general, f_t is nondifferentiable; in fact, f_t is everywhere Fréchet differentiable in H when $0 < t < T$ if and only if $f + (2T)^{-1}\|\cdot\|^2$ is a convex function. (The definitions of Fréchet and Gâteaux differentiability are recalled in Section 2.) In the literature, the functions (1) are also known as the Moreau–Yosida approximations of f in particular in the presence of convexity. For general functions f, an application of Asplund's [7] results on generic differentiability of convex functions shows that $(t, x) \mapsto f_t(x)$ is Fréchet differentiable on a dense G_δ. The connection to the Hamilton–Jacobi equation

$$\frac{\partial S}{\partial t} + \frac{1}{2}\|d_x S\|^2 = 0 \quad \text{in } (0, \infty) \times H, \tag{3}$$
$$\lim_{t \downarrow 0} S(t, x) = f(x) \quad \text{in } H, \tag{4}$$

is classical and well understood in particular when $H = \mathbb{R}^n$. The Moreau envelope $S(t, x) = f_t(x)$ is a viscosity solution of the Cauchy problem (3)–(4) also in infinite dimensions [37, 38]; see Proposition 7.

If H is finite-dimensional, then the infimum (1) is always attained and the following conditions are equivalent:

(i) f_t is Fréchet differentiable at x;

(ii) f_t is Gâteaux differentiable at x;

(iii) $P_{tf}(x)$ is a singleton.

Adopting a term from Hamilton–Jacobi theory, a point $(t, x) \in (0, \infty) \times \mathbb{R}^n$ is called *singular* or a *singularity* if f_t fails to be differentiable at x. We conclude that (t, x) is a singularity if and only if there exists more than one minimizer y in (1).

The situation is less clear-cut and, in fact, intricate when dim $H = \infty$. First, the infimum (1) need not be attained. Still, it is true that (ii) $\Leftarrow$ (i) $\Rightarrow$ (iii) but no other implication holds true between any other pair of these conditions. We choose to distinguish between two kinds of singularities one of which is stronger than the other when dim $H = \infty$.

Definition 1 A point (t, x) is called *singular* or a *singularity* for $S(t, x) = f_t(x)$ if f_t fails to be Fréchet differentiable at x. The set of all singular points in $(0, \infty) \times H$ is denoted by Σ.

Definition 2 A point (t, x) is called *strongly singular* or a *strong singularity* for $S(t, x) = f_t(x)$ if either f_t fails to be Gâteaux differentiable at x or $P_{tf}(x) = \emptyset$. The set of all strongly singular points in $(0, \infty) \times H$ is designated by Σ_s.

On the one hand, Σ and Σ_s are nonempty unless f is a convex function and on the other $\Sigma_s \subseteq \Sigma$; see Proposition 6. Theorem 10 and Example 3 shed light on subtleties regarding Σ and Σ_s for Moreau envelopes f_t showing in particular that Σ and Σ_s disagree in general. When E is nonconvex and $f = \mathfrak{I}_E$, $f_t(\cdot) = d_E^2(\cdot)/(2t)$ is not a globally differentiable function. Clearly, the question of differentiability of f_t is in this case independent of t. With a slight abuse of notation, we therefore understand that

$$\Sigma = \{x \in H : d_E^2 \text{ is not Fréchet differentiable at } x\}$$

and

$$\Sigma_s = \{x \in H : P_E(x) = \emptyset \text{ or } d_E^2 \text{ is not Gâteaux differentiable at } x\}.$$

Needless to say, Σ and E are disjoint sets. It follows from Fitzpatrick's paper [23] that actually $\Sigma = \Sigma_s$ in this case; see Theorem 11. The following fundamental result on singular dynamics for distance functions in a Hilbert space has been an inspiration for this work.

Theorem 1 (Frerking and Westphal [24]) *Suppose that $x_0 \in \Sigma$. Then, unless it is an isolated singularity, x_0 lies on a nonconstant Lipschitz arc each point of which is a member of Σ. Furthermore, x_0 is an isolated singularity if and only if $P_E(x_0) = \mathbb{S}(x_0, d_E(x_0))$, the sphere about x_0 of radius $d_E(x_0)$.*

There is a substantial, highly successful and fast expanding literature on singular dynamics for Hamilton–Jacobi equations set in $\mathbb{R}^n$ covering (3) when dim $H < \infty$; see, e.g., the survey articles [14, 15] by Cannarsa and Cheng. At the heart of this branch

of research there are two main concepts of characteristics as well as some refinements. It has been established when dim $H < \infty$ that singularities propagate along Lipschitz continuous arcs $X(t)$ that are characterized as generalized characteristics. This notion was first defined and studied in $\mathbb{R}^n$ by Albano and Cannarsa in [4], although it agrees for $n = 1$ with Dafermos' concept of generalized characteristics for scalar conservation laws in one space variable [22]. For the Cauchy problem (3)–(4), assuming $H = \mathbb{R}^n$ and denoting by $d^+ f_t$ the Fréchet superdifferential of f_t (see (7) for the definition), for any $(t_0, x_0) \in (0, \infty) \times \mathbb{R}^n$ there exists a unique locally Lipschitz continuous arc $X(t)$ such that

$$X(t_0) = x_0 \quad \text{and} \quad \dot{X}(t) \in d^+ f_t(X(t)) \quad \text{a.e. } t \in [t_0, \infty). \tag{5}$$

It satisfies $(t, X(t)) \in \Sigma$ for all $t \in [t_1, \infty)$ if $t_1 \geq t_0$ and $(t_1, X(t_1)) \in \Sigma$. The fact that the singular propagation continuous without interruption was established by Cannarsa, Mazzola and Sinestrari in [19] (see Theorem 2 below). For a viscosity solution $S(t, x)$ of a general Hamilton–Jacobi equation $\partial S / \partial t + \mathcal{H}(t, x, \nabla S) = 0$ in, say, $(0, \infty) \times \mathbb{R}^n$, a generalized characteristic refers to a locally Lipschitz continuous curve $X(t)$ satisfying

$$\dot{X}(t) \in \text{co} \, \nabla_p \mathcal{H}(t, X(t), d_x^+ S(t, X(t))) \qquad \text{for a.e. } t.$$

It is well-known that the singularities of $S(t, x)$ propagate along generalized characteristics locally in time t; consult, e.g., the seminal paper [4] or the subsequent articles [17, 43] by Cannarsa and Yu or the monograph [16] on semiconcavity by Cannarsa and Sinestrari. In certain special cases research has revealed that the propagation continues for all later times without stopping, see, e.g., [19] and Albano's paper [3]. However, the extent to which the propagation is global in time t remains to this day a vital and central research problem for general Hamilton–Jacobi equations [14, 15]. Nevertheless, Albano actually proved in [2] for a broad class of Hamilton–Jacobi equations that $(t, X(t))$ remains in the closed hull of the singular set (i.e., in the C^1 singular support) for all later times. This is an important achievement especially in the presence of some degree of smoothness. In certain nonsmooth cases, however, there might be a considerable difference between Σ and $\overline{\Sigma}$, inasmuch as int $\Sigma = \emptyset$ whereas $\overline{\Sigma}$ may contain interior points. Santilli [39] has in his investigations of distance functions to C^1 or $C^{1,1}$ hypersurfaces in $\mathbb{R}^n$ obtained several striking results on the denseness of the singular set.

   While the propagation of singularities along arcs may be viewed as a kind of a lower bound on the structure and connectedness of the singular set, rectifiability results constitute upper bounds. The fine estimates obtained by Alberti, Ambrosia and Cannarsa in [6] for general convex functions on $\mathbb{R}^N$ apply to Σ. Furthermore, significant results on the rectifiability of $\overline{\Sigma}$ including sharp bounds for the Hausdorff measure of $\overline{\Sigma} \backslash \Sigma$ for smooth initial-value problems for Hamilton–Jacobi equations appear in the article [18] by Cannarsa, Mennucchi and Sinestrari. A recent paper by Miura and Tanaka [34] on distance functions of closed sets states that Σ is not only covered by but is equal to an at most countable union of Lipschitz hypersurfaces save an exceptional set of codimension two.

Singular dynamics has also been studied for general semiconcave functions on Banach spaces as notably in the paper [1] by Albano and Cannarsa.

In this paper we take a different path and study the motion of singularities in an infinite-dimensional setting by means of so-called intrinsic characteristics. This class of relevant curves was singled out by Cannarsa and Cheng in [13] and was used for topological studies in [20, 21]. It is not based on a differential inclusion but has a purely variational definition.

Definition 3 (Intrinsic characteristic) For any $x_0 \in H$ and any $t_0 > 0$ we define the *intrinsic characteristic* $\boldsymbol{x}(t)$ for $t \in [t_0, \infty)$ by $\boldsymbol{x}(t_0) = x_0$ and

$$\{\boldsymbol{x}(t)\} = \arg\max_{x \in H} \left(f_t(x) - \frac{1}{2(t - t_0)} \|x_0 - x\|^2 \right), \qquad t_0 < t < \infty.$$

Equivalently,

$$\frac{\boldsymbol{x}(t) - x_0}{t - t_0} \in d^+ f_t(\boldsymbol{x}(t)), \qquad t_0 < t < \infty. \tag{6}$$

In (5) and (6), $f_t(x)$ serves as a nonsmooth velocity potential, for momentaneous and average velocities, respectively, and the superdifferential $d^+ f_t$ is defined by (7). The results of this paper are not stated in the introduction except for the main propagation theorem (which reappears as Theorem 4 below). This backbone asserts that singularities, once they come into existence, propagate along intrinsic characteristics instantly transforming into strong singularities.

Theorem. For a given lower semicontinuous function $f : H \to \mathbb{R} \cup \{\infty\}$ meeting (2), let $\boldsymbol{x}(t)$ be the intrinsic characteristic emanating from (t_0, x_0).

(i) If $t_1 \geq t_0$ and $(t_1, \boldsymbol{x}(t_1)) \in \Sigma$, then $(t, \boldsymbol{x}(t)) \in \Sigma_s$ for every $t \in (t_1, \infty)$.
(ii) If $(t_0, x_0) \in \Sigma$ and $0 \notin d^+ f_{t_0}(x_0)$, then $\dot{\boldsymbol{x}}^+(t_0) \neq 0$ and $(t, \boldsymbol{x}(t)) \in \Sigma_s$ for all $t \in (t_0, \infty)$.

Outline of the paper. Section 2 contains some background material mainly on relevant concepts of differentiation, including sub- and superdifferentials. The definition of Asplund's function and some of its basic properties are given in Sect. 3. A major previous result on singular propagation for Moreau envelopes in $\mathbb{R}^n$ is briefly recalled in Sect. 4 where also some remarks are made about the problem of extensions to infinite dimensions. In Sects. 5 through 8 we state *without proofs* our most important principal results, namely Theorems 3 through 9, on singular dynamics along intrinsic characteristics. Sections 5–6 state the basic results for Moreau envelopes and distance functions. The generation of singularities for distance functions and an application to homotopy equivalence is the subject of Sect. 7. Section 8 examines weak limit points of bounded singular arcs $\boldsymbol{x}(t)$ as $t \to \infty$ showing that every such limit point is singular too. This stability property entails that the singular propagation along intrinsic characteristics has a truly global character.

After that point the paper is devoted to a detailed study of Moreau envelopes and intrinsic characteristics *including proofs.* Sects. 9–10 investigate the differentiability

properties of $S(t, x) = f_t(x)$ partly from the perspective of viscosity solution theory. Sections 11–13 define and analyze intrinsic characteristics for Moreau envelopes and distance functions with a focus on singular dynamics. Section 14 adds a further tool for prolonging the singular propagation. Section 15 compiles and completes the postponed proofs of the main results contained in Sects. 5–8. Finally, some examples are gathered in Section 16.

2 Prerequisites

Throughout the paper H stands for a real Hilbert space whose scalar product and norm are denoted by $\langle \cdot, \cdot \rangle$ and $\| \cdot \|$, respectively. The open and closed balls with center at x and of radius R are signified by $\mathbb{B}(x, R)$ and $\overline{\mathbb{B}}(x, R)$, respectively, while the sphere is denoted by $\mathbb{S}(x, R)$. Let $g \colon H \to (-\infty, \infty]$ be a proper function which means that its essential domain $\operatorname{dom} g = \{x \in H : g(x) < \infty\}$ is nonvoid. The Fréchet and Gâteaux differential at a point $x \in \operatorname{dom} g$ are denoted by $dg(x)$ and $\nabla g(x)$, respectively. We recall the definitions. The function g is a said to be *Fréchet differentiable* at $x \in \operatorname{dom} g$ if there exists a vector $dg(x) \in H$ such that

$$g(x + h) - g(x) = \langle dg(x), h \rangle + o(h) \quad \text{as } h \to 0,$$

where as usual $o(h)/\|h\| \to 0$ as $h \to 0$. The function is *Gâteaux differentiable* at $x \in \operatorname{dom} g$ if the directional derivative

$$g'(x, v) = \lim_{\lambda \to 0} \frac{g(x + \lambda v) - g(x)}{\lambda}$$

exists in any direction $v \in H$, and for a certain $\nabla g(x) \in H$ it holds that $g'(x, v) = \langle \nabla g(x), v \rangle$ for all $v \in H$.

The *Fréchet superdifferential* of g is the multivalued mapping $d^+ g \colon H \rightrightarrows H$ defined by

$$d^+ g(x) = \left\{ p \in H : \limsup_{\|h\| \to 0} \frac{g(x + h) - g(x) - \langle h, p \rangle}{\|h\|} \leq 0 \right\} \tag{7}$$

when $x \in \operatorname{dom} g$ while $d^+ g(x) = \emptyset$ if $g(x) = \infty$. The *Fréchet subdifferential* $d^- f$ is defined by replacing "lim sup" and "≤ 0" in (7) by "lim inf" and "≥ 0", respectively. The sets $d^\pm g(x)$ are convex and closed and they are simultaneously nonempty exactly when g is Fréchet differentiable at x, in which case $d^+ g(x) = d^- g(x) = \{dg(x)\}$. Kruger's survey paper [31] furnishes an overview of generalized differentiation.

The *Legendre–Fenchel transform* $g^* \colon H \to (-\infty, \infty]$ of g is the convex and lower semicontinuous function which assigns to each $y \in H$ the value

$$g^*(y) = \sup_{x \in H} (\langle x, y \rangle - g(x)).$$

If g is itself convex and lower semicontinuous, then $d^- g$ agrees with the Fenchel subdifferential ∂g which is defined by

$$\partial g(x) = \{y \in H : \forall z \in H \ g(z) \geq g(x) + \langle z - x, y \rangle\} \quad \text{for every } x \in \text{dom } g,$$

while $\partial g(x) = \emptyset$ otherwise. In this case, $g^{**} = g$ and $\partial g^* = (\partial g)^{-1}$, i.e., $y \in \partial g(x)$ $\Leftrightarrow x \in \partial g^*(y)$.

Let g be continuous and locally semiconcave in an open nonempty subset Ω of H, i.e., for any point $x \in \Omega$ let there exist a ball $B \subseteq \Omega$ centered at x and a constant $c \geq 0$ such that $g - c\| \cdot \|^2/2$ is concave in B. Then, for any $x \in \Omega$, $d^+ g(x)$ is a bounded convex closed nonempty subset of H, which reduces to a singleton if and only if g is Gâteaux differentiable at x; and $d^- g(x)$ is empty unless g is Fréchet differentiable at x. A *reachable gradient* $p \in H$ of g at x, in symbols $p \in d^\bullet g(x)$, is by definition a *weak* limit of a sequence of Fréchet gradients $dg(x_k)$ where $x_k \to x$ strongly; $d^\bullet g(x)$ is a nonvoid subset of $d^+ g(x)$ and $d^+ g(x) = \overline{co} d^\bullet g(x)$. While Fréchet differentiability is a stronger notion than Gâteaux differentiability in general, it is a property of semiconcave functions in finite dimensions that $dg(x)$ exists if and only if $\nabla g(x)$ exists. The main object of this paper, namely $S(t, x) = f_t(x)$, is locally semiconcave in $(0, \infty) \times H$, an open subset of the Hilbert space $\mathbb{R} \times H$.

We next recall Asplund's characterization of Fréchet differentiability of g^* for general proper functions g. We denote by Γ the set of all convex, lower semicontinuous functions $\gamma : [0, \infty) \to [0, \infty]$ such that $\gamma(0) = 0$ and consider the subsets

$$\Gamma_U = \{\gamma \in \Gamma : \gamma(r) > 0 \text{ if } r > 0\} \quad \text{and} \quad \Gamma_L = \{\gamma \in \Gamma : \gamma(r)/r \to 0 \text{ as } r \to 0\}.$$

In defining the conjugate to $\gamma \in \Gamma$ by

$$\gamma^*(s) = \sup_{r \geq 0}(rs - \gamma(r)), \quad 0 \leq s < \infty,$$

$\gamma \in \Gamma_U$ if and only if $\gamma^* \in \Gamma_L$; see Lemma 1 in [7].

Lemma 1 (Asplund [7]) *Let g be a lower semicontinuous proper function on H and consider a pair $(x_0, y_0) \in H \times H$. Then the following conditions are mutually equivalent.*

(i) *g^* is finite and Fréchet differentiable at y_0 with $x_0 = dg^*(y_0)$.*
(ii) *For some $\gamma^* \in \Gamma_L$,*

$$g^*(y) \leq g^*(y_0) + \langle x_0, y - y_0 \rangle + \gamma^*(\|y - y_0\|) \quad \text{for all } y \in H,$$

and $g^*(y_0) \in \mathbb{R}$.
(iii) *For some $\gamma \in \Gamma_U$,*

$$g(x) \geq g(x_0) + \langle x - x_0, y_0 \rangle + \gamma(\|x - x_0\|) \quad \text{for all } x \in H,$$

and $g(x_0) \in \mathbb{R}$.

(iv) g^* *is finite at y_0 and* dom g^* *is radial at y_0; and if*

$$\lim_{j \to \infty} (\langle x_j, y_0 \rangle - g(x_j)) = g^*(y_0),$$

then $x_j \to x_0$ in norm.

Any single one of these four conditions implies $\langle x_0, y_0 \rangle = g(x_0) + g^*(y_0)$ *and* $g(x_0) = g^{**}(x_0)$.

3 Asplund's function

The following is a version of Asplund's function [8].

Lemma 2 *The associated function* $A \colon (0, \infty) \times H \to \mathbb{R}$ *defined by*

$$A(t, x) = \frac{1}{2} \|x\|^2 - t S(t, x), \qquad (t, x) \in (0, \infty) \times H,$$

is convex. In terms of the Legendre–Fenchel transform, $A(t, x) = \left(tf + \frac{1}{2} \| \cdot \|^2 \right)^* (x)$.

Proof Indeed, A can be represented as the pointwise supremum of a family of affine functions of (t, x) in the following way:

$$A(t, x) = \sup_{y \in \text{dom } f} (\langle x, y \rangle - tf(y) - \|y\|^2 / 2).$$

$\square$

In the case of the distance function to a set E, A is independent of t, namely, $A(x) = \frac{1}{2} \|x\|^2 - \frac{1}{2} d_E^2(x)$. In his paper on Chebyshev sets, Asplund [8] made significant use of the fact that A is a convex continuous function such that $P_E \subseteq \partial A$ (see Proposition 6). We give a brief review of the relation between P_E and ∂A. The equality $P_E(x) = \partial A(x)$ holds if and only if $x \notin \Sigma$, in which case $P_E(x) = \{dA(x)\}$. If $\dim H < \infty$, then co $P_E(x) = \partial A(x)$. However, $\overline{co} P_E(x)$ and $\partial A(x)$ do not coincide in general if $\dim H = \infty$. Example 5 (Example 4, respectively) furnishes a case where $P_E(x) = \emptyset$ while $\partial A(x)$ is a singleton ($P_E(x)$ is a singleton while $\partial A(x)$ has more than one element, respectively). In fact, Godini [25] has in any infinite-dimensional Hilbert space H uncovered a proximinal set E (i.e., $P_E(x) \neq \emptyset$ for all $x \in H$) such that $\overline{co} P_E(x)$ is a strict subset of $\partial A(x)$ for some $x \in H \setminus E$. Klee [30] obtained a few years earlier an example of this nature for nonseparable Hilbert spaces.

Lemma 3 *(Berens [10] and Veselý [42]) At every $x \in H$,*

$$\partial A(x) = \bigcap_{R > d_E(x)} \overline{co} \left(E \cap \overline{\mathbb{B}}(x, R) \right) \subseteq \overline{\mathbb{B}}(x, d_E(x)) \cap \overline{co} E$$

and $P_E(x) = \partial A(x) \cap \mathbb{S}(x, d_E(x))$.

We collect a few elementary observations concerning the extreme case when $P_E(x) = \mathbb{S}(x, d_E(x))$ which occurs exactly when $E \cap \overline{\mathbb{B}}(x, d_E(x)) = \mathbb{S}(x, d_E(x))$. A proof can easily be constructed on the basis of Lemma 3.

Lemma 4 *The following conditions are equivalent for any $x \in H$:*

(i) $P_E(x) = \mathbb{S}(x, d_E(x))$;
(ii) $\partial A(x) = \overline{\mathbb{B}}(x, d_E(x))$;
(iii) *The boundary of $\partial A(x)$ is equal to $\mathbb{S}(x, d_E(x))$;*
(iv) *The boundary of $\partial A(x)$ is included in $P_E(x)$.*

The following properties of A^* were elegantly exploited in [7].

Lemma 5 $A^* = (\mathfrak{I}_E + \frac{1}{2}\|\cdot\|^2)^{**}$ *is the supremum of all closed convex functions minorizing $\frac{1}{2}\|\cdot\|^2$ on E. In particular, $A^*(x) = \infty$ and hence $\partial A^*(x) = \emptyset$ for any x in the complement of $\overline{co}E$, i.e.,*

$$\operatorname{dom} \partial A^* \subseteq \operatorname{dom} A^* \subseteq \overline{co}E.$$

4 Remarks on extension from finite to infinite dimensions

The time global propagation results for generalized characteristics $X(t)$ (defined by (5)) obtained in [3, 19, 41] are directly applicable to Moreau envelopes in $\mathbb{R}^n$. Regardless of dimension, as confirmed in Proposition 7 below, singularities of $S(t, x) = f_t(x)$ can be detected from the following dichotomy:

$$\min\{\omega + \|v\|^2/2 \colon (\omega, v) \in d^+ S(t, x)\} \begin{cases} < 0 & \text{if } (t, x) \in \Sigma, \\ = 0 & \text{if } (t, x) \notin \Sigma. \end{cases}$$

We present a version of the global-in-time propagation result in $\mathbb{R}^n$ obtained by Cannarsa, Mazzola and Sinestrari [19]. Albano's paper [3] examines more general equations successfully.

Theorem 2 *Assume that $f \colon \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is a lower semicontinuous function satisfying (2) and set $S(t, x) = f_t(x)$. Consider the generalized characteristic $X \colon [t_0, \infty) \to \mathbb{R}^n$ emanating from the point $(t_0, x_0) \in (0, \infty) \times \mathbb{R}^n$. Then*

$$m(t) := t^2 \cdot \min\{\omega + \|v\|^2/2 \colon (\omega, v) \in d^+ S(t, X(t))\}$$

is a right-continuous, nonincreasing and nonpositive function of $t \in [t_0, \infty)$ such that $m(t) < 0$ exactly if $(t, X(t)) \in \Sigma$. In particular, if $t_1 \geq t_0$ and $(t_1, X(t_1)) \in \Sigma$, then $(t, X(t)) \in \Sigma$ for all $t \in [t_1, \infty)$.

Proof We sketch the proof of the monotonicity presented in [41] leaving out most details. The task is to prove that $m(t) \leq m(t_1)$ for any fixed $t_0 \leq t_1$ and any $t > t_1$. First, $A(t, x)$ is approximated by C^∞ convex functions $A_\varepsilon(t, x)$ by means of integral

convolution with a nonnegative mollifier and then $S_\varepsilon(t, x)$ is defined from $A_\varepsilon(t, x) = \frac{1}{2}\|x\|^2 - t S_\varepsilon(t, x)$. Consider the solution $X_\varepsilon(t)$ to $dx/dt = \nabla_x S_\varepsilon(t, x)$ satisfying $X_\varepsilon(t_1) = X(t_1)$. Then $X_\varepsilon(t) \to X(t)$ locally uniformly as $\varepsilon \downarrow 0$. A calculation reveals that the derivative of

$$m_\varepsilon(t) := t^2 \left(\frac{\partial S_\varepsilon}{\partial t}(t, X_\varepsilon(t)) + \frac{1}{2}\|\nabla_x S_\varepsilon(t, X_\varepsilon(t))\|^2 \right)$$

satisfies

$$\dot{m}_\varepsilon(t) = -t \left\langle D^2 A_\varepsilon(Y_\varepsilon(t)) \dot{Y}_\varepsilon(t), \dot{Y}_\varepsilon(t) \right\rangle \le 0.$$

Here, $D^2 A_\varepsilon(t, x)$ denotes the (positive semidefinite) full Hessian matrix of $A_\varepsilon(t, x)$ and $Y_\varepsilon(t) = (t, X_\varepsilon(t))$. We find that $m_\varepsilon(t) \le m_\varepsilon(t_1)$ for every $t > t_1$. By virtue of the *specific* mollification lemma of [17] the mollifier can be chosen so as to obtain $m_\varepsilon(t_1) \to m(t_1)$ as $\varepsilon \downarrow 0$ making it possible to conclude that $m(t) \le m(t_1)$. $\qquad\square$

As this proof sketch shows, the approximation lemma of [17, 43] is a vital tool for the analysis in $\mathbb{R}^n$. It is a difficulty when attempting to derive results in infinite-dimensional spaces that this tailor-made regularization technique is no longer available. For distance functions, the article [5] proves the indefinite propagation of singularities in $\mathbb{R}^n$ as well as in manifolds. The proofs revolve around relevant ordinary differential equations interpreted in a clever way. The major hurdle in extending propagation results to H is the lack of compactness of bounded sets in H or the lack of weak lower semicontinuity of relevant functions. The choice of this paper is the intrinsic characteristics approach of Cannarsa and Cheng [13] which works well owing to the semiconcavity of f_t (as expressed in Lemma 2) yielding well-behaved global concave maximization problems in H. For instance, this is true of the very definition of the arcs (see Definition 3). The analysis relies on weak convergence and the weak lower semicontinuity that convex closed functions enjoy.

5 Principal results for Moreau envelopes

We proceed by presenting a selection of our most central results about intrinsic characteristics introduced in Definition 3. We define $(\omega^\circ(t, x), v^\circ(t, x))$ as the unique element of $d^+ S(t, x)$ minimizing $\omega + \frac{1}{2}\|v\|^2$, i.e.,

$$\omega^\circ(t, x) + \frac{1}{2}\|v^\circ(t, x)\|^2 \le \omega + \frac{1}{2}\|v\|^2 \quad \text{for all } (\omega, v) \in d^+ S(t, x).$$

We observe that $v^\circ(t, x) \in d^+ f_t(x)$.

Theorem 3 *Assume that $f: H \to \mathbb{R} \cup \{\infty\}$ is a lower semicontinuous function satisfying (2). Let $\boldsymbol{x}(t)$ signify the intrinsic characteristic emanating from a point $x_0 \in H$ at time $t_0 > 0$. Then the following assertions are fulfilled:*

(i) $\mathbf{x}(t) = dF^{t_0 \to t}(x_0)$ *when* $t \geq t_0$ *where*

$$F^{t_0 \to t} = \left(\frac{t - t_0}{t} A(t, \cdot) + \frac{t_0}{2t} \| \cdot \|^2 \right)^*$$

or, setting $J^{t_0 \to t} = dF^{t_0 \to t}$, $\mathbf{x}(t) = J^{t_0 \to t}(x_0)$ *where*

$$J^{t_0 \to t} = \left(\frac{t - t_0}{t} \partial_x A(t, \cdot) + \frac{t_0}{t} I \right)^{-1} = \left(I - (t - t_0)d^+ f_t \right)^{-1}.$$

The convex function $F^{t_0 \to t} \colon H \to \mathbb{R}$ is Fréchet differentiable and $J^{t_0 \to t} = dF^{t_0 \to t}$ is globally Lipschitz continuous of rate t/t_0.

(ii) $t \mapsto \mathbf{x}(t)$ *is Hölder continuous with exponent* $1/2$ *in* $[t_0, T]$ *for every* $t_0 < T < \infty$.

(iii) The right derivative $\dot{\mathbf{x}}^+(t_0)$ exists and $\dot{\mathbf{x}}^+(t_0) = v^\circ(t_0, x_0)$; in particular, $\dot{\mathbf{x}}^+(t_0) \neq 0$ if $0 \notin d^+ f_{t_0}(x_0)$.

Corollary 1 (Lipschitz and monotone dependence on initial data) *Let* $t_0 > 0$ *and consider the intrinsic characteristics* $\mathbf{x}(t)$ *and* $\mathbf{y}(t)$ *issuing from* (t_0, x_0) *and* (t_0, y_0), *respectively. Then*

$$\frac{1}{t} \| \mathbf{x}(t) - \mathbf{y}(t) \| \leq \frac{1}{t_0} \| x_0 - y_0 \| \quad and \quad \langle \mathbf{x}(t) - \mathbf{y}(t), x_0 - y_0 \rangle \geq 0$$

for all $t \in [t_0, \infty)$.

The following is our principal result on singular dynamics for Moreau envelopes.

Theorem 4 *(Propagation of singularities) For a given lower semicontinuous function* $f \colon H \to \mathbb{R} \cup \{\infty\}$ *meeting* (2)*, let* $\mathbf{x}(t)$ *be the intrinsic characteristic emanating from* (t_0, x_0).

(i) If $t_1 \geq t_0$ and $(t_1, \mathbf{x}(t_1)) \in \Sigma$, then $(t, \mathbf{x}(t)) \in \Sigma_s$ for every $t \in (t_1, \infty)$.

(ii) If $(t_0, x_0) \in \Sigma$ and $0 \notin d^+ f_{t_0}(x_0)$, then $\dot{\mathbf{x}}^+(t_0) = v^\circ(t_0, x_0) \neq 0$ and $(t, \mathbf{x}(t)) \in \Sigma_s$ for all $t \in (t_0, \infty)$.

Remark 1 Example 6 manifests that generalized characteristics and intrinsic characteristics are different concepts and that, in general, $J^{t_1 \to t_2} \circ J^{t_0 \to t_1} \neq J^{t_0 \to t_2}$ when $t_0 < t_1 < t_2$.

Remark 2 As an alternative to $\mathbf{x}(t) = J^{t_0 \to t}(x_0)$ we could construct a singular curve by choosing a sequence $0 < t_0 < t_1 < \cdots < t_k \to \infty$, setting $\boldsymbol{\xi}(t_0) = x_0$ and proceeding recursively by defining

$$\boldsymbol{\xi}(t) = J^{t_k \to t}(\boldsymbol{\xi}(t_k)) \quad \text{for } t \in (t_k, t_{k+1}], \ k = 0, 1, 2, \ldots.$$

Cf. Example 6.

6 Basic results for distance functions

We turn to the spreading of singularities of distance functions. We remind the reader that $\Sigma = \Sigma_s$ in this case (Theorem 11). For any $x_0 \in H \setminus E$ and any $t_0 > 0$, the intrinsic characteristic $\boldsymbol{x}(t)$ satisfying the initial condition $\boldsymbol{x}(t_0) = x_0$ is in this case given by

$$\{\boldsymbol{x}(t)\} = \arg\max_{x \in H} \left(\frac{1}{2t} d_E^2(x) - \frac{1}{2(t-t_0)} \|x_0 - x\|^2 \right), \qquad t_0 < t < \infty. \qquad (8)$$

Equivalently,

$$\boldsymbol{x}(t) = J^{t_0 \to t}(x_0) \quad \text{where} \quad J^{t_0 \to t} = \left(\frac{t-t_0}{t} \partial A + \frac{t_0}{t} I \right)^{-1}, \qquad t_0 \le t < \infty, \quad (9)$$

where Asplund's function is $A(x) = \frac{1}{2} \|x\|^2 - \frac{1}{2} d_E^2(x)$ [8]. We remark that the choice of $t_0 > 0$ is arbitrary because $J^{\mu t_0 \to \mu t} = J^{t_0 \to t}$ for any $\mu > 0$.

Definition 4 A point $x_0 \in H$ is termed *critical* if $0 \in d^+ d_E^2(x_0)$.

A trivial consequence of (6) is that $\boldsymbol{x}(t) = x_0$ for all $t \in [t_0, \infty)$ if x_0 is a critical point.

Proposition 1 *In terms of Asplund's function* $A(x) = \frac{1}{2} \|x\|^2 - \frac{1}{2} d_E^2(x)$, x_0 *is a critical point if and only $x_0 \in \partial A(x_0)$ if and only if $x_0 \in \partial A^*(x_0)$.*

Proposition 2 *Let E be a closed nonempty subset of H. If x_0 is a critical point, then $x_0 \in \text{dom } \partial A^* \subseteq \overline{co}E$. In particular, there exists no critical point outside of $\overline{co}E$.*

Proof Let x_0 be a critical point. Then $x_0 \in \partial A^*(x_0)$ and hence $x_0 \in \text{dom } \partial A^* \subseteq \overline{co}E$ by Lemma 5. The conclusion also follows from Lemma 3 applied to $x_0 \in \partial A(x_0)$. □

We let $\boldsymbol{r}(x)$ denote the norm minimal element of $d^+ d_E^2(x)/2$. As $d^+ d_E^2/2 = d_E d^+ d_E$ and d_E is Lipschitz continuous with constant 1, we have $\|\boldsymbol{r}(x)\| \le d_E(x)$. If $x \notin \Sigma$, then $\boldsymbol{r}(x) = x - y$ where y is the closest point to x in E, thus, $\|\boldsymbol{r}(x)\| = d_E(x)$. By contrast, if $x \in \Sigma$, then $\|\boldsymbol{r}(x)\| < d_E(x)$ as $d^+ d_E(x)$ is a convex closed subset of $\overline{\overline{\mathbb{B}}}(0, 1)$ with more than one element. On account of Theorem 3(iii), $\dot{\boldsymbol{x}}^+(t_0) = v^\circ(t_0, x_0) = \boldsymbol{r}(x_0)/t_0$.

The crude Hölder estimate of Theorem 3(ii) can be radically improved for distance functions.

Theorem 5 (Lipschitz continuity) *Let E be a closed nonempty subset of H. For any $t_0 < s < t < \infty$ consider the initial velocity $v_0 = \boldsymbol{r}(x_0)/t_0$ and the average velocities*

$$v_{t_0 \to s} = \frac{\boldsymbol{x}(s) - x_0}{s - t_0} \quad \text{and} \quad v_{s \to t} = \frac{\boldsymbol{x}(t) - \boldsymbol{x}(s)}{t - s}.$$

Then

$$\|v_{s\to t}\|^2 \le \|v_{t_0\to s}\|^2 - \|v_{s\to t} - v_{t_0\to s}\|^2; \tag{10}$$

in particular, $\|v_{s\to t}\| < \|v_{t_0\to s}\|$ unless $v_{s\to t} = v_{t_0\to s}$. Furthermore,

$$\|v_{t_0\to t}\|^2 \le \|v_0\|^2 - \|v_{t_0\to t} - v_0\|^2 \tag{11}$$

as well as

$$\|v_{s\to t}\|^2 \le \|v_0\|^2 - \|v_{t_0\to s} - v_0\|^2 - \|v_{s\to t} - v_{t_0\to s}\|^2; \tag{12}$$

in particular, $\|v_{s\to t}\| < \|v_0\|$ unless $v_{s\to t} = v_{t_0\to s} = v_0$.

Theorem 6 *The intrinsic characteristic emanating from a noncritical point x_0 possesses the following properties.*

(i) *$x(t) \ne x_0$ for all $t \in (t_0, \infty)$ and the initial velocity $\dot{x}^+(t_0)$ is given by*

$$\dot{x}^+(t_0) = v_0 := v^\circ(t_0, x_0) = r(x_0)/t_0 \ne 0.$$

   In particular, $\|\dot{x}^+(t_0)\| = \|v_0\| \in (0, d_E(x_0)/t_0]$.
(ii) *$\|x(t) - x(s)\| \le \|v_0\|(t - s)$ for all $t_0 \le s < t < \infty$.*
(iii) *The distances $d_E(x(t))$ and $\|x_0 - x(t)\|$ are nondecreasing functions of $t \in [t_0, \infty)$. In fact, if $t_0 \le t_1 < t_2$, then*

$$d_E(x(t_1)) < d_E(x(t_2)) \quad and \quad \|x_0 - x(t_1)\| < \|x_0 - x(t_2)\|$$

   unless $x(t_1) = x(t_2)$. Furthermore, the functions $d_E(x(t))$ and $\|x_0 - x(t)\|$ are either simultaneously bounded or simultaneously unbounded.
(iv) *(Propagation of singularities) If $x_0 \in \Sigma$, then $0 < \|\dot{x}^+(t_0)\| < d_E(x_0)/t_0$ and $x(t) \in \Sigma\backslash\{x_0\}$ for all $t \in (t_0, \infty)$.*

Similar results were derived for generalized characteristics of distance functions in Riemannian manifolds in the article [5] by Albano, Cannarsa, Nguyen and Sinestrari.

7 Generation of singularities for distance functions and homotopy equivalence

If $P_E(x_0)$ is a singleton $\{y_0\}$, then we can extend $x(\cdot)$ to $[0, \infty)$ by setting,

$$\text{for } t \in [0, t_0), \quad x(t) = y_0 + tv_0 \quad \text{where} \quad v_0 = \frac{x_0 - y_0}{t_0}. \tag{13}$$

Our next theorem asserts that if $x(t)$ starts from a nonsingular point $x_0 \in \complement E$, whose nearest point in E is y_0, then $x(t)$ will eventually become singular unless E is supported at the boundary point y_0 by the hyperplane whose normal vector is $r(x_0) = x_0 - y_0$. (Since $x_0 \notin \Sigma$, $P_E(x_0) = \{dA(x_0)\} = \{y_0\}$.)

Theorem 7 (Generation and propagation of singularities) *Let E be a closed nonempty subset of H and suppose that $x_0 \in \complement(\Sigma \cup E)$. Let $\boldsymbol{x}(\cdot)$ be extended to $[0, \infty)$ by (13) where $P_E(x_0) = \{y_0\}$. Then, either*

(i) *$\boldsymbol{x}(t)$ remains nonsingular for all $t \in (0, \infty)$, $P_E(\boldsymbol{x}(t)) = \{y_0\}$ and $\boldsymbol{x}(t) = y_0 + tv_0 = x_0 + (t - t_0)v_0$ for all $t \in [0, \infty)$, or*
(ii) *there exists a $t^* \in [t_0, \infty)$ such that $\boldsymbol{x}(t)$ is nonsingular and $\boldsymbol{x}(t) = x_0 + (t - t_0)v_0$ for all $t \in (0, t^*)$ whereas $\boldsymbol{x}(t)$ is singular for all $t \in (t^*, \infty)$.*

Case (i) happens if and only if

$$E \subseteq \{x \in H : \langle x - y_0, v_0 \rangle \leq 0\}.$$

In case (ii), the cut point $\boldsymbol{x}(t^*)$ lies on the ray from y_0 through x_0, $y_0 \in P_E(\boldsymbol{x}(t^*))$ while $y_0 \notin P_E(\boldsymbol{x}(t))$ for every $t > t^*$.

Theorem 7 has topological implications. While $\complement E$ and Σ are clearly not homeomorphic (as $\complement E$ is an open set while the interior of Σ is empty), it was proved by Lieutier in [33] that they still are homotopy equivalent if $\complement E$ is a bounded nonvoid subset of $\mathbb{R}^n$. Lieutier's theorem was extended to Riemannian manifolds in [5]. In our Hilbert space setting we shall construct the required homotopy drawing on Theorem 7, assuming a weaker condition than the boundedness of $\complement E$, namely that $\rho(\complement E) < \infty$ where

$$\rho(\complement E) := \sup_{x \in \complement E} d_E(x) = \sup_{x \in H} d_E(x). \tag{14}$$

Equivalently,

$$\rho(\complement E) = \sup\{R > 0 : \mathbb{B}(x, R) \subseteq \complement E \text{ for some } x \in \complement E\}.$$

The first step states that $\boldsymbol{x}(t)$ will be trapped in Σ for all times t such that $t/t_0 > \rho(\complement E)/d_E(x_0)$ provided (14) is finite.

Corollary 2 *Assume that the complement of E is open, nonempty and such that $\rho(\complement E) < \infty$. Let $x_0 \in \complement(\Sigma \cup E)$. Then alternative (ii) in Theorem 7 is in force. Moreover, $\boldsymbol{x}(t^*) = y_0 + t^*v_0$ and $t^*/t_0 \leq \rho(\complement E)/d_E(x_0)$.*

Proof Alternative (i) is ruled out by $d_E(\boldsymbol{x}(t)) \leq \rho(\complement E) < \infty$. We have $\boldsymbol{x}(t) = y_0 + tv_0$ when $t \in [t_0, t^*]$ and

$$\rho(\complement E) \geq d_E(\boldsymbol{x}(t^*)) = \|\boldsymbol{x}(t^*) - y_0\| = t^*\|v_0\| = t^* d_E(x_0)/t_0.$$

It ensues that $t^*/t_0 \leq \rho(\complement E)/d_E(x_0)$. □

The second step summarizes what we know about the Lipschitz continuity of $J^{t_0 \to t}(x_0) = J^{1 \to t/t_0}(x_0)$. A direct combination of Theorem 5 and Corollary 1 yields the following conclusion.

Proposition 3 *The mapping* $\Phi \colon [1, \infty) \times H \to H$ *defined by* $\Phi(\tau, x) = J^{1 \to \tau}(x)$ *is Lipschitz continuous:*

$$\|\Phi(\tau_1, x) - \Phi(\tau_0, x)\| \le \|r(x)\| |\tau_1 - \tau_0| \le d_E(x) |\tau_1 - \tau_0| \tag{15}$$

as well as

$$\|\Phi(\tau, x_1) - \Phi(\tau, x_0)\| \le \tau \|x_1 - x_0\| \tag{16}$$

for all τ *and* τ_j *in* $[1, \infty)$ *and all* x *and* x_j *in* H.

We next generalize Lieutier's result to the Hilbert space H. Since $\Sigma \subset \complement E$, the proof amounts to exhibiting a continuous mapping $F \colon [0, 1] \times \complement E \to \complement E$ such that $F(0, \cdot)$ is the identity mapping on $\complement E$ while $F(1, \cdot)$ maps $\complement E$ into Σ and $F(\theta, \cdot)$ maps Σ into Σ for every $\theta \in [0, 1]$.

Theorem 8 (Homotopy equivalence) *Assume that the complement of E is an open nonempty set such that* $\rho(\complement E) < \infty$. *Then* $\complement E$ *and* Σ *are of the same homotopy type because the mapping* $F \colon [0, 1] \times \complement E \to \complement E$ *defined by*

$$F(\theta, x) = J^{1 \to 1 - \theta + \theta R / d_E(x)}(x) \quad \text{for all } (\theta, x) \in [0, 1] \times \complement E,$$

for any fixed $R > \rho(\complement E)$, *is continuous and satisfies, on the one hand,* $F(0, x) = x$ *and* $F(1, x) \in \Sigma$ *for all* $x \in \complement E$ *and, on the other,* $F(\theta, x) \in \Sigma$ *whenever* $(\theta, x) \in [0, 1] \times \Sigma$. *In fact, the homotopy F is Lipschitz continuous away from the boundary of E.*

8 Weak limit points as $t \to \infty$ for distance functions

This section investigates the behavior of $x(t)$ as $t \to \infty$ assuming that the initial point $x_0 \in \complement E$ be noncritical. Being a nondecreasing function, the distance $\|x_0 - x(t)\|$ either approaches ∞ or remains bounded as t tends to ∞. Let us assume the latter alternative, i.e., that $x(t)$ stays in some ball $\overline{\mathbb{B}}(x_0, R)$ for all $t \ge t_0$. Under these circumstances, $x(t) \in \Sigma$ for all sufficiently large t by virtue of Theorem 7. By the weak sequential compactness of closed balls, for any $t_j \to \infty$ the bounded sequence $x(t_j)$ possesses a weakly convergent subsequence. It turns out that each weak limit point belongs to $\partial A^*(x_0)$ (see Lemma 5 for basic properties of this set). Moreover, it holds that $\partial A^*(x_0) \subseteq \Sigma \setminus \{x_0\}$. Thus, each weak limit point is singular. By basic convex function theory, $x_0 \in \text{int dom } A^* \Leftrightarrow A^*$ is finite and continuous at $x_0 \Leftrightarrow A^*$ is Lipschitz continuous on some neighborhood of x_0 (see, e.g., [12, Sect. 4.1]). Either of these conditions implies that $\partial A^*(x_0)$ is bounded and nonempty.

Theorem 9 (Weak limit points) *For the distance function d_E to a closed nonempty set $E \subset H$, let $x(t)$ be the intrinsic characteristic emanating at time $t_0 > 0$ from a noncritical point $x_0 \in \complement E$. Assume that $\|x_0 - x(t)\|$ stays bounded as t tends to ∞.*

Then, the set of weak limit points

$$W = \{\bar{x} \in H : \bar{x} \text{ is the weak limit of } \boldsymbol{x}(t_j) \text{ for some sequence } t_j \to \infty\}$$

is a weakly closed nonempty subset of $\Sigma \setminus \{x_0\}$. *In fact,*

$$\emptyset \neq W \subseteq \partial A^*(x_0) \subseteq \Sigma \setminus \{x_0\}. \tag{17}$$

In particular, $x_0 \in \overline{co}E$.

Corollary 3 *Under the hypotheses of Theorem 9, the following assertions hold true.*

(a) *If* $\partial A^*(x_0)$ *is a bounded set, then W is weakly compact.*
(b) *If* A^* *is Gâteaux differentiable at* x_0, *then* $W = \{\nabla A^*(x_0)\}$ *and* $\boldsymbol{x}(t) \to \nabla A^*(x_0)$ *weakly as* $t \to \infty$.
(c) *If* A^* *is Fréchet differentiable at* x_0, *then* $\boldsymbol{x}(t) \to dA^*(x_0)$ *strongly as* $t \to \infty$.

Corollary 4 *Every intrinsic characteristic* $\boldsymbol{x}(t)$ *with an initial point* $x_0 \notin \overline{co}E$ *is unbounded, i.e.,* $\|x_0 - \boldsymbol{x}(t)\| \to \infty$ *and* $d_E(\boldsymbol{x}(t)) \to \infty$ *as* $t \to \infty$.

Proof By contraposition of Theorem 9, we conclude that $\|x_0 - \boldsymbol{x}(t)\|$ is unbounded when $x_0 \notin \overline{co}E$. Theorem 6 (or Proposition 15) asserts that $d_E(\boldsymbol{x}(t))$ is unbounded too.   □

Corollary 5 *Every singular intrinsic characteristic* $\boldsymbol{x}(t)$ *with an initial point* $x_0 \notin \overline{co}E$ *carries singularities to infinity, i.e., if* $t_1 \geq t_0$ *and* $\boldsymbol{x}(t_1) \in \Sigma$, *then* $\boldsymbol{x}(t) \in \Sigma$ *for all* $t \in [t_1, \infty)$, $\|x_0 - \boldsymbol{x}(t)\| \to \infty$ *and* $d_E(\boldsymbol{x}(t)) \to \infty$ *as* $t \to \infty$.

Proof See Theorem 4 or Corollary 7 for the singular propagation.   □

Cf. Examples 6–7 at the end of the paper. The following remark addresses the important question of the global character of the singular dynamics.

Remark 3 The stability property presented in Theorem 9 can be utilized to exclude that a singular intrinsic characteristic merely is a rescaled or reparameterized purely local singular arc. To explain this, suppose that $\xi : [0, \infty) \to H$ is a continuous nonconstant curve such that, for some $\bar{\tau} > 0$, $\xi(\tau) \in \Sigma$ when $\tau \in [0, \bar{\tau})$ but $\xi(\bar{\tau}) \notin \Sigma$. In this hypothetical situation, $\xi(\cdot)$ propagates singularities locally but not globally in τ as $\xi(\tau)$ exits Σ at $\tau = \bar{\tau}$. Let $\phi : [0, \infty) \to [0, \bar{\tau})$ be an increasing homeomorphism (e.g., $\phi(\sigma) = \bar{\tau}(1 - e^{-\sigma})$ when $\sigma \in [0, \infty)$) and set $\chi = \xi \circ \phi$. Clearly, $\chi(\sigma) \in \Sigma$ for all $\sigma \in [0, \infty)$ yet the propagation is not genuinely global as $\chi(\infty) := \lim_{\sigma \to \infty} \chi(\sigma) = \xi(\bar{\tau}) \notin \Sigma$.

By Theorem 9, if it exists, the strong (or weak) limit $\boldsymbol{x}(\infty)$ is a member of Σ for any singular bounded intrinsic characteristic $\boldsymbol{x}(\cdot)$. Hence, $\boldsymbol{x}(\cdot)$ is not merely a rescaled local singular arc in the above sense. Furthermore, if $\bar{x} = \boldsymbol{x}(\infty)$ is noncritical, then the singular propagation can be genuinely prolonged by using $\bar{x}$ as a new initial point. Indeed, the curve $\boldsymbol{\xi}(s) = J^{s_0 \to s}(\bar{x})$, defined for some fixed $s_0 > 0$ and all $s \geq s_0$, is a nonconstant singular arc such that $d_E(\boldsymbol{\xi}(s))$ is nondecreasing with $d_E(\boldsymbol{\xi}(s)) > d_E(\bar{x}) \geq d_E(\boldsymbol{x}(t))$ for all $t > t_0$ and all $s > s_0$ (see Theorem 6).

Let us take a look at the exceptional case where $x(t)$ moves inside a "spherical cavity" (see Theorem 1 and Lemma 4). This is the only way that an intrinsic characteristic can reach an isolated singularity.

Example 1 Assume that $E \cap \overline{\mathbb{B}}(z_0, R) = \mathbb{S}(z_0, R)$ and select an initial point x_0 in $\mathbb{B}(z_0, R)$, $x_0 \neq z_0$. Then $x(t)$ moves with constant velocity to the center z_0 of the ball, i.e.,

$$x(t) = x_0 + (t - t_0)v_0 \quad \text{when } t \in [t_0, t^*]$$

where

$$t^* = \frac{t_0 R}{R - \|z_0 - x_0\|}, \quad t^* - t_0 = \frac{t_0\|z_0 - x_0\|}{R - \|z_0 - x_0\|} \quad \text{and} \quad v_0 = \frac{z_0 - x_0}{t^* - t_0}.$$

Moreover, $x(t) = z_0$ for all $t \in [t^*, \infty)$ and z_0 is an isolated singularity because $d_E(x) = R - \|z_0 - x\|$ for all $x \in \mathbb{B}(z_0, R)$.

9 Differentiability of Moreau envelopes

This section is concerned with the differentiability properties of the Moreau–Yosida approximations (1) of a lower semicontinuous function f subject to condition (2). Clearly, at a point x, the differentiability of f_t is equivalent to that of $A(t, \cdot)$; see Lemma 2. From now on proofs will be included.

Proposition 4 *Assume that $S(t_0, \cdot) = f_{t_0}(\cdot)$ is Fréchet differentiable at x_0. Then the infimum* (1) *for $t = t_0$ is achieved at a unique point y_0, i.e., $P_{t_0 f}(x_0) = \{y_0\}$ and*

$$d_x S(t_0, x_0) = df_{t_0}(x_0) = \frac{x_0 - y_0}{t_0}.$$

Moreover, at (t_0, x_0), the partial derivative $\partial S/\partial t$ exists too and the Hamilton–Jacobi equation $\partial S/\partial t + \frac{1}{2}\|d_x S\|^2 = 0$ is satisfied.

Proof By Lemma 2, likewise as $S(t, x) = f_t(x)$, the convex function $A: (t, x) \mapsto (tf + \frac{1}{2}\| \cdot \|^2)^*(x)$ is locally Lipschitz continuous. Since f_{t_0} is Fréchet differentiable at x_0, owing to (6), so is $A(t_0, \cdot)$, say, $y_0 = d_x A(t_0, x_0)$. By Lemma 1, therefore,

$$\langle x_0, y_0 \rangle = \left(t_0 f + \frac{1}{2}\| \cdot \|^2\right)^*(x_0) + t_0 f(y_0) + \frac{1}{2}\|y_0\|^2 \tag{18}$$

which converts to

$$t_0 f_{t_0}(x_0) = \frac{1}{2}\|x_0\|^2 - \left(t_0 f + \frac{1}{2}\| \cdot \|^2\right)^*(x_0) = t_0 f(y_0) + \frac{1}{2}\|x_0 - y_0\|^2, \tag{19}$$

demonstrating that the infimum (1) for $t = t_0$ is attained at y_0. In order to investigate the derivative of $f_t(x_0)$ at $t = t_0$ we consider $t \neq t_0$ close to t_0 and select $y_t \in H$ such that

$$f(y_t) + \frac{1}{2t}\|x_0 - y_t\|^2 < f_t(x_0) + (t - t_0)^2, \tag{20}$$

which translates to

$$\langle x_0, y_t \rangle - tf(y_t) - \frac{1}{2}\|y_t\|^2 > \left(tf + \frac{1}{2}\|\cdot\|^2\right)^*(x_0) - t(t - t_0)^2. \tag{21}$$

We claim that $y_t \to y_0$ in norm as $t \to t_0$. On the basis of (21) we find that

$$\liminf_{t \to t_0}\left(\langle x_0, y_t \rangle - t_0 f(y_t) - \frac{1}{2}\|y_t\|^2\right) = \liminf_{t \to t_0}\left(\langle x_0, y_t \rangle - tf(y_t) - \frac{1}{2}\|y_t\|^2\right)$$

$$\geq \liminf_{t \to t_0}\left(tf + \frac{1}{2}\|\cdot\|^2\right)^*(x_0) = \left(t_0 f + \frac{1}{2}\|\cdot\|^2\right)^*(x_0). \tag{22}$$

By virtue of (18), (22) and the Fréchet differentiability of $(t_0 f + \frac{1}{2}\|\cdot\|^2)^*$ at x_0, by invoking Lemma 1, we may now infer that $y_t \to y_0$ in norm, as claimed.

On account of (19) and (20) it holds that

$$f(y_t) + \frac{1}{2t}\|x_0 - y_t\|^2 - (t - t_0)^2 - \left(f(y_t) + \frac{1}{2t_0}\|x_0 - y_t\|^2\right)$$

$$\leq f_t(x_0) - f_{t_0}(x_0)$$

$$\leq f(y_0) + \frac{1}{2t}\|x_0 - y_0\|^2 - \left(f(y_0) + \frac{1}{2t_0}\|x_0 - y_0\|^2\right)$$

which reduces to

$$-\frac{t - t_0}{2tt_0}\|x_0 - y_t\|^2 - (t - t_0)^2 \leq f_t(x_0) - f_{t_0}(x_0) \leq -\frac{t - t_0}{2tt_0}\|x_0 - y_0\|^2;$$

hence,

$$\lim_{t \to t_0}\frac{f_t(x_0) - f_{t_0}(x_0)}{t - t_0} = -\frac{1}{2t_0^2}\|x_0 - y_0\|^2 = -\frac{1}{2}\|df_{t_0}(x_0)\|^2.$$

$\square$

Example 2 Differentiability of $t \mapsto S(t, x_0) = f_t(x_0)$ at $t = t_0$ does not imply that of $x \mapsto S(t_0, x)$ at the point x_0. Indeed, consider $f = -\|\cdot\|$ and $f_t(x) = -t/2 - \|x\|$ at $x_0 = 0$.

Remark 4 The hypothesis of Fréchet differentiability in Proposition 4 cannot be relaxed to Gâteaux differentiability. Indeed, let f be the indicator function of a closed

nonconvex set $E \subset H$ such that d_E^2 is Gâteaux but not Fréchet differentiable at a point $x_0 \in H \setminus E$. Then, by virtue of Theorem 11, $P_E(x_0) = \emptyset$, $\|\nabla d_E(x_0)\| < 1$ and, thus,

$$\frac{\partial S}{\partial t}(t, x_0) + \frac{1}{2}\|\nabla_x S(t, x_0)\|^2 = -\frac{d_E^2(x_0)}{2t^2} + \frac{d_E^2(x_0)}{2t^2}\|\nabla d_E(x_0)\|^2 < 0.$$

A closed subset E of $H = \ell^2$ such that $d_E(0) = 1$ and $\nabla d_E(0) = 0$ is furnished by Example 5. See Corollary 6 below for a more general statement.

Proposition 5 *Suppose that $y_0 \in P_{t_0 f}(x_0)$. Then, at each point of the form $X(t) = y_0 + (t/t_0)(x_0 - y_0)$ where $0 < t < t_0$, f_t is Fréchet differentiable with*

$$df_t(X(t)) = \frac{X(t) - y_0}{t} = \frac{x_0 - y_0}{t_0}.$$

In particular, $P_{tf}(X(t)) = \{y_0\}$ when $t \in (0, t_0)$.

Proof We notice that the supremum for $A(t_0, x_0) = (t_0 f + \frac{1}{2}\|\cdot\|^2)^*(x_0)$ is attained at the same point, y_0, i.e.,

$$\langle x_0, y \rangle - t_0 f(y) - \frac{1}{2}\|y\|^2 \le \langle x_0, y_0 \rangle - t_0 f(y_0) - \frac{1}{2}\|y_0\|^2 \qquad (23)$$

for all $y \in H$. It suffices to verify that $x \mapsto A(t, x)$ is differentiable at $X(t)$ with $d_x A(t, X(t)) = y_0$ for fixed $0 < t < t_0$. To this end, we set

$$\mathcal{D}(y) := tf(y) + \frac{1}{2}\|y\|^2 - tf(y_0) - \frac{1}{2}\|y_0\|^2 - \langle X(t), y - y_0 \rangle.$$

We take (23) into account to derive the inequality

$$\begin{aligned}
\mathcal{D}(y) &= \frac{t}{t_0}(t_0 f(y) - t_0 f(y_0)) + \frac{1}{2}\|y\|^2 - \frac{1}{2}\|y_0\|^2 - \langle X(t), y - y_0 \rangle \\
&\ge \frac{t}{t_0}\left(-\frac{1}{2}\|y\|^2 + \frac{1}{2}\|y_0\|^2 + \langle x_0, y - y_0 \rangle\right) + \frac{1}{2}\|y\|^2 - \frac{1}{2}\|y_0\|^2 - \langle X(t), y - y_0 \rangle \\
&= \frac{t_0 - t}{2t_0}\|y - y_0\|^2,
\end{aligned}$$

which shows that $\mathcal{D}(y) \ge \gamma(\|y - y_0\|)$ for a quadratic function $\gamma \in \Gamma_U$. By appealing to Lemma 1 we may conclude that f_t is Fréchet differentiable at $X(t)$. $\qquad \square$

We recall that, for every $x \in H$, $P_{tf}(x)$ is included in the closed convex nonempty set $\partial_x A(t, x)$ of subgradients. In particular, P_{tf} is a cyclically monotone mapping.

Proposition 6 *Let $f : H \to \mathbb{R} \cup \{\infty\}$ be a lower semicontinuous function satisfying (2). Then the following statements are fulfilled.*

(i) *$\overline{co} P_{tf}(x) \subseteq \partial_x A(t, x)$ for every $t > 0$ and $x \in H$.*

(ii) *A point (t, x) is not strongly singular if and only if f_t is Gâteaux differentiable at x and $P_{tf}(x) \neq \emptyset$.*
(iii) *$\Sigma_s \subseteq \Sigma$ when $\dim H = \infty$ while $\Sigma_s = \Sigma$ when $\dim H < \infty$.*
(iv) *If (t, x) is not a strong singularity, then $P_{tf}(x)$ is a singleton $\{y\}$, $\nabla f_t(x) = (x - y)/t$ and $P_{tf}(x) = \partial_x A(t, x) = \{\nabla_x A(t, x)\}$.*

Proof (i) Let $y \in P_{tf}(x)$; then $f(y) < \infty$ and

$$
\begin{aligned}
& f_t(x + h) - f_t(x) \\
& \leq f(y) + \frac{1}{2t}\|x + h - y\|^2 - \left(f(y) + \frac{1}{2t}\|x - y\|^2 \right) \\
& = \left\langle \frac{x - y}{t}, h \right\rangle + \frac{\|h\|^2}{2t}.
\end{aligned}
$$

It ensues that $(x - y)/t \in d^+ f_t(x)$ which translates to $y \in \partial_x A(t, x)$. Part (ii) is a trivial consequence of the definition.

(iii)–(iv) The Gâteaux differential $\nabla f_t(x)$, if it exists and if $y \in P_{tf}(x)$, must be equal to $(x - y)/t$ whence $P_{tf}(x) = \{y\} = \partial_x A(t, x)$. By Proposition 4, if $(t, x) \notin \Sigma$ then the infimum defining $f_t(x)$ is uniquely attained and so $(t, x) \notin \Sigma_s$. □

10 Viscosity solution

We next confirm that $S(t, x) = f_t(x)$ is a viscosity solution of (3) based on $d^\pm S$.

Proposition 7 *The following assertions are true for the Moreau envelopes of any lower semicontinuous function $f : H \to \mathbb{R} \cup \{\infty\}$ subject to (2).*

(i) *$S(t, x) = f_t(x)$ is a viscosity solution of (3), i.e., for any $(t, x) \in (0, \infty) \times H$, $\omega + \frac{1}{2}\|v\|^2 \leq 0$ if $(\omega, v) \in d^+ S(t, x)$ while $\omega + \frac{1}{2}\|v\|^2 \geq 0$ if $(\omega, v) \in d^- S(t, x)$.*
(ii) *A point $(t, x) \in (0, \infty) \times H$ belongs to Σ if and only if $\omega + \frac{1}{2}\|v\|^2 < 0$ for some $(\omega, v) \in d^+ S(t, x)$.*

Proof (i) By the convexity of $(\omega, v) \mapsto \omega + \frac{1}{2}\|v\|^2$ and $d^+ S(t, x) = \overline{co}d^\bullet S(t, x)$, it suffices to demonstrate the subsolution inequality assuming that $(\omega, v) \in d^\bullet S(t, x)$. In this case, there exists a sequence $\Sigma \not\ni (t_k, x_k) \to (t, x)$ such that $dS(t_k, x_k)$ converges weakly to (ω, p). Proposition 4 ensures that the Hamilton–Jacobi equation is satisfied at each (t_k, x_k); hence, $\omega + \frac{1}{2}\|v\|^2 \leq 0$ is obtained in the limit as $k \to \infty$. Second, if $(\omega, v) \in d^- S(t, x)$, then S is Fréchet differentiable at (t, x) and the Hamilton–Jacobi equation $\omega + \frac{1}{2}\|v\|^2 = 0$ is satisfied.

(ii) Fixing the point (t, x), the task amounts to showing the implication (a) $\Rightarrow$ (b) where

(a) $\omega + \frac{1}{2}\|v\|^2 = 0$ for all $(\omega, v) \in d^+ S(t, x)$;
(b) S is Fréchet differentiable at (t, x).

It is elementary to verify, for any two points $(\omega_k, v_k) \in \mathbb{R} \times H$, that the system of equations

$$0 = \omega_0 + \frac{1}{2}\|v_0\|^2 = \omega_1 + \frac{1}{2}\|v_1\|^2 = \frac{\omega_0 + \omega_1}{2} + \frac{1}{2}\left\|\frac{v_0 + v_1}{2}\right\|^2 \qquad (24)$$

implies $(\omega_0, v_0) = (\omega_1, v_1)$. Assume (a) and choose $(\omega_k, v_k) \in d^+ S(t, x)$, $k = 0, 1$, arbitrarily. Then also $\frac{1}{2}(\omega_0 + \omega_1, v_0 + v_1)$ belongs to $d^+ S(t, x)$. By (a), (24) is fulfilled forcing $(\omega_0, v_0) = (\omega_1, v_1)$. Thus $d^+ S(t, x)$ is a singleton $\{(\omega, v)\}$ which means that S is Gâteaux differentiable at (t, x) with gradient (ω, v). Then $d^\bullet S(t, x) = \{(\omega, v)\}$. Thus, for any sequence $(t_j, x_j) \notin \Sigma$ converging strongly to (t, x), we have $(\omega_j, v_j) := dS(t_j, x_j) \to (\omega, v)$ weakly. Owing to (a) we have $\omega + \frac{1}{2}\|v\|^2 = 0$ and using $\omega_j + \frac{1}{2}\|v_j\|^2 = 0$ we find that $\|v_j\|^2 \to \|v\|^2$ as $j \to \infty$ since $\omega_j \to \omega$. This observation elevates the weak convergence $v_j \to v$ to strong convergence. By Šmulian's theorem on Fréchet differentiability [40], [12, Thm. 4.2.10] (which extends from convex to semiconvex and semiconcave functions) we may conclude that S is Fréchet differentiable at (t, x). □

Theorem 10 *Let f be a proper lower semicontinuous function on H such that (2) is fulfilled. Then the following conditions are equivalent for any $(t_0, x_0) \in (0, \infty) \times H$.*

(i) *$S(t_0, \cdot)$ is Fréchet differentiable at x_0.*
(ii) *S is Fréchet differentiable at (t_0, x_0).*
(iii) *S is Gâteaux differentiable at (t_0, x_0) and $P_{t_0 f}(x_0)$ is a singleton.*
(iv) *S is Gâteaux differentiable at (t_0, x_0) and $P_{t_0 f}(x_0) \neq \emptyset$.*
(v) *S is Gâteaux differentiable at (t_0, x_0) and the Gâteaux gradient (ω, p) satisfies $\omega + \frac{1}{2}\|v\|^2 = 0$.*
(vi) *$\omega + \frac{1}{2}\|v\|^2 = 0$ for all $(\omega, v) \in d^+ S(t, x)$.*

Proof Proposition 4 ensures that (ii) $\Rightarrow$ (iii) while the implication (iii) $\Rightarrow$ (iv) is immediate. The implication (iv) $\Rightarrow$ (v) follows since, at (t_0, x_0), the Gâteaux gradient $(\partial S/\partial t, \nabla S)$ is equal to $(-\|x_0 - y_0\|^2/(2t_0^2), (x_0 - y_0)/t_0)$ if (iv) holds and $y_0 \in P_{t_0 f}(x_0)$ (the proof is similar to that of Proposition 6). Furthermore, (v) $\Rightarrow$ (vi) holds as $d^+ S(t_0, x_0)$ is a singleton when S is Gâteaux differentiable at (t_0, x_0). The implication (vi) $\Rightarrow$ (ii) was a key step in the proof of Proposition 7. As (ii) $\Rightarrow$ (i) holds trivially, we shall close the circle by demonstrating that (i) implies (vi). Assuming (i), pick any $(\omega, v) \in d^+ S(t_0, x_0)$. Then $v \in d_x^+ S(t_0, x_0)$ and $\omega \in d_t^+ S(t_0, x_0)$. It follows that $v = d_x S(t_0, x_0)$ and taking into account Proposition 4 we also find that $\omega = (\partial S/\partial t)(t_0, x_0)$ and $\omega + \frac{1}{2}\|v\|^2 = 0$. □

The equation $\omega + \frac{1}{2}\|v\|^2 = 0$ may not hold for any $(\omega, v) \in d^+ S(t, x)$.

Corollary 6 *If S is Gâteaux but not Fréchet differentiable at (t_0, x_0), then $P_{t_0 f}(x_0) = \emptyset$ and*

$$\frac{\partial S}{\partial t}(t_0, x_0) + \frac{1}{2}\|\nabla_x S(t_0, x_0)\|^2 < 0.$$

Example 3 shows in particular that Σ and Σ_s are different in general and that the condition

- "$S(t_0, \cdot)$ is Gâteaux differentiable at x_0 and $P_{t_0 f}(x_0)$ is a singleton"

is not equivalent to any of the conditions presented in Theorem 10.

Theorem 10 is a generalization to Moreau envelopes of Fitzpatrick's characterization of Fréchet differentiability of distance functions.

Theorem 11 (Fitzpatrick [23]) *Let E be a nonvoid closed subset of H. The following conditions are equivalent for any $x_0 \in H \setminus E$.*

(i) *d_E is Fréchet differentiable at x_0.*
(ii) *d_E is Gâteaux differentiable at x_0 and $P_E(x_0)$ is nonempty.*
(iii) *d_E is Gâteaux differentiable at x_0 and $P_E(x_0)$ is a singleton.*
(iv) *d_E is Gâteaux differentiable at x_0 and $\|\nabla d_E(x_0)\| = 1$.*
(v) *The metric projection P_E is continuous at x_0 in the sense that $P_E(x_0)$ is a singleton and $y_j \to P_E(x_0)$ whenever $x_j \to x_0$ and $y_j \in P(x_j)$.*

In particular, $\Sigma = \Sigma_s$.

For more on the theme of differentiability of Moreau envelopes, see [26].

11 Basic analysis of intrinsic characteristics for Moreau envelopes

We are now in the position to apply, on a detailed level, the ideas and methods set forth by Cannarsa and Cheng in [13] to the viscosity solution $S(t, x) = f_t(x)$ of (3)–(4). A formal time reversal motivates the definition

$$g^s(x) = \sup_{y \in H} \left(g(y) - \frac{1}{2s} \|x - y\|^2 \right).$$

In [32], Lasry and Lions introduced a regularization and approximation scheme for lower semicontinuous functions f on a Hilbert space by defining $f_{t,s} = (f_t)^s$ for $0 < s < t$. These double indexed approximations of f enjoy remarkable properties. For instance, if (2) is met, $f_{t,s} \to f$ pointwise as $0 < s < t \downarrow 0$ and $f_{t,s}$ is differentiable with a globally Lipschitz continuous differential $df_{t,s}$. Reconnecting to the approach of [13] we consider, for a fixed initial point (t_0, x_0) and any $t > t_0$, the maximization problem

$$\arg \max_{x \in H} \left(f_t(x) - \frac{1}{2(t - t_0)} \|x_0 - x\|^2 \right), \tag{25}$$

which is well-behaved since the objective function is uniformly concave by virtue of the expansion

$$x \mapsto f_t(x) - \frac{1}{2(t - t_0)} \|x_0 - x\|^2 = \text{ a concave function } - \left(\frac{1}{t - t_0} - \frac{1}{t} \right) \|x\|^2 / 2.$$

It ensues that there exists a unique maximizer in (25), which we shall denote by $x = \boldsymbol{x}(t)$, i.e.,

$$\{\boldsymbol{x}(t)\} = \arg\max_{x \in H} \left(f_t(x) - \frac{1}{2(t - t_0)} \|x_0 - x\|^2 \right), \qquad t_0 < t < \infty.$$

Clearly, $\boldsymbol{x}(t)$ is characterized as the unique solution x of

$$\frac{x - x_0}{t - t_0} \in d^+ f_t(x) \tag{26}$$

and it is straightforward to show that $\boldsymbol{x}(t_0) := \lim_{t \downarrow t_0} \boldsymbol{x}(t) = x_0$. The mapping $\boldsymbol{x} \colon [t_0, \infty) \to H$ is referred to as the *intrinsic characteristic* emanating from x_0 (thus repeating Definition 3).

We expect that minimizing straight arcs are intrinsic characteristics.

Proposition 8 *Suppose $y_1 \in P_{t_1 f}(x_1)$. Consider a point of the form*

$$x_0 = y_1 + t_0 v \quad where\ 0 < t_0 < t_1\ and\ v = \frac{x_1 - y_1}{t_1}.$$

Then the intrinsic characteristic $\boldsymbol{x}(t)$ with $\boldsymbol{x}(t_0) = x_0$ is given by

$$\boldsymbol{x}(t) = x_0 + (t - t_0)v = y_1 + tv \quad when\ \ t \in [t_0, t_1]. \tag{27}$$

Moreover, f_t is Fréchet differentiable at $\boldsymbol{x}(t)$ and

$$\dot{\boldsymbol{x}}(t) = df_t(\boldsymbol{x}(t)) = v \quad when\ \ t \in [t_0, t_1).$$

Proof By Proposition 5, f_t is Fréchet differentiable along the straight line (27) with $df_t(\boldsymbol{x}(t)) = v$ for every $t \in [t_0, t_1)$. The intrinsic characteristic is characterized as the unique solution x of (26) which is $x = x_0 + (t - t_0)v$ when $t \in (t_0, t_1)$. □

If the intrinsic characteristic $(t, \boldsymbol{x}(t))$ that starts at (t_0, x_0) passes through a non-singular point at $(t_1, \boldsymbol{x}(t_1))$ then it has to agree with a nonsingular line segment when $t \in [t_0, t_1]$.

Theorem 12 *Consider the intrinsic characteristic $\boldsymbol{x}(t)$ issuing from an arbitrary point $(t_0, x_0) \in (0, \infty) \times H$. Assume that $(t_1, \boldsymbol{x}(t_1))$ is not a strong singularity for a certain $t_1 > t_0$, which means on the one hand that $P_{t_1 f}(\boldsymbol{x}(t_1)) = \{y_1\}$ for some $y_1 \in H$ and, on the other, that the Gâteaux differential $\nabla f_{t_1}(\boldsymbol{x}(t_1))$ exists and equals $v := (\boldsymbol{x}(t_1) - y_1)/t_1$. Then*

$$\boldsymbol{x}(t) = y_1 + tv = x_0 + (t - t_0)v \quad when\ t \in [t_0, t_1],$$

while f_t is Fréchet differentiable at $\boldsymbol{x}(t)$ with $df_t(\boldsymbol{x}(t)) = v = \dot{\boldsymbol{x}}(t)$ for every $t \in [t_0, t_1)$. Furthermore, $P_{tf}(\boldsymbol{x}(t)) = \{y_1\}$ for all $t \in [t_0, t_1]$.

Proof Let $0 < t_0 < t_1$ and assume that $(t_1, \boldsymbol{x}(t_1))$ is not a strong singularity. The point $x = \boldsymbol{x}(t_1)$ is characterized by (26), i.e.,

$$\nabla f_{t_1}(\boldsymbol{x}(t_1)) = \frac{\boldsymbol{x}(t_1) - x_0}{t_1 - t_0}. \tag{28}$$

As regards $\nabla f_{t_1}(\boldsymbol{x}(t_1))$, Proposition 6 ensures that $P_{t_1 f}(\boldsymbol{x}(t_1))$ is a singleton $\{y_1\}$ and

$$\nabla f_{t_1}(\boldsymbol{x}(t_1)) = \frac{\boldsymbol{x}(t_1) - y_1}{t_1}. \tag{29}$$

It follows from (28) and (29) that

$$\frac{\boldsymbol{x}(t_1) - x_0}{t_1 - t_0} = \frac{\boldsymbol{x}(t_1) - y_1}{t_1}$$

which expresses that (t_0, x_0) is an interior point of the straight line segment

$$\left(t, \, y_1 + t \frac{\boldsymbol{x}(t_1) - y_1}{t_1} \right) = (t, \, y_1 + tv), \qquad 0 \le t \le t_1,$$

connecting $(0, y_1)$ with $(t_1, \boldsymbol{x}(t_1))$. Thus, on account of Propositions 5 and 8, first,

$$\boldsymbol{x}(t) = y_1 + tv = x_0 + (t - t_0)v \quad \text{when} \, t \in [t_0, t_1],$$

and, secondly, f_t is Fréchet differentiable at $\boldsymbol{x}(t)$ with $df_t(\boldsymbol{x}(t)) = v$ for every $t \in [t_0, t_1)$. $\qquad\square$

We may now conclude that singularities propagate forward in time along intrinsic arcs turning immediately into strong singularities.

Corollary 7 (Propagation of singularities) *Let $\boldsymbol{x}(t)$ be the intrinsic characteristic emanating from (t_0, x_0). If $t_1 \ge t_0$ and $(t_1, \boldsymbol{x}(t_1)) \in \Sigma$, then $(t, \boldsymbol{x}(t)) \in \Sigma_s$ for every $t \in (t_1, \infty)$.*

Proof The conclusion follows from Theorem 12 by contraposition. $\qquad\square$

A complementary observation concerning the points of Σ_s of vacuous proximal mapping reads as follows:

Corollary 8 *Suppose that $P_{t_1 f}(\boldsymbol{x}(t_1)) = \emptyset$. Then there exists some $\tau \in [t_0, t_1)$ such that $(t, \boldsymbol{x}(t)) \in \Sigma_s$ for all $t \in (\tau, \infty)$.*

Proof Devising the proof by contraposition, assume that $(t, \boldsymbol{x}(t)) \notin \Sigma_s$ for all $t \in [t_0, t_1)$. Owing to Corollary 7, the task at hand boils down to inferring that $P_{t_1 f}(\boldsymbol{x}(t_1)) \ne \emptyset$. Let $t_0 < \tau < t_1$. By Theorem 12 applied to the point $(\tau, \boldsymbol{x}(\tau)) \notin \Sigma_s$, we deduce that $\boldsymbol{x}(t)$ agrees with a straight line segment of the form $y(\tau) + tv(\tau)$ and that $P_{tf}(\boldsymbol{x}(t)) = \{y(\tau)\}$ when $t \in [t_0, \tau]$. We claim that $y(\tau)$ and $v(\tau)$ are in fact

independent of $\tau \in (t_0, t_1)$. To this end, pick $t_0 < \tau_0 < \tau_1 < t_1$ arbitrarily. Then, in particular,

$$\boldsymbol{x}(t) = y(\tau_0) + tv(\tau_0) = y(\tau_1) + tv(\tau_1) \quad \text{when } t \in [t_0, \tau_0],$$

forcing $y(\tau_0) = y(\tau_1)$ and $v(\tau_0) = v(\tau_1)$, and the claim follows. We conclude that there exist constant vectors $y \in H$ and $v \in H$ such that $\boldsymbol{x}(t) = y + tv$ and $P_{tf}(\boldsymbol{x}(t)) = \{y\}$ for all $t \in [t_0, t_1)$. Taking the limit as $t \uparrow t_1$ in the inequality

$$\forall z \in H \quad f(y) + \frac{1}{2t}\|\boldsymbol{x}(t) - y\|^2 \leq f(z) + \frac{1}{2t}\|\boldsymbol{x}(t) - z\|^2,$$

which is in force when $t_0 < t < t_1$, using the continuity of $\boldsymbol{x}(t)$ proved in Proposition 11, implies $y \in P_{t_1 f}(\boldsymbol{x}(t_1))$. $\qquad\square$

We next examine the intersection of intrinsic characteristics noticing that the inverse multivalued mapping $(J^{t_0 \to t})^{-1} \colon H \rightrightarrows H$, given by

$$(J^{t_0 \to t})^{-1} = I - (t - t_0)d^+ f_t,$$

maps points to bounded convex nonempty sets.

Proposition 9 *The mapping* $J^{t_0 \to t_1} \colon H \to H$ *is surjective for any* $0 < t_0 < t_1$. *In other words, through any given point* $(t_1, x_1) \in (0, \infty) \times H$ *with* $0 < t_0 < t_1$ *there passes at least one intrinsic characteristic issuing at* $t = t_0$ *from* some *initial point* x_0; *more precisely,*

$$J^{t_0 \to t_1}(x_0) = x_1$$

if and only if

$$x_0 \in (J^{t_0 \to t_1})^{-1}(x_1) = x_1 - (t_1 - t_0)d^+ f_{t_1}(x_1).$$

Furthermore, the intrinsic characteristic starting at $t = t_0$ *passing through* (t_1, x_1) *is unique exactly if* f_{t_1} *is Gâteaux differentiable at* x_1. *In this case, it is given by* $\boldsymbol{x}(t) = x_1 + (t - t_1)\nabla f_{t_1}(x_1)$ *when* $t \in [t_0, t_1]$ *and* $(t, x(t)) \notin \Sigma$ *when* $t \in [t_0, t_1)$ *if, in addition,* $P_{t_1 f}(x_1) \neq \emptyset$. *If, on the other hand,* f_{t_1} *is Gâteaux differentiable at* x_1 *while* $P_{t_1 f}(x_1) = \emptyset$, *then* $(t, \boldsymbol{x}(t)) \in \Sigma_s$ *when* $t \in (\tau, \infty)$ *for some* $\tau \in (t_0, t_1)$.

Proof Owing to Theorem 3, the description of the inverse mapping of $J^{t_0 \to t_1}$ is straightforward. The inverse mapping is single-valued at x_1 if and only if the convex function $A(t_1, \cdot)$ is Gâteaux differentiable at x_1. In this case, let the unique intrinsic characteristic be denoted by $\boldsymbol{x}(t)$, $t_0 \leq t \leq t_1$. On account of Theorem 12, it is given by $\boldsymbol{x}(t) = x_1 + (t - t_1)\nabla f_{t_1}(x_1)$ when $t \in [t_0, t_1]$ if, in addition, $P_{t_1 f}(x_1) \neq \emptyset$. The conclusion about the case when $P_{t_1 f}(x_1) = \emptyset$ comes from Corollary 8. $\qquad\square$

12 Regularity of intrinsic characteristics

We recall that $(\omega°(t, x), v°(t, x)) \in d^+ S(t, x)$ satisfies

$$\omega°(t, x) + \frac{1}{2}\|v°(t, x)\|^2 \leq \omega + \frac{1}{2}\|v\|^2 \quad \text{for all } (\omega, v) \in d^+ S(t, x).$$

In the terminology of [27, 28], $v°(t, x)$ is the admissible velocity. In $H = \mathbb{R}^n$, the following result was obtained for certain Hamilton–Jacobi equations of the form $\mathcal{H}(x, \nabla S(x)) = 0$ in the seminal paper [13, Prop. 3.4] on intrinsic characteristics. The initial velocity is the right derivative

$$\dot{x}^+(t_0) = \lim_{t \downarrow t_0} \frac{x(t) - x_0}{t - t_0}$$

at the initial point.

Proposition 10 *The right derivative $\dot{x}^+(t_0)$ exists and $\dot{x}^+(t_0) = v°(t_0, x_0)$.*

Proof Let $t_k \downarrow t_0$ and consider

$$v_k = \frac{x(t_k) - x_0}{t_k - t_0}, \quad k = 1, 2, \ldots.$$

The task is to establish that $v_k \to v°(t_0, x_0)$ strongly. Choose $\delta > 0$ so large that $(t_k, x(t_k)) \in \Omega = [t_0, t_0 + \delta] \times \overline{\mathbb{B}}(x_0, \delta)$ for all $k \geq 1$. The function $S(t, x) = f_t(x)$ is Lipschitz continuous and semiconcave in Ω. From (26) we see that $v_k \in d_x^+ S(t_k, x(t_k))$ and, thus, (v_k) is a bounded sequence possessing a weakly convergent subsequence, still labelled by (v_k), i.e., $v_k \to v_0$ weakly. In fact, the semiconcavity of S ensures that $(\omega_k, v_k) \in d^+ S(t_k, x(t_k))$ for some bounded real sequence (ω_k). By extracting a further subsequence, we may assume that $\omega_k \to \omega_0$ as $k \to \infty$. On account of the semiconcavity,

$$\begin{aligned}
S(t_0, x_0) \leq\ & S(t_k, x(t_k)) + \omega_k(t_0 - t_k) + \langle v_k, x_0 - x(t_k)\rangle \\
& + C((t_0 - t_k)^2 + \|x_0 - x(t_k)\|^2)
\end{aligned} \tag{30}$$

Similarly, for any $(\omega, v) \in d^+ S(t_0, x_0)$ it holds that

$$\begin{aligned}
S(t_k, x(t_k)) \leq\ & S(t_0, x_0) + \omega(t_k - t_0) + \langle v, x(t_k) - x_0\rangle \\
& + C((t_k - t_0)^2 + \|x(t_k) - x_0\|^2).
\end{aligned} \tag{31}$$

In (30) and (31), C is a constant of semiconcavity of S on Ω. Inequalities (30) and (31) add to

$$0 \leq (\omega - \omega_k)(t_k - t_0) + \langle v - v_k, x(t_k) - x_0\rangle + 2C((t_k - t_0)^2 + \|x(t_k) - x_0\|^2)$$

or, dividing through by $t_k - t_0 > 0$, to

$$0 \leq \omega - \omega_k + \langle v - v_k, v_k \rangle + 2C(1 + \|v_k\|^2)(t_k - t_0).$$

Rearranging terms and passing to the limit yields

$$\|v_0\|^2 \leq \liminf_{k \to \infty} \|v_k\|^2 \leq \limsup_{k \to \infty} \|v_k\|^2$$

$$\leq \limsup_{k \to \infty} \Big(\omega - \omega_k + \langle v, v_k \rangle + 2C(1 + \|v_k\|^2)(t_k - t_0) \Big) = \omega - \omega_0 + \langle v, v_0 \rangle.$$

(32)

Thus, we have demonstrated that

$$\omega - \omega_0 + \langle v - v_0, v_0 \rangle \geq 0 \quad \text{for all } (\omega, v) \in d^+ S(t_0, x_0). \tag{33}$$

Owing to the outer semicontinuity of the multivalued mapping $d^+ S$, the weak limit (ω_0, v_0) belongs to $d^+ S(t_0, x_0)$; this conclusion is confirmed in Lemma 6 below. To prove that actually $v_k \to v_0$ in norm it is sufficient to notice that each inequality in (32) is an equality when $(\omega, v) = (\omega_0, v_0)$. Indeed, this choice shows that

$$\lim_{k \to \infty} \|v_k\|^2 = \|v_0\|^2$$

which together with $v_k \to v_0$ weakly implies that $v_k \to v_0$ in norm.

Finally, inequality (33) implies that $(\omega_0, v_0) = (\omega^\circ(t_0, x_0), v^\circ(t_0, x_0))$. Indeed, by expansion of $\|v - v_0\|^2$ the following inequality holds for all $(\omega, v) \in \mathbb{R} \times H$ with $v \neq v_0$:

$$\omega + \frac{1}{2} \|v\|^2 - \left(\omega_0 + \frac{1}{2} \|v_0\|^2 \right) > \omega - \omega_0 + \langle v - v_0, v_0 \rangle,$$

which combined with (33) yields $(\omega_0, v_0) = (\omega^\circ(t_0, x_0), v^\circ(t_0, x_0))$. $\qquad\square$

Lemma 6 *Let $\mathbb{B}$ be an open ball in a Hilbert space $\mathbb{H}$. Let $F : \mathbb{H} \to \mathbb{R}$ be such that $F(X) - C\|X\|^2/2$ is concave and continuous in the closed ball $\overline{\mathbb{B}}$ for some constant C. If $X_j \to \bar{X} \in \mathbb{B}$ in norm, $P_j \in d^+ F(X_j)$ and $P_j \to \bar{P}$ weakly, then $\bar{P} \in d^+ F(\bar{X})$.*

Proof We convert to the convex function defined as $G(X) = C\|X\|^2/2 - F(X)$ in the closed ball $\overline{\mathbb{B}}$ and as $G(X) = \infty$ otherwise. We set $Q_j = CX_j - P_j$ which converges weakly to $\bar{Q} = C\bar{X} - \bar{P}$. Then, $Q_j \in \partial G(X_j)$ in the sense of convex analysis which means that

$$G(X) \geq G(X_j) + \langle X - X_j, Q_j \rangle \quad \text{for all } X \in \mathbb{H}.$$

Sending $j \to \infty$ yields

$$G(X) \geq G(\bar{X}) + \langle X - \bar{X}, \bar{Q} \rangle \quad \text{for all } X \in \mathbb{H},$$

which says that $\bar{Q} \in \partial G(\bar{X})$ and, thus, that $\bar{P} \in d^+ F(\bar{X})$. □

Remark 5 The generalized characteristic defined by (5) actually satisfies $\dot{X}^+(t) = v^\circ(t, X(t))$ for all $t \in [t_0, \infty)$ owing to its uniqueness [17, Cor. 3.4].

We proceed to the regularity of $x(t)$. We base the proof partly on the regularity of $A(t, x)$. The most advantageous case is when $f = \mathfrak{I}_E$ since then A is independent of t.

Proposition 11 *The intrinsic characteristic $x(t)$ issuing from (t_0, x_0) is locally Hölder continuous with exponent $1/2$. It is locally Lipschitz continuous if $\partial A/\partial t$ exists and is a locally Lipschitz continuous function of $(t, x) \in [t_0, \infty) \times H$. Moreover, if f is the indicator function of a closed nonempty set $E \subset H$, or if $\partial A/\partial t$ is a constant, then $x(t)$ satisfies the Lipschitz estimates of Theorem 5.*

Proof For any $t_0 < s < t$ we set

$$p_t = x(t) - \frac{t(x(t) - x_0)}{t - t_0} \quad \text{and} \quad p_s = x(s) - \frac{s(x(s) - x_0)}{s - t_0}$$

noting that $p_t \in \partial_x A(t, x(t))$ and $p_s \in \partial_x A(s, x(s))$ on account of (49). Since A is convex there exist real numbers α_t, α_s such that $(\alpha_t, p_t) \in \partial A(t, x(t))$, $(\alpha_s, p_s) \in \partial A(s, x(s))$, and hence

$$(\alpha_t - \alpha_s)(t - s) + \langle p_t - p_s, x(t) - x(s) \rangle \geq 0. \tag{34}$$

We next expand

$$\mathcal{Q} := \|x(t) - x(s)\|^2 \tag{35}$$

to

$$\mathcal{Q} = -\frac{(t - t_0)\langle p_t - p_s, x(t) - x(s) \rangle}{t_0} + \left\langle \frac{x(s) - x_0}{s - t_0}, x(t) - x(s) \right\rangle (t - s). \tag{36}$$

Taking into account (34) and

$$\frac{x(s) - x_0}{s - t_0} \in d^+ f_s(x(s))$$

we can derive upper bounds for (36) as follows

$$\mathcal{Q} \leq \frac{(t - t_0)(\alpha_t - \alpha_s)(t - s)}{t_0} + \left\langle \frac{x(s) - x_0}{s - t_0}, x(t) - x(s) \right\rangle (t - s) \tag{37}$$

$$\leq \left(\frac{(t - t_0)(\alpha_t - \alpha_s)}{t_0} + M(s)\|x(t) - x(s)\| \right) (t - s) \tag{38}$$

where

$$M(s) = \sup_{v \in d^+ f_s(\boldsymbol{x}(s))} \|v\|.$$

Using that $|\alpha_t - \alpha_s|$ and $M(s)$ are bounded when $t_0 \leq s < t \leq T$ we infer, from (35) and (37)–(38), the Hölder estimate $\|\boldsymbol{x}(t) - \boldsymbol{x}(s)\| \leq C_T (t-s)^{1/2}$ for $t_0 \leq s < t \leq T < \infty$.

Let us now add the hypothesis that $\partial A / \partial t$ be locally Lipschitz continuous when $t \geq t_0$. We claim that the average velocity

$$v_{s \to t} = \frac{\boldsymbol{x}(t) - \boldsymbol{x}(s)}{t - s}$$

is then bounded when $t_0 \leq s < t \leq T$. Observing that

$$\alpha_t - \alpha_s = \frac{\partial A}{\partial t}(t, \boldsymbol{x}(t)) - \frac{\partial A}{\partial t}(s, \boldsymbol{x}(s)) \leq C_T ((t-s) + \|\boldsymbol{x}(t) - \boldsymbol{x}(s)\|)$$

when $t_0 \leq s < t \leq T$ we find, by means of (38), that

$$\|v_{s \to t}\|^2 \leq \frac{C_T (t - t_0)}{t_0} (1 + \|v_{s \to t}\|) + M_T \|v_{s \to t}\|.$$

We may now infer the speed bound $\|v_{s \to t}\| \leq L_T$ when $t_0 \leq s < t \leq T < \infty$ for some constant L_T depending on T.

Finally we assume instead that $\alpha_t \leq \alpha_s$, which certainly holds if $\partial A / \partial t$ is a constant. For instance, if $f = \mathfrak{I}_E$, then $A(t, x)$ is independent of t and $\alpha_t = 0 = \alpha_s$. In this special case we can derive the inequality

$$\|v_{s \to t}\|^2 \leq \langle v_{t_0 \to s}, v_{s \to t} \rangle \tag{39}$$

for the average velocities

$$v_{t_0 \to s} = \frac{\boldsymbol{x}(s) - x_0}{s - t_0} \quad \text{and} \quad v_{s \to t} = \frac{\boldsymbol{x}(t) - \boldsymbol{x}(s)}{t - s}.$$

Indeed, combining (35) with (37) yields

$$\|\boldsymbol{x}(t) - \boldsymbol{x}(s)\|^2 \leq \left\langle \frac{\boldsymbol{x}(s) - x_0}{s - t_0}, \boldsymbol{x}(t) - \boldsymbol{x}(s) \right\rangle (t - s)$$

and (39) ensues by division with $(t - s)^2$. Inequality (39) is equivalent to (10), i.e., to

$$\|v_{s \to t}\|^2 - \|v_{t_0 \to s}\|^2 \leq -\|v_{s \to t} - v_{t_0 \to s}\|^2, \tag{40}$$

which is confirmed by an expansion of the right-hand side of (40); in particular, $\|v_{s \to t}\| < \|v_{t_0 \to s}\|$ unless $v_{s \to t} = v_{t_0 \to s}$. $\qquad \square$

13 A few propositions about distance functions

We recall that $\Sigma_s = \Sigma$ in the case of distance functions by Theorem 11. In this case, $x(t)$ is given by (8) or, equivalently, in setting

$$\theta(t) = \frac{t}{t - t_0}, \qquad t_0 < t < \infty,$$

$x = x(t)$ if and only if

$$\theta(t)(x - x_0) \in d^+ d_E^2(x)/2.$$

The function $\theta(\cdot)$ is decreasing and its range is $(1, \infty)$. The intrinsic characteristic $x(t)$ does not return to the initial point x_0 unless $x(t) \equiv x_0$. Recall that a point $x_0 \in H$ is called critical if $0 \in d^+ d_E^2(x_0)$.

Proposition 12 *If x_0 is a critical point then $x(t) = x_0$ for all $t \in [t_0, \infty)$. Otherwise, $x(t) \neq x_0$ for all $t \in (t_0, \infty)$ and $\dot{x}^+(t_0) = v^\circ(t_0, x_0) = r(x_0)/t_0 \neq 0$.*

Proposition 13 *Assume that x_0 is not a critical point, $t_0 < t_1 < t_2$ and $x(t_1) = x(t_2)$. Then $x(t_j) \in \Sigma$ and x is constant on $[t_1, t_2]$.*

Proof Setting $\bar{x} = x(t_1) = x(t_2)$ we have $\bar{x} \neq x_0$ and $\theta(t_j)(\bar{x} - x_0) \in d^+ d_E^2(\bar{x})/2$ for $j = 1, 2$, implying $\bar{x} \in \Sigma$. Let $t_1 < t < t_2$; then also $\theta(t)(\bar{x} - x_0) \in d^+ d_E^2(\bar{x})/2$, and hence $x(t) = \bar{x}$, because $\theta(t_2) < \theta(t) < \theta(t_1)$ and $d^+ d_E^2(\bar{x})/2$ is a convex set. $\square$

If $x(t)$ reaches a critical point, then it will come to a stop there.

Proposition 14 *Assume that $x(t_1)$ is a critical point for some $t_1 \in (t_0, \infty)$. Then $x(t) = x(t_1)$ for all $t \in [t_1, \infty)$.*

Proof The point $x(t_1)$ is characterized by

$$\theta(t_1)(x(t_1) - x_0) \in d^+ d_E^2(x(t_1))/2;$$

hence

$$\text{co}\,\{0, \theta(t_1)(x(t_1) - x_0)\} \subset d^+ d_E^2(x(t_1))/2. \tag{41}$$

In order to demonstrate that $x(t) = x(t_1)$ for any $t > t_1$ it suffices to show that

$$\theta(t)(x(t_1) - x_0) \in d^+ d_E^2(x(t_1))/2. \tag{42}$$

To this end we need only observe that the left-hand side of (42) is an element of the left-hand side of (41). Indeed,

$$\theta(t)(x(t_1) - x_0) = (1 - \lambda)0 + \lambda\theta(t_1)(x(t_1) - x_0)$$

for the scalar $\lambda = \theta(t)/\theta(t_1) \in (0, 1)$. $\square$

Proposition 15 *The distances $d_E(\boldsymbol{x}(t))$ and $\|x_0 - \boldsymbol{x}(t)\|$ are nondecreasing functions of $t \in [t_0, \infty)$. In fact, if $t_0 \leq t_1 < t_2$, then*

$$d_E(\boldsymbol{x}(t_1)) < d_E(\boldsymbol{x}(t_2)) \quad and \quad \|x_0 - \boldsymbol{x}(t_1)\| < \|x_0 - \boldsymbol{x}(t_2)\|$$

unless $\boldsymbol{x}(t_1) = \boldsymbol{x}(t_2)$. Moreover, the functions $d_E(\boldsymbol{x}(t))$ and $\|x_0 - \boldsymbol{x}(t)\|$ are either simultaneously bounded or simultaneously unbounded.

Proof For arbitrary fixed $t_0 < t_1 < t_2$ we set $F(x) = d_E^2(x)/2$, $G(x) = \|x_0 - x\|^2/2$ and, for $j = 1, 2$, $x_j = \boldsymbol{x}(t_j)$ and $\theta_j = \theta(t_j)$. Then $\theta_2 < \theta_1$ and

$$\{x_j\} = \arg\max_{x \in H} (F(x) - \theta_j G(x)).$$

Suppose that $x_1 \neq x_2$; then

$$F(x_1) - \theta_1 G(x_1) > F(x_2) - \theta_1 G(x_2)$$

as well as

$$F(x_2) - \theta_2 G(x_2) > F(x_1) - \theta_2 G(x_1),$$

which combine to

$$\theta_2(G(x_2) - G(x_1)) < F(x_2) - F(x_1) < \theta_1(G(x_2) - G(x_1)). \tag{43}$$

Taking into account that $\theta_1 > \theta_2 > 0$ it follows that each membrum of (43) is positive. Thus, $F(x_2) > F(x_1)$ and $G(x_2) > G(x_1)$. To prove the boundedness assertion, we start by recalling that by (8), for any fixed $t > t_0$, $\boldsymbol{x}(t)$ is the maximizer of

$$\Psi(x) = \frac{1}{2t} d_E^2(x) - \frac{1}{2(t - t_0)} \|x_0 - x\|^2.$$

As $\Psi(\boldsymbol{x}(t)) = \sup \Psi \geq \Psi(x_0)$ we must have

$$\frac{1}{2t} d_E^2(\boldsymbol{x}(t)) - \frac{1}{2(t - t_0)} \|x_0 - \boldsymbol{x}(t)\|^2 \geq \frac{1}{2t} d_E^2(x_0)$$

implying

$$\|x_0 - \boldsymbol{x}(t)\|^2 \leq \frac{(t - t_0)(d_E^2(\boldsymbol{x}(t)) - d_E^2(x_0))}{t} \leq d_E^2(\boldsymbol{x}(t)) - d_E^2(x_0).$$

In particular, if $d_E(\boldsymbol{x}(t))$ is a bounded function of $t \in [t_0, \infty)$ then so is $\|x_0 - \boldsymbol{x}(t)\|$. Conversely, it is easily verified that the boundedness of $\|x_0 - \boldsymbol{x}(t)\|$ implies that of $d_E(\boldsymbol{x}(t))$. $\qquad\square$

14 A further tool for propagation of singularities

The propagation results of Theorems 4–7 have the strength of being global in time t. Even if $\boldsymbol{x}(t)$ eventually becomes constant, the propagation still persists along $(t, \boldsymbol{x}(t))$; see Example 8 in the final section for an elementary illustration. In the case of distance functions, we also know that the speed of propagation along $\boldsymbol{x}(t)$ does not exceed the initial speed. Still there is a flaw in this picture. We are primarily interested in the propagation along $\boldsymbol{x}(t)$ in $\complement E$, thinking of t as a parameter. It may happen that $\boldsymbol{x}(t)$ becomes constant after some time while singularities may continue propagating along some other arc. Suppose, for the sake of argument, that an intrinsic characteristic $\boldsymbol{x}(t)$ is nonsingular when $t < t_1$ but $\boldsymbol{x}(t_1) = J^{t_0 \to t_1}(x_0)$ is a critical point implying that $\boldsymbol{x}(t) = \boldsymbol{x}(t_1)$ for all $t \in [t_1, \infty)$, as confirmed by Proposition 14. It is clear that the arc $t \mapsto J^{t_1 \to t}(\boldsymbol{x}(t_1))$, with $\boldsymbol{x}(t_1)$ as the new initial point, is constant as well when $t \geq t_1$. Still there exists a nonconstant singular Lipschitz arc emanating from $\boldsymbol{x}(t_1)$ provided $P_E(\boldsymbol{x}(t_1)) \neq \mathbb{S}(\boldsymbol{x}(t_1), d_E(\boldsymbol{x}(t_1)))$ (which rules out the situation of Example 1); see Theorem 1. In our next theorem we use the operator $J^{t_0 \to t_1}$ again to construct such a singular arc by letting it act on a certain straight line segment rather than the point x_0 alone.

Theorem 13 *Let E be a closed nonempty subset of H. Assume that $\boldsymbol{x}(t_1) = J^{t_0 \to t_1}(x_0) \in \Sigma$ where $t_0 < t_1$ and $P_E(\boldsymbol{x}(t_1)) \neq \mathbb{S}(\boldsymbol{x}(t_1), d_E(\boldsymbol{x}(t_1)))$. Then a singular Lipschitz arc $X(s) \in \Sigma$, $0 \leq s \leq s_0$, satisfying $X(0) = \boldsymbol{x}(t_1)$ and $X(s) \neq \boldsymbol{x}(t_1)$ for all $s \in (0, s_0]$ is obtained by defining*

$$X(s) = J^{t_0 \to t_1}\left(\frac{t_0}{t_1}\boldsymbol{x}(t_1) + \frac{t_1 - t_0}{t_1}(y_0 + su)\right), \quad 0 \leq s \leq s_0,$$

where $y_0 \in \mathbb{B}(\boldsymbol{x}(t_1), d_E(\boldsymbol{x}(t_1)))$ is a boundary point of $\partial A(\boldsymbol{x}(t_1))$, $0 < s_0 < d_E(y_0)/2$, while u is any unit vector such that $y_0 + su \notin \partial A(\boldsymbol{x}(t_1))$ for all $s > 0$.

Proof We modify the proof in [24]. Abbreviating $J^{t_0 \to t_1}$ to J, Theorem 3 tells us that

$$\boldsymbol{x}(t_1) = J(x_0) = ((1 - \lambda)\partial A + \lambda I)^{-1}(x_0)$$

for $\lambda = t_0/t_1$. On account of Lemmas 3–4, while included in $\overline{\mathbb{B}}(\boldsymbol{x}(t_1), d_E(\boldsymbol{x}(t_1)))$, the boundary of $\partial A(\boldsymbol{x}(t_1))$ intersects the open ball $\mathbb{B}(\boldsymbol{x}(t_1), d_E(x_1))$. We may therefore select a boundary point y_0 of $\partial A(\boldsymbol{x}(t_1))$ and a unit vector u satisfying $\|\boldsymbol{x}(t_1) - y_0\| < d_E(\boldsymbol{x}(t_1))$ and $y_s := y_0 + su \notin \partial A(\boldsymbol{x}(t_1))$ for all $s > 0$. In particular, $y_0 \notin E$, and choosing $0 < s_0 < d_E(y_0)/2$, we find that, for any $s \in [0, s_0]$,

$$|d_E(y_s) - d_E(y_0)| \leq \|y_s - y_0\| = s \leq s_0 < d_E(y_0)/2$$

forcing

$$s < d_E(y_s). \tag{44}$$

On the sole basis of the definition of J, for any $x \in H$ and $y \in H$ it holds that

$$J(\lambda x + (1 - \lambda)y) = x \quad \Leftrightarrow \quad y \in \partial A(x) \tag{45}$$

as well as

$$x - \lambda J(x) \in (1 - \lambda)\partial A(J(x)). \tag{46}$$

We set

$$z_s = \lambda \boldsymbol{x}(t_1) + (1 - \lambda)y_s, \qquad w_s = \frac{z_s - \lambda J(z_s)}{1 - \lambda},$$

and examine the Lipschitz continuous curve $X(s) = J(z_s)$ for $s \in [0, s_0]$. The equivalence (45) ensures that $X(0) = \boldsymbol{x}(t_1)$ and $X(s) \neq \boldsymbol{x}(t_1)$ for every $s \in (0, s_0]$. Thus $X(\cdot)$ does not reduce to a single point. It remains only to prove that $X(s)$ is a singular arc. On the one hand, (46) ensures that $w_s \in \partial A(X(s))$ and, on the other, taking into account the Lipschitz continuity of J (with constant $1/\lambda$ by Theorem 3) and (44),

$$\|y_s - w_s\| = \frac{\lambda}{1 - \lambda}\|J(z_s) - J(z_0)\| \leq \frac{\lambda}{1 - \lambda}\frac{1}{\lambda}\|(1 - \lambda)(y_s - y_0)\| = s < d_E(y_s)$$

implying $w_s \notin E$. Hence, A is not Fréchet differentiable at $X(s)$, because if it were then $w_s = dA(X(s)) = P_E(X(s)) \in E$ by Proposition 6. Thereby, we have demonstrated that $X(s) \in \Sigma$ for every $s \in [0, s_0]$, concluding the proof. □

Remark 6 In Example 9 in the final section Σ is a hyperplane consisting of critical points. If the initial point x_0 is nonsingular, then $\boldsymbol{x}(t)$ reaches Σ in a finite time and comes to a stop in Σ. By contrast, $X(s)$ is a nonconstant singular arc.

A version of Theorem 13 for Moreau envelopes concludes this section. It asserts the existence of a nonconstant singular arc in $\{t_1\} \times H$ starting from a strictly singular point (t_1, x_1). By Proposition 6, $P_{tf}(x) \subseteq \partial_x A(t, x)$.

Theorem 14 *Given a lower semicontinuous function $f \colon H \to \mathbb{R} \cup \{\infty\}$ fulfilling (2), let (t_1, x_1) be a strictly singular point such that some point of the boundary of $\partial_x A(t_1, x_1)$ is not a member of $P_{t_1 f}(x_1)$. Then there exists a Lipschitz curve $X(s)$ defined for $s \in [0, s_0]$ such that $X(0) = x_1$ while $X(s) \neq x_1$ and $(t_1, X(s)) \in \Sigma_s$ for all $s \in (0, s_0]$.*

Proof Fix $0 < t_0 < t_1$, set $\lambda = t_0/t_1$ and $J = J^{t_0 \to t_1}$. Then, for any $x \in H$ and $y \in H$ it holds that

$$J(\lambda x + (1 - \lambda)y) = x \quad \Leftrightarrow \quad y \in \partial_x A(t_1, x) \tag{47}$$

as well as

$$x - \lambda J(x) \in (1 - \lambda)\partial_x A(t_1, J(x)). \tag{48}$$

Let y_0 be a boundary point of $\partial_x A(t_1, x_1)$ such that $y_0 \notin P_{t_1 f}(x_1)$ and select a unit vector u such that $y_0 + su \notin \partial_x A(t_1, x_1)$ for all $s > 0$. We set

$$z_s = \lambda x_1 + (1 - \lambda)(y_0 + su), \qquad w_s = \frac{z_s - \lambda J(z_s)}{1 - \lambda},$$

and investigate the Lipschitz continuous arc $X(s) = J(z_s)$ for $s \in [0, s_0]$ where the value of s_0 is specified below. The equivalence (47) ensures that $X(0) = x_1$ and $X(s) \neq x_1$ for every $s \in (0, s_0]$. To prove that $X(s)$ is a strictly singular arc, first, (48) shows that $w_s \in \partial_x A(t_1, X(s))$. Furthermore,

$$m(s) = \inf_{y \in H} \left(f(y) + \frac{1}{2t_1} \|X(s) - y\|^2 \right)$$

is an upper semicontinuous function and, as $w_0 = y_0 \notin P_{t_1 f}(x_1)$,

$$m(0) = \inf_{y \in H} \left(f(y) + \frac{1}{2t_1} \|x_1 - y\|^2 \right) < f(w_0) + \frac{1}{2t_1} \|x_1 - w_0\|^2 =: M.$$

Hence, $\limsup_{s \downarrow 0} m(s) \leq m(0) < M$ and so, for some $\varepsilon > 0$ and some $s_0 > 0$, $m(s) < M - \varepsilon$ for all $s \in [0, s_0]$. Moreover, by lower semicontinuity,

$$M \leq \liminf_{s \downarrow 0} \left(f(w_s) + \frac{1}{2t_1} \|X(s) - w_s\|^2 \right);$$

making s_0 smaller if necessary, it ensues that $w_s \notin P_{t_1 f}(X(s))$ for all $s \in [0, s_0]$. Hence $(t_1, X(s)) \in \Sigma_s$, otherwise $w_s = \nabla_x A(t_1, X(s)) = P_{t_1 f}(X(s))$ by Proposition 6. $\quad\square$

15 Proofs of Theorems 3–9 and Corollary 3

We are now ready to present complete proofs of our most central results. We may convert (11) to the uniformly convex minimization problem

$$\{\boldsymbol{x}(t)\} = \arg\min_{x \in H} \left(\frac{1}{t} A(t, x) - \frac{1}{2t} \|x\|^2 + \frac{1}{2(t - t_0)} \|x_0 - x\|^2 \right). \qquad (49)$$

Proof of Theorem 3 (i) From (49) we infer that $\boldsymbol{x}(t)$ is given by

$$\{\boldsymbol{x}(t)\} = \left(\frac{t - t_0}{t} \partial_x A(t, \cdot) + \frac{t_0}{t} I \right)^{-1} (x_0)$$

$$= \partial \left(\frac{t - t_0}{t} A(t, \cdot) + \frac{t_0}{2t} \| \cdot \|^2 \right)^* (x_0) = \partial F(x_0).$$

We notice that $F = G^*$ where $G - \alpha \| \cdot \|^2 / 2$ is a convex function for $\alpha = t_0/t$. Hence, by the duality theory for the Legendre–Fenchel transform, F is Fréchet differentiable

and $dF = dG^*$ is Lipschitz continuous with constant $1/\alpha = t/t_0$. Parts (ii)–(iii) are covered by Propositions 10–11. □

Proof of Theorem 4 Statement (i) is identical to Corollary 7. To demonstrate (ii) we need only recall that $\dot{x}^+(t_0) = v^\circ(t_0, x_0) \in d^+ f_{t_0}(x_0)$ by Proposition 10. □

Proof of Theorem 5 Proposition 11 covers this case. Estimate (10) implies (11) in the limit as $s \downarrow t_0$ while (12) is obtained when (11) is substituted into (10). □

Proof of Theorem 6 Parts (i) and (ii) are covered by Propositions 10–11. For (iii) see Proposition 15. Fort part (iv) see Corollary 7. □

Proof of Theorem 7 See Proposition 8 and Theorem 12. Case (i) is equivalent to

$$\|x - (y_0 + tv_0)\|^2 > t^2\|v_0\|^2 \quad \text{for all } t > 0, x \in E, x \neq y_0,$$

which is fulfilled if and only if

$$\langle x - y_0, v_0 \rangle < \frac{1}{2t}\|x - y_0\|^2 \quad \text{for all } t > 0, x \in E, x \neq y_0,$$

or, equivalently, exactly when $\langle x - y_0, v_0 \rangle \leq 0$ for all $x \in E$. Assuming alternative (ii), Proposition 8 implies that $P_E(x(t))$ cannot contain y_0 for any $t > t^*$. □

Proof of Theorem 8 As $R/d_E(x) > 1$ for all $x \in \complement E$ (owing to $R > \rho(\complement E)$) it follows from Proposition 3 that the composite mapping

$$F(\theta, x) = \Phi(1 - \theta + \theta R/d_E(x), x) \quad \text{where} \quad \Phi(\tau, x) = J^{1 \to \tau}(x)$$

is well-defined on $[0, 1] \times \complement E$ and continuous. In addition, Proposition 15 ensures that $F(\theta, x) \in \complement E$ for all $(\theta, x) \in [0, 1] \times \complement E$. Obviously, $F(0, \cdot)$ is the identity mapping while $F(\theta, \cdot)$ maps Σ into Σ for every $\theta \in [0, 1]$ by Corollary 7. As regards $F(1, \cdot)$, $F(1, x) = J^{1 \to R/d_E(x)}(x) \in \Sigma$ for any $x \in \complement E$ owing to Corollary 2 as $t^*/t_0 < R/d_E(x)$ when $x \notin \Sigma$.

To derive Lipschitz estimates for F, we invoke Proposition 3. First, for any $\theta_j \in [0, 1]$ and any $x \in \complement E$, by virtue of (15),

$$\|F(\theta_1, x) - F(\theta_0, x)\|$$
$$\leq d_E(x) \left| (1 - \theta_1) + \frac{\theta_1 R}{d_E(x)} - \left((1 - \theta_0) + \frac{\theta_0 R}{d_E(x)} \right) \right| = (R - d_E(x))|\theta_1 - \theta_0|.$$

Secondly, for arbitrary $\theta \in [0, 1]$ and $x_j \in \complement E$, owing to (15)–(16),

$$\|F(\theta, x_1) - F(\theta, x_0)\| \leq \|\Phi(1 - \theta + \theta R/d_E(x_1), x_1) - \Phi(1 - \theta + \theta R/d_E(x_1), x_0)\|$$
$$+ \|\Phi(1 - \theta + \theta R/d_E(x_1), x_0) - \Phi(1 - \theta + \theta R/d_E(x_0), x_0)\|$$
$$\leq \left(1 - \theta + \frac{\theta R}{d_E(x_1)} \right) \|x_1 - x_0\| + d_E(x_0) \left| \frac{\theta R}{d_E(x_1)} - \frac{\theta R}{d_E(x_0)} \right|$$

$$= \left(1 - \theta + \frac{\theta R}{d_E(x_1)}\right) \|x_1 - x_0\| + \theta R \frac{|d_E(x_0) - d_E(x_1)|}{d_E(x_1)}$$

$$\leq \left(1 - \theta + \frac{2\theta R}{d_E(x_1)}\right) \|x_1 - x_0\| \leq \frac{2R}{d_E(x_1)} \|x_1 - x_0\|.$$

Taking into account the symmetry of the left-hand side, the denominator on the last lines may be replaced by $d_E(x_0)$ or by $\max(d_E(x_1), d_E(x_0))$.   □

Proof of Theorem 9 By (9),

$$x_0 - \frac{t_0}{t} \boldsymbol{x}(t) \in \frac{t - t_0}{t} \partial A(\boldsymbol{x}(t))$$

which yields

$$y_t := \frac{t x_0 - t_0 \boldsymbol{x}(t)}{t - t_0} \in \partial A(\boldsymbol{x}(t)).$$

We note that $y_t \to x_0$ strongly as $t \to \infty$ since $\boldsymbol{x}(\cdot)$ is bounded by hypothesis. Consider a weakly convergent sequence $\boldsymbol{x}(t_j) \to \bar{x}$ where $t_j \to \infty$ as $j \to \infty$. Passing to the limit along the sequence $t = t_j$ in the subgradient inequality

$$A(x) \geq A(\boldsymbol{x}(t)) + \langle x - \boldsymbol{x}(t), y_t \rangle \quad \text{for all } x \in H,$$

yields, owing to the weak lower semicontinuity of the convex function A,

$$A(x) \geq A(\bar{x}) + \langle x - \bar{x}, x_0 \rangle \quad \text{for all } x \in H,$$

telling us that $x_0 \in \partial A(\bar{x})$. From convex analysis we know that $x_0 \in \partial A(\bar{x})$ is equivalent to $\bar{x} \in \partial A^*(x_0)$. In particular, $x_0 \in \overline{co}E$ because $\operatorname{dom} \partial A^* \subseteq \overline{co}E$ by Lemma 5. We claim that $\partial A^*(x_0) \subseteq \Sigma \setminus \{x_0\}$. Let $y_0 \in \partial A^*(x_0)$. Then y_0 must be a singular point. Indeed, if A were Fréchet differentiable at y_0, then $x_0 = dA(y_0)$ and $P_E(y_0) = \{x_0\}$ (see Proposition 6) violating the assumption that $x_0 \in \complement E$. We also note that $y_0 \neq x_0$, otherwise x_0 would be a critical point by Proposition 1, establishing $\partial A^*(x_0) \subseteq \Sigma \setminus \{x_0\}$ and (17). Being the set of weak limit points, W is weakly closed.   □

Proof of Corollary 3 (a) W and $\partial A^*(x_0)$ are weakly compact because $\partial A^*(x_0)$ is a closed convex and bounded set in this case. To demonstrate part (b) it suffices to note that $\partial A^*(x_0)$ consists of $\nabla A^*(x_0)$ alone when $\nabla A^*(x_0)$ exists. As regards (c), returning to the proof of Theorem 9 above, $\boldsymbol{x}(t_j) \in \partial A^*(y_{t_j})$ where $y_{t_j} \to x_0$ strongly and $\boldsymbol{x}(t_j) \to \bar{x}$ weakly. By Šmulian's theorem [12, Thm. 4.2.10], if A^* is Fréchet differentiable at x_0, then $\boldsymbol{x}(t_j) \to \bar{x} = dA^*(x_0)$ strongly.   □

16 Examples

We return to the logical relations between the following conditions when dim $H = \infty$:

(i) f_t is Fréchet differentiable at x;
(ii) f_t is Gâteaux differentiable at x;
(iii) $P_{tf}(x)$ is a singleton.

While it is correct that (ii) $\Leftarrow$ (i) $\Rightarrow$ (iii), no other implication is valid between any other pair of these conditions. To justify this assertion we give three examples, two of which are imported from [23]. In Example 3, for a certain t, (ii) and (iii) are in force for all $x \in H$ yet (i) fails for some $x \in H$. In Example 4, only condition (iii) is satisfied. Example 5 displays a situation where only (ii) is fulfilled.

Example 3 Assuming dim $H = \infty$, we select a Gâteaux differentiable continuous convex function $g \colon H \to \mathbb{R}$ which is not everywhere Fréchet differentiable and which satisfies $c\| \cdot \|^2/2 \le g(\cdot) \le \| \cdot \|^2/2$ for some constant $0 < c < 1$. We may actually choose g as the square of a certain equivalent norm on H; consult [11] by Borwein and Fabian for a paper on this topic. Then, $\| \cdot \|^2/2 \le g^* \le \| \cdot \|^2/(2c)$, g^* is strictly convex, and the supremum for the bi-conjugate

$$g^{**}(x) = \sup_{y \in H} \big(\langle x, y \rangle - g^*(y) \big) \tag{50}$$

is uniquely attained for each $x \in H$. Define, for a fixed $t_0 > 0$,

$$f(x) = \Big(g^*(x) - \|x\|^2/2 \Big)/t_0, \qquad x \in H.$$

Then, f is real-valued, $f \ge 0$ and f_{t_0} is everywhere Gâteaux differentiable yet fails to be Fréchet differentiable throughout H. To see this, we need only observe that

$$t_0 f_{t_0}(x) = \frac{1}{2}\|x\|^2 - (t_0 f + \| \cdot \|^2/2)^*(x) = \frac{1}{2}\|x\|^2 - g^{**}(x) = \frac{1}{2}\|x\|^2 - g(x)$$

since $g^{**} = g$. For each $x \in H$, the infimum for $f_{t_0}(x)$ is uniquely attained because it boils down to the supremum (50). To summarize,

$$\Sigma_s \bigcap (\{t_0\} \times H) = \emptyset \quad \text{whereas} \quad \Sigma \bigcap (\{t_0\} \times H) \ne \emptyset.$$

By Theorem 10, $S(t, x) = f_t(x)$ fails to be jointly Gâteaux differentiable at some point of $\{t_0\} \times H$.

Remark 7 By Theorem 11, (i) $\Leftrightarrow$ (ii) $\wedge$ (iii) in the case $f = \mathfrak{I}_E$. This equivalence fails in general as Example 3 shows.

Example 4 (Fitzpatrick [23]) In the Hilbert space $H = \ell^2$ let

$$E = \{e_1\} \cup \{r_n e_n \colon n = 2, 3, \ldots\}$$

where $1 < r_n \downarrow 1$ as $n \to \infty$. Then $P_E(0) = \{e_1\}$, a singleton, yet d_E^2 is not Gâteaux differentiable at 0. Indeed,

$$d_E^2(\lambda e_1) - d_E^2(0) = \min\left\{(\lambda - 1)^2, \inf_{n \geq 2}(\lambda^2 + r_n^2)\right\} - 1 = \min\{\lambda^2 - 2\lambda, \lambda^2\}$$

and we conclude that

$$\lim_{\lambda \to 0} \frac{d_E^2(\lambda e_1) - d_E^2(0)}{\lambda} \quad \text{does not exist.}$$

Example 5 (Fitzpatrick [23]) In $H = \ell^2$, let

$$E = \{r_n e_n : n = 1, 2, \ldots\}$$

where $r_n = (n+1)/n$. Then $P_E(0) = \emptyset$ yet d_E^2 is Gâteaux differentiable at 0 with $\nabla d_E^2(0) = 0$. In this case, setting $I(x) = -2A(x)$, $x = (x_k)_{k=1}^\infty \in \ell^2$,

$$I(x) = \inf_{k \geq 1}(r_k^2 - 2r_k x_k)$$

is a concave function with its global maximum at $x = 0$ on account of $I(0) = 1$ and

$$I(x) \leq \liminf_{k \to \infty}(r_k^2 - 2r_k x_k) = 1 \quad \text{for all } x \in \ell^2,$$

owing to $1 < r_k \to 1$ and $x_k \to 0$ as $k \to \infty$. In particular, if its exists, $\nabla I(0)$ must be 0. The confirmation of $\nabla d_E^2(0) = 0$ amounts to demonstrating that the directional derivative

$$\lim_{\lambda \to 0} \frac{I(\lambda v) - I(0)}{\lambda} = \lim_{\lambda \to 0} \frac{I(\lambda v) - 1}{\lambda} = 0 \tag{51}$$

in any direction $v \in \ell^2$. To this end, let $w_k = v_k$ for all $1 \leq k \leq N$ and $w_k = 0$ for all $k > N$. The truncated sequence $w = w^N$ satisfies $I(\lambda w) = 1$ for all sufficiently small $|\lambda|$. Indeed,

$$\inf_{k > N}(r_k^2 - 2\lambda w_k r_k) = \inf_{k > N} r_k^2 = 1$$

and

$$\min_{1 \leq k \leq N}(r_k^2 - 2\lambda w_k r_k) > 1$$

for all λ in a neighborhood of 0 (by continuity since the left-hand side is equal to $r_N^2 > 1$ when $\lambda = 0$). Returning to v we next find that

$$\left|\frac{I(\lambda v) - I(0)}{\lambda}\right| \leq \left|\frac{I(\lambda v) - I(\lambda w)}{\lambda}\right| + \left|\frac{I(\lambda w) - I(0)}{\lambda}\right| \leq C\|v - w\| + 0$$

when $0 < |\lambda|$ is small enough, whence $\limsup_{\lambda \to 0} |(I(\lambda v) - I(0))/\lambda| \le C\|v - w\|$. Finally, (51) follows when $N \to \infty$.

We close the paper by giving four examples of intrinsic characteristics.

Example 6 Let E be the complement of the open first quadrant in $H = \mathbb{R}^2$ and hence $\Sigma = \{(x_1, x_2) \in H : x_2 = x_1 > 0\}$. The intrinsic characteristic subject to the initial condition $\boldsymbol{x}(t_0) = x_0 = (\xi_0, \xi_0) \in \Sigma$ is given by

$$\boldsymbol{x}(t) = \frac{2t}{t_0 + t} x_0, \qquad t_0 \le t < \infty, \tag{52}$$

Indeed, it is a singular arc whose points $x = (\xi, \xi)$ are calculated by maximizing

$$\mathbb{R} \ni \xi \mapsto \frac{1}{2t}\xi^2 - \frac{1}{2(t - t_0)}2(\xi_0 - \xi)^2, \qquad t_0 < t < \infty. \tag{53}$$

We notice that $\boldsymbol{x}(\cdot)$ traces out only $[1, 2)x_0 \subset \Sigma$ and that $\boldsymbol{x}(\infty) = 2x_0$. (By means of the recursive approach explained in Remark 2 a singular curve $\boldsymbol{\xi}(t)$ can be obtained which traces out the ray $[1, \infty)x_0 \subset \Sigma$. Such a singular arc can also be constructed by iterating the step sketched in Remark 3.) By contrast, the unique generalized characteristic emanating from x_0 at time t_0 is

$$X(t) = \sqrt{\frac{t}{t_0}} x_0, \qquad t_0 \le t < \infty. \tag{54}$$

Indeed, if $t > 0$ and $x = (\xi, \xi)$, then the norm minimal element of $d^+ d_E^2(x)/2 = \mathrm{co}\{(\xi, 0), (0, \xi)\}$ is $\frac{1}{2}(\xi, \xi)$ and, thus, $v^\circ(t, x) = (2t)^{-1}x$. Solving $d\xi/dt = (2t)^{-1}\xi$ with initial condition $\xi(t_0) = \xi_0$ yields (54). Obviously, $\boldsymbol{x}$ and X are distinct. Furthermore, $J^{t_1 \to t_2} \circ J^{t_0 \to t_1}(x_0) \neq J^{t_0 \to t_2}(x_0)$ when $t_0 < t_1 < t_2$.

Example 7 Again in $H = \mathbb{R}^2$, let E be the L-shaped set defined for some $a > 0$ as

$$E = [0, 2a] \times \{0\} \bigcup \{0\} \times [0, 2a],$$

whose singular set is $\Sigma = \{(x_1, x_2) \in H : x_2 = x_1 > 0\}$. We set $z_0 = (a, a)$. For the initial point $x_0 = (\xi_0, \xi_0) \in \Sigma$ we distinguish between three cases:

(a) $x_0 \in \mathrm{co}\, E$, i.e., $0 < \xi_0 \le a$. In this case, $\boldsymbol{x}(t)$ is given by (52) and its limit is $\boldsymbol{x}(\infty) = 2x_0$. We observe that $\boldsymbol{x}(\infty)$ remains in $\mathrm{co}\, E$ only if $\xi_0 \le a/2$.

(b) $x_0 \notin \mathrm{co}\, E$ and $a < \xi_0 < 2a$. The intrinsic characteristic $\boldsymbol{x}(t)$ agrees with (52) until it reaches the point $2z_0$ at time $t_1 = at_0/(\xi_0 - a)$. After that, when $t \ge t_1$, $\boldsymbol{x}(t)$ coincides with (55).

(c) $x_0 \notin \mathrm{co}\, E$ and $\xi_0 \ge 2a$. Similarly as is in (53), the singular arc is obtained by maximizing

$$\xi \mapsto \frac{1}{2t}\left((\xi - 2a)^2 + \xi^2\right) - \frac{1}{2(t - t_0)}2(\xi_0 - \xi)^2, \qquad t_0 < t < \infty.$$

A calculation results in

$$\mathbf{x}(t) = x_0 + (t - t_0)v_0, \qquad t_0 \leq t < \infty, \tag{55}$$

where the constant velocity is $v_0 = (x_0 - z_0)/t_0$.

Clearly, the singular arcs of (b) and (c) are unbounded which is consistent with Theorem 9. In case (c), the arc has constant nonzero velocity.

Example 8 Let $f(x) = \exp(-\|x\|^2/2)$. Then it can be checked that $\Sigma = \Sigma_s = (1, \infty) \times \{0\}$. If $t_0 > 0$ and $x_0 = 0$ then $(t, \mathbf{x}(t)) = (t, 0)$ for all $t \in [t_0, \infty)$. Assuming $0 < t_0 < 1$, this shows that $\mathbf{x}(t)$ can be constant even though $(t, \mathbf{x}(t))$ starts off as nonsingular and later becomes singular (when $t > 1$). If we extend each intrinsic characteristic to $[0, \infty)$, then the resulting family of curves can be parameterized by their initial points $\mathbf{x}(0) = y$. This procedure results in the general form

$$\mathbf{x}_y(t) = \begin{cases} (1 - te^{-\|y\|^2/2})y & \text{if } 0 \leq t < e^{\|y\|^2/2}, \\ 0 & \text{if } t \geq e^{\|y\|^2/2}. \end{cases}$$

Indeed, $(t, \mathbf{x}_y(t))$ stays nonsingular with constant velocity $df(y) = -e^{-\|y\|^2/2}y$ when $0 \leq t < e^{\|y\|^2/2}$ until entering Σ at time $t = e^{\|y\|^2/2}$. After that it will remain in $\Sigma = \Sigma_s$, hence, $\mathbf{x}_y(t) = 0$ when $t \geq e^{\|y\|^2/2}$. Furthermore, $\mathbf{x}_y(t)$ coincides in this case with the unique generalized characteristic for any $y \in H$ although $\mathbf{x}_y(t)$ switches from nonsingular to singular and $\dot{\mathbf{x}}_y^+(t)$ is not a constant for any $y \neq 0$.

Example 9 Let f be the indicator function of the following union of two half-spaces

$$E = \{x \in H : \langle x, z \rangle \geq 1 \text{ or } \langle x, z \rangle \leq -1\}$$

for a certain fixed $z \in H$ with $\|z\| = 1$. Then

$$d_E^2(x) = \min\{(\langle x, z \rangle - 1)^2, (\langle x, z \rangle + 1)^2\} \quad \text{when} -1 < \langle x, z \rangle < 1.$$

Note, first, that d_E^2 fails to be Gâteaux differentiable on the orthogonal complement $\{z\}^\perp$ on account of

$$(\langle x, z \rangle - 1)^2 = (\langle x, z \rangle + 1)^2 \quad \Leftrightarrow \quad \langle x, z \rangle = 0$$

and, secondly, that $d_E^2 = 1$ on $\{z\}^\perp$ but $d_E^2 < 1$ elsewhere. Hence, if $x_0 \in \{z\}^\perp$, then x_0 is a critical point and $\mathbf{x}(t) = x_0$ for all $t \geq t_0$. Still, assuming $\dim H > 1$, there emanates from x_0 a variety of nonconstant singular Lipschitz arcs in the orthogonal complement of $\{z\}$. In fact, every curve X with $X(t) \in \{z\}^\perp$ for all t is a singular curve.

Declarations

Conflict of interest The author declares no conflict of interest.

References

1. Albano, P., Cannarsa, P.: Singularities of semiconcave functions in Banach spaces. Stochastic analysis, control, optimization and applications, 171–190, Systems Control Found. Appl., Birkhäuser Boston, Boston, MA (1999)
2. Albano, P.: Propagation of singularities for solutions of Hamilton-Jacobi equations. J. Math. Anal. Appl. **411**, 684–687 (2014)
3. Albano, P.: Global propagation of singularities for solutions of Hamilton-Jacobi equations. J. Math. Anal. Appl. **444**, 1462–1478 (2016)
4. Albano, P., Cannarsa, P.: Propagation of singularities for solutions of nonlinear first order partial differential equations. Arch. Ration. Mech. Anal. **162**, 1–23 (2002)
5. Albano, P., Cannarsa, P., Nguyen, K., Sinestrari, C.: Singular gradient flow of the distance function and homotopy equivalence. Math. Ann. **356**, 23–43 (2013)
6. Alberti, G., Ambrosio, L., Cannarsa, P.: On the singularities of convex functions. Manuscripta Math. **76**, 421–435 (1992)
7. Asplund, E.: Fréchet differentiability of convex functions. Acta Math. **121**, 31–47 (1968)
8. Asplund, E.: Čebyšev sets in Hilbert space. Trans. Am. Math. Soc. **144**, 235–240 (1969)
9. Balaganskii, V., Vlasov, L.: The problem of the convexity of Chebyshev sets. Russian Math. Surv. **51**, 1127–1190 (1996)
10. Berens, H.: Ein Problem über die beste Approximation in Hilberträumen. Functional analysis and approximation (Oberwolfach, 1980), pp. 247-254, Internat. Ser. Numer. Math., 60, Birkhäuser, Basel-Boston, Mass (1981)
11. Borwein, J.M., Fabián, M.: On convex functions having points of Gateaux differentiability which are not points of Fréchet differentiability. Canad. J. Math. **45**, 1121–1134 (1993)
12. Borwein, J.M., Vanderwerff, J.D.: Convex functions: constructions, characterizations and counterexamples. Encyclopedia of Mathematics and its Applications, 109. Cambridge University Press, Cambridge, (2010)
13. Cannarsa, P., Cheng, W.: Generalized characteristics and Lax-Oleinik operators: global theory. Calc. Var. Partial Differential Equations 56, no. 5, Paper No. 125, 31 pp (2017)
14. Cannarsa, P., Cheng, W.: On and beyond propagation of singularities of viscosity solutions. In: Proceedings of the International Consortium of Chinese Mathematicians 2017, 141-157, Int. Press, Boston, MA (2020)
15. Cannarsa, P., Cheng, W.: Singularities of solutions of Hamilton-Jacobi equations. Milan J. Math. **89**, 187–215 (2021)

16. Cannarsa, P., Sinestrari, S.: Semiconcave functions, Hamilton–Jacobi equations, and optimal control. In: Progress in Nonlinear Differential Equations and their Applications, 58, Birkhäuser Boston, MA (2004)
17. Cannarsa, P., Yu, Y.: Singular dynamics for semiconcave functions. J. Eur. Math. Soc. (JEMS) **11**, 999–1024 (2009)
18. Cannarsa, P., Mennucci, A., Sinestrari, C.: Regularity results for solutions of a class of Hamilton-Jacobi equations. Arch. Rational Mech. Anal. **140**, 197–223 (1997)
19. Cannarsa, P., Mazzola, M., Sinestrari, C.: Global propagation of singularities for time dependent Hamilton-Jacobi equations. Discrete Contin. Dyn. Syst. **35**, 4225–4239 (2015)
20. Cannarsa, P., Cheng, W., Fathi, A.: On the topology of the set of singularities of a solution to the Hamilton-Jacobi equation. C. R. Math. Acad. Sci. Paris **355**, 176–180 (2017)
21. Cannarsa, P., Cheng, W., Fathi, A.: Singularities of solutions of time dependent Hamilton-Jacobi equations. Applications to Riemannian geometry. Publ. Math. Inst. Hautes Études Sci. **133**, 327-366 (2021)
22. Dafermos, C.M.: Generalized characteristics and the structure of solutions of hyperbolic conservation laws. Indiana Univ. Math. J. **26**, 1097–1119 (1977)
23. Fitzpatrick, S.: Metric projections and the differentiability of distance functions. Bull. Austral. Math. Soc. **22**, 291–312 (1980)
24. Frerking, J., Westphal, U.: On a property of metric projections onto closed subsets of Hilbert spaces. Proc. Am. Math. Soc. **105**, 644–651 (1989)
25. Godini, G.: On a problem of H. Berens. Math. Ann. **263**, 279–281 (1983)
26. Jourani, A., Thibault, L., Zagrodny, D.: Differential properties of the Moreau envelope. J. Funct. Anal. **266**, 1185–1237 (2014)
27. Khanin K., Sobolevski, A.: Particle dynamics inside shocks in Hamilton–Jacobi equations. Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. **368**, 1579–1593 (2010)
28. Khanin, K., Sobolevski, A.: On dynamics of Lagrangian trajectories for Hamilton-Jacobi equations. Arch. Ration. Mech. Anal. **219**, 861–885 (2016)
29. Klee, V.: Convexity of Chevyshev sets. Math. Ann. **142**, 292-304 (1960/61)
30. Klee, V.: Dispersed Chebyshev sets and coverings by balls. Math. Ann. **257**, 251–260 (1981)
31. Kruger, A. Ya.: On Fréchet subdifferentials. Optimization and related topics, 3. J. Math. Sci. (N.Y.) **116**, 3325-3358 (2003)
32. Lasry, J.-M., Lions, P.-L.: A remark on regularization in Hilbert spaces. Israel J. Math. **55**, 257–266 (1986)
33. Lieutier, A.: Any open bounded subset of $\mathbb{R}^n$ has the same homotopy type as its medial axis. Comput. Aided Des. **36**, 1029–1046 (2004)
34. Miura, T., Tanaka, M.: Strongly Lipschitz structure of the singular set of distance functions (2022). Preprint available at arXiv:2204.10449
35. Moreau, J.J.: Fonctions convexes duales et points proximaux dans un espace hilbertien. C. R. Acad. Sci. Paris **255**, 2897–2899 (1962)
36. Moreau, J.J.: Proximité et dualité dans un espace hilbertien. Bull. Soc. Math. France **93**, 273–299 (1965)
37. Penot, J.-P., Volle, M.: Explicit solutions to Hamilton-Jacobi equations under mild continuity and convexity assumptions. J. Nonlinear Convex Anal. **1**, 177–199 (2000)
38. Penot, J.-P., Volle, M.: Duality methods for the study of Hamilton–Jacobi equations. In: CSVAA 2004-control set-valued analysis and applications, 96-142, ESAIM Proc., 17, EDP Sci., Les Ulis (2007)
39. Santilli, M.: Distance functions with dense singular sets. Comm. Partial Diffe. Equ. **46**, 1319–1325 (2021)
40. Šmulian, V. L.: Sur la structure de la sphère unitaire dans l'espace de Banach. Rec. Math. [Mat. Sbornik] N.S. 9 (51), 545–561 (1941)
41. Strömberg, T.: A new proof of indefinite propagation of singularities for a Hamilton-Jacobi equation. J. Evol. Equ. **16**, 895–903 (2016)
42. Veselý, L.: A connectedness property of maximal monotone operators and its application to approximation theory. Proc. Am. Math. Soc. **115**, 663–667 (1992)
43. Yu, Y.: A simple proof of the propagation of singularities for solutions of Hamilton–Jacobi equations. Ann. Sci. Norm. Super. Pisa Cl. Sci. **5**, 439–444 (2006)