Journal of Convex Analysis
Volume 31 (2024), No. 3, 827–846

Singularities of Fitzpatrick and Convex Functions

Dmitry Kramkov*
Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, U.S.A.
kramkov@cmu.edu

Mihai Sîrbu†
Department of Mathematics, The University of Texas at Austin, U.S.A.
sirbu@math.utexas.edu

Received: December 17, 2022
Accepted: August 23, 2023

In a pseudo-Euclidean space with scalar product $S(\cdot,\cdot)$, we show that the singularities of projections on S-monotone sets and of the associated Fitzpatrick functions are covered by countable $c-c$ surfaces having positive normal vectors with respect to the S-product. By L. Zajíček [*On the differentiation of convex functions in finite and infinite dimensional spaces*, Czechoslovak Math. J. 29/104 (1979) 340–348], the singularities of a convex function f can be covered by a countable collection of $c-c$ surfaces. We show that the normal vectors to these surfaces are restricted to the cone generated by $F-F$, where $F := \operatorname{cl}\operatorname{range}\nabla f$, the closure of the range of the gradient of f.

Keywords: Convexity, subdifferential, Fitzpatrick function, projection, pseudo-Euclidean space, normal vector, singularity.

2020 Mathematics Subject Classification: 26B25, 26B05, 47H05, 52A20.

1. Introduction

A locally Lipschitz function $f = f(x)$ on $\mathbb{R}^d$ is differentiable almost everywhere, according to the Rademacher's Theorem. The set of its singularities

$$\Sigma(f) := \left\{x \in \mathbb{R}^d \middle| \, f \text{ is not differentiable at } x\right\}$$

can be quite irregular. For instance, for $d = 1$, Zahorski [23] (see Fowler and Preiss [11] for a simple proof) shows that *any* $G_{\delta\sigma}$ set (countable union of countable intersections of open sets) of Lebesgue measure zero is the singular set of *some* Lipschitz function.

By Zajíček [24], the $\Sigma(f)$ of a convex function $f = f(x)$ on $\mathbb{R}^d$ has $c-c$ structure: it can be covered by a countable collection of the graphs of the differences of two convex functions of dimension $d-1$. Short proofs of the Zajíček theorem can be found in Benyamini and Lindenstrauss [5, Theorem 4.20, p. 93], Thibault [22, Theorem 12.22, p. 1147], and Hajłasz [13]. Alberti [2] shows that, except for sets with zero $\mathcal{H}^{d-1}$ measure, the Hausdorff measure of dimension $d-1$, the covering can be achieved

* The author also has a research position at the University of Oxford.

† The research of this author was supported in part by the National Science Foundation under Grant DMS 1908903.

ISSN 0944-6532 / $ 2.50

with smooth surfaces. These results yield sharp conditions for the existence and uniqueness of optimal maps in $\mathcal{L}_2$ optimal transport, see Brenier [6] and Ambrosio and Gigli [3, Theorem 2.26, p. 18].

Let E be a closed subset of $\mathbb{R}^d$ and

$$d_E(x) := \inf_{y \in E} |x-y|, \quad P_E(x) := \{y \in E | \; |x-y| = d_E(x)\},$$

be the Euclidean distance to E and the projection on E, respectively. Erdös [8] proves that the singular set

$$\Sigma(P_E) := \left\{x \in \mathbb{R}^d | \; P_E(x) \text{ contains at least two points}\right\}$$

can be covered by countable sets with finite $\mathcal{H}^{d-1}$ measure. Hajłasz [13] uses [24] and the observation of Asplund [4] that the function

$$\psi_E(x) := \frac{1}{2}|x|^2 - \frac{1}{2}d_E^2(x), \quad x \in \mathbb{R}^d,$$

is convex, to conclude that $\Sigma(P_E)$ has the $c-c$ structure. Albano and Cannarsa [1] obtain a lower bound on the size of the set $\Sigma(d_E)$, where d_E is not differentiable.

Let S be a $d \times d$ invertible symmetric matrix with $m \in \{0, 1, \dots, d\}$ positive eigenvalues. Motivated by applications to backward martingale transport in [15], [16], and [17], we investigate in this paper the singularities of projections on monotone sets in the pseudo-Euclidean space with the scalar product

$$S(x,y) := \langle x, Sy \rangle = \sum_{i,j=1}^{d} x^i S^{ij} y^j, \quad x, y \in \mathbb{R}^d.$$

Let $G \subset \mathbb{R}^d$ be an S-monotone or S-positive set:

$$S(x-y, x-y) \geq 0, \quad x, y \in G.$$

For every $x \in \mathbb{R}^d$, the scalar square to G and the projection on G are given by

$$\begin{aligned} \phi_G(x) &:= \inf_{y \in G} S(x-y, x-y), \\ P_G(x) &:= \arg\min_{y \in G} S(x-y, x-y). \end{aligned}$$

Note that S-monotonicity is equivalent to $x \in G \implies x \in P_G(x)$ and that

$$\psi_G(x) := \frac{1}{2}S(x,x) - \frac{1}{2}\phi_G(x), \quad x \in \mathbb{R}^d,$$

is the Fitzpatrick function studied in [10], [21], and [18].

The singularities of the projection P_G can be classified as

$$\begin{aligned} \Sigma(P_G) &:= \left\{x \in \mathbb{R}^d | \; P_G(x) \text{ contains at least two points}\right\} \\ &= \Sigma_0(P_G) \cup \Sigma_1(P_G), \\ \Sigma_0(P_G) &:= \{x \in \Sigma(P_G) | \; S(y_1 - y_2, y_1 - y_2) = 0 \text{ for } \textit{all } y_1, y_2 \in P_G(x)\}, \\ \Sigma_1(P_G) &:= \{x \in \Sigma(P_G) | \; S(y_1 - y_2, y_1 - y_2) > 0 \text{ for } \textit{some } y_1, y_2 \in P_G(x)\}. \end{aligned}$$

By Theorem 4.6, $\Sigma_1(P_G)$ is contained in a countable union of $c-c$ surfaces having strictly positive normal vectors in the S-space. The structure of the zero-order singularities is described in Theorems 4.8 and 4.7. If $m = 1$, then $\Sigma_0(P_G)$ is covered by a countable number of hyperplanes having isotropic normal vectors in the S-space. If $m \geq 2$, then $\Sigma_0(P_G)$ is covered by a countable family of $c-c$ surfaces whose normal vectors are positive and almost isotropic in the S-space. These results yield sharp conditions for the existence and uniqueness of backward martingale maps in [16].

Using similar tools, in Theorem 3.1, we improve the $c-c$ description of singularities of general convex functions $f = f(x)$ from Zajíček [24] by showing that the covering surfaces have normal vectors belonging to the cone generated by $F - F$, where $F := \operatorname{cl} \operatorname{range} \nabla f$, the closure of the range of the gradient of f.

2. Parametrization of singularities

We say that a function $g = g(x)$ on $\mathbb{R}^d$ has *linear growth* if there is a constant $K = K(g) > 0$ such that

$$|g(x)| \leq K(1 + |x|), \quad x \in \mathbb{R}^d.$$

We write $\operatorname{dom} \nabla g$ for the set of points where g is differentiable.

Let $j \in \{1, \ldots, d\}$. We denote by $\mathcal{C}^j$ the collection of compact sets C in $\mathbb{R}^d$ such that

$$y^j = 1, \quad y \in C.$$

Any compact set $C \subset \left\{x \in \mathbb{R}^d | \; x^j > 0\right\}$ can be rescaled as

$$\theta^j(C) := \left\{\frac{y}{y^j} \mid y \in C\right\} \in \mathcal{C}^j.$$

For $x \in \mathbb{R}^d$, we denote by x^{-j} its sub-vector without the jth coordinate:

$$x^{-j} := \left(x^1, \ldots, x^{j-1}, x^{j+1}, \ldots, x^d\right) \in \mathbb{R}^{d-1}.$$

For $C \in \mathcal{C}^j$, we write $\mathcal{H}^j_C$ for the family of functions $h = h(x)$ on $\mathbb{R}^d$ having the decomposition:

$$h(x) = x^j + g_1(x^{-j}) - g_2(x^{-j}), \quad x \in \mathbb{R}^d, \tag{1}$$

where the functions g_1 and g_2 on $\mathbb{R}^{d-1}$ are convex, have linear growth, and

$$\nabla h(x) \in C, \quad x^{-j} \in \operatorname{dom} \nabla g_1 \cap \operatorname{dom} \nabla g_2. \tag{2}$$

The latter property has a clear geometric interpretation. Let H be a closed set in $\mathbb{R}^d$ and $x \in H$. Following [20, Definition 6.3 on page 199], we call a vector $w \in \mathbb{R}^d$ *regular normal to H at x* if

$$\limsup_{\substack{H \ni y \to x \\ y \neq x}} \frac{\langle w, y - x\rangle}{|y - x|} \leq 0.$$

A vector $w \in \mathbb{R}^d$ is called *normal to H at x* if there exist $x_n \in H$ and a regular normal vector w_n to H at x_n such that $x_n \to x$ and $w_n \to w$.

For a set B in $\mathbb{R}^d$, we denote by $\operatorname{conv} B$ its convex hull.

Lemma 2.1. *Let $j \in \{1, \dots, d\}$, $C \in \mathcal{C}^j$, h be given by (1) for convex functions g_1 and g_2 with linear growth, and H be the zero-level set of h:*

$$H := \left\{ x \in \mathbb{R}^d \middle| \, h(x) = 0 \right\}.$$

Then $h \in \mathcal{H}^j_C$, that is, (2) holds, if and only if for every $x \in H$, there exists a normal vector $w \in C$ to H at x.

Proof. We can assume that $j = 1$. Denote $f := g_1 - g_2$, so that $h(x) = x^1 + f(x^{-1})$. We have that $\operatorname{dom} \nabla h = \mathbb{R} \times \operatorname{dom} \nabla f$, $\nabla h(x) = (1, \nabla f(x^{-1}))$, and

$$H = \left\{ (-f(u), u) \middle| \, u \in \mathbb{R}^{d-1} \right\}.$$

Denote also $U := \operatorname{dom} \nabla g_1 \cap \operatorname{dom} \nabla g_2 \subset \operatorname{dom} \nabla f$.

$\Longrightarrow$: Clearly, $\nabla h(x)$ is a regular normal vector to H at $x \in \operatorname{dom} \nabla h \cap H$. As U is dense in $\mathbb{R}^{d-1}$, the result holds by standard compactness arguments.

$\Longleftarrow$: Let $u \in U$, $v \in \mathbb{R}^{d-1}$, and $w = (1, v) \in C$ be a normal vector to H at $x = (-f(u), u)$. Take a sequence (x_n, w_n), $n \geq 1$, that converges to (x, w) and where w_n is a regular normal vector to H at $x_n \in H$ with $w^1_n = 1$. Such a sequence exists by the definition of a normal vector. We can represent $x_n = (-f(u_n), u_n)$ for $u_n \in \mathbb{R}^{d-1}$ and $w_n = (1, v_n)$ for $v_n \in \mathbb{R}^{d-1}$.

We claim that v_n belongs to the Clarke gradient of f at u_n:

$$v_n \in \overline{\partial} f(u_n) := \operatorname{conv} \left\{ \lim_m \nabla f(r_m) \middle| \, \operatorname{dom} \nabla f \ni r_m \to u_n \right\}.$$

The continuity of ∇f at $u \in U$ then yields that $v_n \to \nabla f(u)$. Hence, $\nabla f(u) = v$, implying that $\nabla h(x) = (1, \nabla f(u)) = w \in C$.

In order to prove the claim, we write the definition of $w_n = (1, v_n)$ being regular normal to H at $x_n = (-f(u_n), u_n)$ as

$$\limsup_{\substack{r \to u_n \\ r \neq u_n}} \frac{-\left(f(r) - f(u_n)\right) + \langle v_n, r - u_n \rangle}{|f(r) - f(u_n)| + |r - u_n|} \leq 0.$$

Using the Lipschitz property of f, we obtain that

$$\langle v_n, s \rangle \leq \limsup_{\delta \downarrow 0} \frac{f(u_n + \delta s) - f(u_n)}{\delta}, \quad s \in \mathbb{R}^{d-1}.$$

By [7, Corollary 1.10], we have that $v_n \in \overline{\partial} f(u_n)$, as claimed. $\square$

We recall that the *subdifferential* ∂f of a closed convex function $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ is defined as

$$\partial f(x) := \left\{ y \in \mathbb{R}^d \middle| \, \langle z, y \rangle \leq f(x + z) - f(x), \, z \in \mathbb{R}^d \right\}, \quad x \in \mathbb{R}^d.$$

Clearly, $\operatorname{dom} \partial f := \left\{ x \in \mathbb{R}^d \middle| \, \partial f(x) \neq \emptyset \right\} \subset \operatorname{dom} f := \left\{ x \in \mathbb{R}^d \middle| \, f(x) < \infty \right\}$.

The following theorem is our main technical tool for the study of singularities of convex and Fitzpatrick functions in Sections 3 and 4.

Theorem 2.2. *Let $f: \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be a closed convex function, $j \in \{1, \dots, d\}$, and C_1 and C_2 be compact sets in $\mathbb{R}^d$ such that*

$$y^j > 0, \quad y \in C_2 - C_1.$$

There exists a function $h \in \mathcal{H}^j_{\theta^j(C_2 - C_1)}$ such that

$$\begin{aligned} \Sigma_{C_1,C_2}(\partial f) &:= \{x \in \operatorname{dom} \partial f(x) |\ \partial f(x) \cap C_i \neq \emptyset,\ i = 1,2\} \\ &\subset \left\{x \in \mathbb{R}^d |\ h(x) = 0\right\}. \end{aligned}$$

The proof of the theorem relies on Lemma 2.3. For a closed set $A \subset \mathbb{R}^d$, we denote

$$f_A(x) := \sup_{y \in A} \left(\langle x, y\rangle - f^*(y)\right), \quad x \in \mathbb{R}^d, \tag{3}$$

where f^* is the convex conjugate of f:

$$f^*(y) := \sup_{x \in \mathbb{R}^d} \left(\langle x, y\rangle - f(x)\right) \in \mathbb{R} \cup \{+\infty\}, \quad y \in \mathbb{R}^d.$$

Obviously, f_A is a closed convex function taking values in $\mathbb{R} \cup \{+\infty\}$ if and only if

$$A \cap \operatorname{dom} f^* = \{x \in A |\ f^*(x) < \infty\} \neq \emptyset;$$

otherwise, $f_A = -\infty$. We recall that

$$f(x) = \langle x, y\rangle - f^*(y) \iff y \in \partial f(x) \iff x \in \partial f^*(y). \tag{4}$$

Lemma 2.3. *Let $f: \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be a closed convex function and C be a compact set in $\mathbb{R}^d$ such that $C \cap \operatorname{dom} f^* \neq \emptyset$. Then f_C has linear growth and for every $x \in \mathbb{R}^d$,*

$$\partial f_C(x) \cap C = \operatorname{Arg}_C(x) := \arg\max_{y \in C} \left(\langle x, y\rangle - f^*(y)\right) \neq \emptyset,$$

$$\partial f_C(x) = \operatorname{conv}\left(\partial f_C(x) \cap C\right),$$

$$\partial f(x) \cap C \neq \emptyset \iff f(x) = f_C(x) \iff \partial f(x) \cap C = \partial f_C(x) \cap C.$$

In particular, f_C is differentiable at x if and only if $\partial f_C(x) \cap C$ is a singleton, in which case

$$\partial f_C(x) = \{\nabla f_C(x)\} \in C.$$

Proof. Since C is a compact set, f^* is a closed convex function, and $C \cap \operatorname{dom} f^* \neq \emptyset$, we have that

$$\sup_{y \in C} |y| < \infty, \quad \inf_{x \in C} f^*(x) > -\infty,$$

and that $\operatorname{Arg}_C(x)$ is a non-empty compact. Let $y_0 \in C \cap \operatorname{dom} f^*$. From the definition of f_C we deduce that

$$-|x|\,|y_0| - f^*(y_0) \leq f_C(x) \leq |x| \sup_{y \in C} |y| - \inf_{y \in C} f^*(y), \quad x \in \mathbb{R}^d.$$

It follows that f_C has linear growth.

The function $h(x,y) := \langle x,y\rangle - f^*(y) \in \mathbb{R}\cup\{-\infty\}, \quad x,y\in\mathbb{R}^d$,
is linear in x and concave and upper semi-continuous in y. Fix $x \in \mathbb{R}^d$. We can choose K large enough such that for

$$E := \{z \in C|\ f^*(z) \le K\},$$

we have that $f_C = f_E$ in a neighborhood of x and

$$\mathrm{Arg}_C(x) = \mathrm{Arg}_E(x) := \arg\max_{y\in E}\left(\langle x,y\rangle - f^*(y)\right).$$

Since E is compact and the function $h(\cdot,y)$ is finite for $y \in E$, the classical envelope theorem [14, Theorem 4.4.2, p. 189] yields that

$$\begin{aligned}\partial f_C(x) &= \partial f_E(x) = \partial \max_{y\in E} h(x,y) = \mathrm{conv} \bigcup_{y\in \mathrm{Arg}_E(x)} \partial_x h(x,y)\\ &= \mathrm{conv}\,\mathrm{Arg}_E(x) = \mathrm{conv}\,\mathrm{Arg}_C(x).\end{aligned}$$

From the concavity of $h(x,\cdot)$ we deduce that

$$\partial f_C(x)\cap C = (\mathrm{conv}\,\mathrm{Arg}_C(x))\cap C = \mathrm{Arg}_C(x).$$

Clearly, $f_C \le f$. If $y \in \partial f(x)\cap C$, then $f(x) = \langle x,y\rangle - f^*(y) \le f_C(x)$. Hence, $f_C(x) = f(x)$ and $y \in \mathrm{Arg}_C(x) \subset \partial f_C(x)$.

Conversely, let $f_C(x) = f(x)$. For every $y \in \partial f_C(x)\cap C$, we readily conclude that $f(x) = f_C(x) = \langle x,y\rangle - f^*(y)$, and then that $y \in \partial f(x)\cap C$.

Finally, being a convex function, f_C is differentiable at x if and and only if $\partial f_C(x)$ is a singleton. In this case, $\partial f_C(x) = \{\nabla f_C(x)\}$. □

Proof of Theorem 2.2. Hereafter, $i = 1, 2$. We assume that $\Sigma_{C_1,C_2}(\partial f) \neq \emptyset$ as otherwise, there is nothing to prove. This implies that $C_i \cap \mathrm{dom}\, f^* \neq \emptyset$. Let $C := C_1 \cup C_2$. Lemma 2.3 yields that

$$\Sigma_{C_1,C_2}(\partial f) = \Sigma_{C_1,C_2}(\partial f_C)\cap\left\{x\in\mathbb{R}^d|\ f(x) = f_C(x)\right\}. \tag{5}$$

To simplify notations we assume that $j = 1$. Denote by

$$a := \max_{y\in C_1} y^1 < \min_{y\in C_2} y^1 := b.$$

We write $x \in \mathbb{R}^d$ as (t,u), where $t\in\mathbb{R}$ and $u \in \mathbb{R}^{d-1}$, and define the saddle function

$$g(t,u) := \inf_{s\in\mathbb{R}}\left(f_C(s,u) - st\right), \quad a < t < b, u \in \mathbb{R}^{d-1}.$$

Select $y_i = (q_i, z_i) \in C_i \cap \mathrm{dom}\, f^*$. We have that

$$f_C(s,u) \ge \max_{i=1,2}\left(sq_i + \langle u, z_i\rangle - f^*(y_i)\right).$$

Since $q_1 \le a < b \le q_2$, it follows from the definition of g that

$$-\infty < g(t,u) \le f_C(0,u), \quad a < t < b, u\in\mathbb{R}^{d-1}.$$

By Lemma 2.3, f_C has linear growth. The classical results on saddle functions, Theorems 33.1 and 37.5 in [19], imply that

(i) For every $a < t < b$, the function $g(t, \cdot)$ is convex and has linear growth on $\mathbb{R}^{d-1}$.

(ii) For every $u \in \mathbb{R}^{d-1}$, the function $g(\cdot, u)$ is concave and finite on (a, b).

(iii) For any $a < t < b$, we have that $(t, v) \in \partial f_C(s, u)$ if and only if $v \in \partial_u g(t, u)$ and $-s \in \partial_t g(t, u)$. In this case, $g(t, u) = f_C(s, u) - st$.

We take $a < r_1 < r_2 < b$ and denote $g_i := g(r_i, \cdot)$. We verify the assertions of the theorem for the function

$$h(s, u) := s + \frac{1}{r_2 - r_1}\left(g_2(u) - g_1(u)\right), \quad s \in \mathbb{R}, u \in \mathbb{R}^d.$$

More precisely, we show that

$$\Sigma_{C_1,C_2}(\partial f_C) = \left\{x \in \mathbb{R}^d \middle|\, h(x) = 0\right\},$$

which together with (5) implies the result.

Let $x = (s, u)$ and $(t_i, v_i) \in \partial f_C(x) \cap C_i$. As $t_1 \le a < r_i < b \le t_2$, the convexity of subdifferentials yields that $(r_i, w_i) \in \partial f_C(x)$, where

$$w_i = \frac{t_2 - r_i}{t_2 - t_1} v_1 + \frac{r_i - t_1}{t_2 - t_1} v_2. \tag{6}$$

By (iii), $g_i(u) = f_C(x) - sr_i$ and then $h(x) = 0$.

Conversely, let $x = (s, u)$ be such that $h(x) = 0$ or, equivalently,

$$-s = \frac{g_2(u) - g_1(u)}{r_2 - r_1} = \frac{g(r_2, u) - g(r_1, u)}{r_2 - r_1}.$$

The mean-value theorem yields $r \in (r_1, r_2)$ such that $-s \in \partial_t g(r, u)$. Observe that we have $a < r < b$. Taking any $w \in \partial_u g(r, u)$, we deduce from condition (iii) that $y := (r, w) \in \partial f_C(x)$. As $\partial f_C = \operatorname{conv}(\partial f_C \cap C)$, the point y is a convex combination of some $y_i \in \partial f_C(x) \cap C_i$, $i = 1, 2$. In particular, $x \in \Sigma_{C_1,C_2}(\partial f_C)$.

Finally, let $x = (s, u)$ be such that $h(x) = 0$ and g_1 and g_2 are differentiable at u. As we have already shown, the gradients $w_i := \nabla g_i(u)$ are given by (6) for some $(t_i, v_i) \in \partial f_C(x) \cap C_i$. It follows that

$$\nabla h(x) = \left(1, \frac{w_2 - w_1}{r_2 - r_1}\right) = \left(1, \frac{v_2 - v_1}{t_2 - t_1}\right) \in \theta^1(C_2 - C_1).$$

Hence, $h \in \mathcal{H}^1_{\theta^1(C_2 - C_1)}$. □

3. Singular points of convex functions

For a multi-function $\Psi : \mathbb{R}^d \rightrightarrows \mathbb{R}^d$ taking values in closed subsets of $\mathbb{R}^d$, we denote its domain by

$$\operatorname{dom} \Psi := \left\{x \in \mathbb{R}^d \middle|\, \Psi(x) \neq \emptyset\right\}.$$

Given an index $j \in \{1, \dots, d\}$ and a closed set A in $\mathbb{R}^d$, the *singular* set of Ψ is defined as

$$\Sigma^j_A(\Psi) := \left\{x \in \operatorname{dom} \Psi \middle|\, \exists y_1, y_2 \in \Psi(x) \cap A \text{ with } y_1^j \neq y_2^j\right\}.$$

We also write $\Sigma^j(\Psi) = \Sigma^j_{\mathbb{R}^d}(\Psi)$ and $\Sigma(\Psi) = \cup_{j=1,\dots,d}\Sigma^j(\Psi)$.

Let $f:\ \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be a closed convex function such that its domain has a non-empty interior:

$$D := \operatorname{int}\operatorname{dom} f \neq \emptyset.$$

It is well-known that $\operatorname{dom}\nabla f := \{x \in D |\ \nabla f(x) \text{ exists}\}$ is dense in D and

$$D \setminus \operatorname{dom}\nabla f = \Sigma(\partial f) \cap D = \{x \in D |\ \partial f(x) \text{ is not a point}\}.$$

According to [24], this set of *interior* singularities can be covered by countable $c-c$ surfaces $H_n = \left\{x \in \mathbb{R}^d |\ h_n(x) = 0\right\}$, $n \geq 1$, where

$$h_n(x) = x^j + g_{n,1}(x^{-j}) - g_{n,2}(x^{-j}), \quad x \in \mathbb{R}^d,$$

for some $j \in \{1, \dots, d\}$ and finite convex functions $g_{n,1}$ and $g_{n,2}$ on $\mathbb{R}^{d-1}$.

Theorem 3.1 and Lemma 2.1 describe the *orientation* of the covering surface H_n by showing that, at any point, it has a normal vector w with $w^j = 1$, that belongs to the cone K generated by $F - F$, where

$$F := \operatorname{cl}\operatorname{range}\nabla f = \left\{\lim_n \nabla f(x_n) \mid x_n \in \operatorname{dom}\nabla f\right\}.$$

Of course, this information has some value only if K is distinctively smaller than $\mathbb{R}^d$. This is the case for Fitzpatrick functions in the pseudo-Euclidean space S, where according to Theorem 4.5, K contains only S-non-negative vectors. Proposition 3.5 provides the geometric interpretation of the range ∇f and Lemma 3.3 explains the special feature of its closure F.

It turns out that the same surfaces (H_n) also cover the singularities of the Clarke-type subdifferential $\overline{\partial} f$ defined on the *closure* of D:

$$\overline{\partial} f(x) := \operatorname{cl}\operatorname{conv}\left\{\lim_n \nabla f(x_n) |\ \operatorname{dom}\nabla f \ni x_n \to x\right\}, \quad x \in \operatorname{cl} D.$$

By Theorem 25.6 in [19],

$$\partial f(x) = \overline{\partial} f(x) + N_{\operatorname{cl} D}(x), \quad x \in \operatorname{cl} D,$$

where $N_A(x)$ denotes the normal cone to a closed convex set $A \subset \mathbb{R}^d$ at $x \in A$:

$$\begin{aligned} N_A(x) &:= \left\{s \in \mathbb{R}^d |\ \langle s, y - x\rangle \leq 0 \text{ for all } y \in A\right\} \\ &= \left\{s \in \mathbb{R}^d |\ s \text{ is a regular normal vector to } A \text{ at } x\right\}. \end{aligned}$$

Recalling that $0 \in N_A(x)$ for $x \in A$ and $N_A(x) = \{0\}$ for $x \in \operatorname{int} A$, we deduce that

$$\operatorname{dom}\overline{\partial} f = \operatorname{dom}\partial f, \tag{7}$$

$$\overline{\partial} f(x) \subset \partial f(x), \quad x \in \operatorname{cl} D, \tag{8}$$

$$\overline{\partial} f(x) = \partial f(x), \quad x \in D. \tag{9}$$

The *diameter* of a set E is denoted by $\operatorname{diam} E := \sup_{x,y \in E} |x - y|$.

Theorem 3.1. *Let* $f\colon \mathbb{R}^d \to \mathbb{R}\cup\{+\infty\}$ *be a closed convex function such that* $D := \operatorname{int}\operatorname{dom} f \neq \emptyset$*. Let* $j \in \{1,\dots,d\}$ *and* A *be a closed set in* $\mathbb{R}^d$ *containing* $\operatorname{range}\nabla f$*. Then*

$$\Sigma^j(\partial f)\cap D = \Sigma^j(\overline{\partial} f)\cap D = \Sigma^j_A(\partial f)\cap D, \tag{10}$$

$$\Sigma^j(\overline{\partial} f) \subset \Sigma^j_A(\partial f). \tag{11}$$

If $y^j = z^j$ *for all* $y,z \in A$*, then, clearly,* $\Sigma^j_A(\partial f) = \emptyset$*. Otherwise, for every* $n \geq 1$*, there exist a compact set* $C_n \subset A - A$ *with* $y^j > 0$*,* $y \in C_n$*, and a function* $h_n \in \mathcal{H}^j_{\theta^j(C_n)}$ *such that*

$$\Sigma^j_A(\partial f) \subset \bigcup_n \left\{x\in\mathbb{R}^d \,|\, h_n(x) = 0\right\}. \tag{12}$$

For any $\epsilon > 0$*, all* C_n *can be chosen so that* $\operatorname{diam}\theta^j(C_n) < \epsilon$*.*

Taking the unions over $j \in \{1,\dots,d\}$, we obtain the descriptions of the *full* singular sets $\Sigma(\partial f)$ on D, $\Sigma(\overline{\partial} f)$, and $\Sigma_A(\partial f)$. Taking a smaller $\epsilon > 0$ in the last sentence of the theorem, we make the directions of the normal vectors to the covering surface H_n closer to each other. As a result, H_n gets approximated by a hyperplane.

Theorem 3.1 uses a larger closed set A instead of F to allow for more flexibility in the treatment of singularities of ∂f on the boundary of D. Lemma 3.4 shows that

$$\partial f(x)\cap A = \operatorname{Arg}_A(x) := \arg\max_{y\in A}\left(\langle x,y\rangle - f^*(y)\right), \quad x\in \operatorname{cl} D.$$

In the framework of Fitzpatrick functions in Section 4, Arg_A becomes a projection on the monotone set A in a pseudo-Euclidean space.

The proof of Theorem 3.1 relies on some lemmas. We start with a simple fact from convex analysis.

Lemma 3.2. *Let* $f:\ \mathbb{R}^d \to \mathbb{R}\cup\{+\infty\}$ *be a closed convex function attaining a strict minimum at a point* x_0*:* $\quad f(x_0) < f(x), \quad x\in\mathbb{R}^d,\ x\neq x_0.$

Then $\quad f(x_0) < \inf_{|x-x_0|\geq\epsilon} f(x), \quad \epsilon > 0.$

Proof. If the conclusion is not true, then there exist $\epsilon > 0$ and a sequence (x_n) with $|x_n - x_0| \geq \epsilon$ such that $f(x_n)\to f(x_0)$. As

$$z_n := x_0 + \frac{\epsilon}{|x_n - x_0|}(x_n - x_0)$$

is a convex combination of x_0 and x_n, we deduce that

$$f(z_n) \leq \max(f(x_0), f(x_n)) = f(x_n) \to f(x_0).$$

By compactness, $z_n \to z_0$ over a subsequence. Clearly, $|z_0 - x_0| = \epsilon$, while by the lower semi-continuity, $f(z_0) \leq \liminf f(z_n) \leq f(x_0)$. We have arrived to a contradiction. $\square$

The following result explains the special role played by $\operatorname{cl}\operatorname{range}\nabla f$. Recall the notation f_A from (3).

Lemma 3.3. *Let $f\colon \mathbb{R}^d \to \mathbb{R}\cup\{+\infty\}$ be a closed convex function such that $D := \operatorname{int}\operatorname{dom} f \neq \emptyset$ and A be a closed set in $\mathbb{R}^d$. Then*

$$f_A(x) = f(x),\ x \in \operatorname{cl} D \iff \operatorname{range}\nabla f \subset A.$$

In other words, $F := \operatorname{cl}\operatorname{range}\nabla f$ is the minimal closed set such that $f_F = f$ on $\operatorname{cl} D$.

Proof. $\Longleftarrow$: If $x \in \operatorname{dom}\nabla f$, then $y := \nabla f(x) \in A$ and (4) yields that

$$f(x) = \langle x, y\rangle - f^*(y) = f_A(x).$$

Since $\operatorname{dom}\nabla f$ is dense in D, the closed convex functions f_A and f coincide on $\operatorname{cl} D$.

$\Longrightarrow$: We fix $x_0 \in \operatorname{dom}\nabla f$ and set $y_0 := \nabla f(x_0)$. By the assumption we have $f_A(x_0) = f(x_0)$. If $y_0 \notin A$, then the distance between y_0 and A is at least $\epsilon > 0$. According to (4), the concave upper semi-continuous function

$$y \to \langle x_0, y\rangle - f^*(y)$$

attains a strict global maximum at y_0 with maximum value $f(x_0)$. By Lemma 3.2,

$$\sup_{|y-y_0|\geq\epsilon} \left(\langle x_0, y\rangle - f^*(y)\right) < f(x_0).$$

As $A \subset \left\{y \in \mathbb{R}^d |\ |y-y_0| \geq \epsilon\right\}$, we arrive to a contradiction:

$$f_A(x_0) = \sup_{y\in A}\left(\langle x_0, y\rangle - f^*(y)\right) \leq \sup_{|y-y_0|\geq\epsilon}\left(\langle x_0, y\rangle - f^*(y)\right) < f(x_0).$$

Hence, $y_0 \in A$, as required. $\square$

Lemma 3.4. *Let $f : \mathbb{R}^d \to \mathbb{R}\cup\{+\infty\}$ be a closed convex function such that $D := \operatorname{int}\operatorname{dom} f \neq \emptyset$. Let A be a closed set in $\mathbb{R}^d$ such that $f = f_A$ on $\operatorname{cl} D$. Then*

$$\partial f(x) \cap A = \operatorname{Arg}_A(x) := \arg\max_{y\in A}\left(\langle x, y\rangle - f^*(y)\right), \quad x \in \operatorname{cl} D,$$
$$\operatorname{dom}\operatorname{Arg}_A \cap \operatorname{cl} D = \operatorname{dom}\partial f = \operatorname{dom}\overline{\partial} f,$$
$$\overline{\partial} f(x) = \partial f(x) = \operatorname{conv}\left(\partial f(x)\cap A\right), \quad x \in D,$$
$$\overline{\partial} f(x) \subset \operatorname{cl}\operatorname{conv}\left(\partial f(x)\cap A\right), \quad x \in \operatorname{dom}\overline{\partial} f.$$

Proof. If $x \in \operatorname{dom}\operatorname{Arg}_A \cap \operatorname{cl} D$ and $y \in \operatorname{Arg}_A(x)$, then

$$f(x) = f_A(x) = \langle x, y\rangle - f^*(y). \tag{13}$$

Hence, $y \in \partial f(x) \cap A$, by the properties of subdifferentials in (4).

Conversely, let $x \in \operatorname{dom}\partial f$. Lemma 3.3 shows that $F := \operatorname{cl}\operatorname{range}\nabla f \subset A$. Accounting for (7) and (8) and the definition of $\overline{\partial} f(x)$, we obtain that

$$\partial f(x) \cap A \supset \partial f(x) \cap F \supset \overline{\partial} f(x) \cap F \neq \emptyset.$$

Let $y \in \partial f(x) \cap A$. From (4) we deduce (13) and then that $y \in \mathrm{Arg}_A(x)$. We have proved the first two assertions of the lemma.

The fact that $\overline{\partial} f = \partial f$ on D has been already stated in (9). To prove the second equality in the third assertion, we fix $x_0 \in D$ and choose $\epsilon > 0$ such that

$$B(x_0, \epsilon) := \left\{x \in \mathbb{R}^d | \ |x - x_0| \leq \epsilon\right\} \subset D.$$

The uniform boundedness of ∂f on compacts in D implies the existence of a constant $K > 0$ such that

$$|y| \leq K < \infty, \quad y \in \partial f(x), \, x \in B(x_0, \epsilon).$$

Denote by $A_K := A \cap \left\{x \in \mathbb{R}^d | \ |x| \leq K\right\}$. As range $\nabla f \subset A$, we deduce from (4) that

$$f(x) = f_{A_K}(x), \quad x \in B(x_0, \epsilon) \cap \mathrm{dom} \nabla f,$$

and then, by the density of $\mathrm{dom} \nabla f$ in D, that $f = f_{A_K}$ on $B(x_0, \epsilon)$. Finally, Lemma 2.3 shows that

$$\partial f(x_0) = \partial f_{A_K}(x_0) = \mathrm{conv}\left(\partial f_{A_K}(x_0) \cap A_K\right) = \mathrm{conv}\left(\partial f(x_0) \cap A\right).$$

If $\mathrm{dom} \nabla f \ni x_n \to x$ and $\nabla f(x_n) \to y$, then, clearly, $y \in F \subset A$. Moreover, $y \in \partial f(x)$, by the continuity of subdifferentials. The last assertion of the lemma readily follows. □

We are ready to finish the proof of Theorem 3.1.

Proof of Theorem 3.1. Lemma 3.3 shows that $f = f_A$ on $\mathrm{cl}\, D$. By Lemma 3.4,

$$\overline{\partial} f(x) = \partial f(x) = \mathrm{conv}\left(\partial f(x) \cap A\right), \quad x \in D,$$

$$\overline{\partial} f(x) \subset \mathrm{cl}\,\mathrm{conv}\left(\partial f(x) \cap A\right), \quad x \in \mathrm{dom}\, \overline{\partial} f.$$

Recalling that $\mathrm{dom}\, \partial f = \mathrm{dom}\, \overline{\partial} f$, we deduce (10) and (11).

Fix $\epsilon > 0$. Let (x_n) be a dense sequence in A and (r_n) be an enumeration of all positive rationals. Denote by $\alpha := (m, n, k, l)$ the indexes for which the compacts

$$C_1^\alpha := \{x \in A | \ |x - x_m| \leq r_k\}, \quad C_2^\alpha := \{x \in A | \ |x - x_n| \leq r_l\},$$

satisfy the constraints:

$$\mathrm{diam}\, \theta^j(C_2^\alpha - C_1^\alpha) < \epsilon \text{ and } x^j > 0, \ x \in C_2^\alpha - C_1^\alpha.$$

We have that

$$\Sigma_A^j(\partial f) = \bigcup_\alpha \Sigma_{C_1^\alpha, C_2^\alpha}(\partial f) = \bigcup_\alpha \left\{x \in \mathbb{R}^d | \ \partial f(x) \cap C_i^\alpha \neq \emptyset, \, i = 1, 2\right\}.$$

For every index α, Theorem 2.2 yields a function $h \in \mathcal{H}^j_{\theta^j(C_2^\alpha - C_1^\alpha)}$ such that

$$\Sigma_{C_1^\alpha, C_2^\alpha}(\partial f) \subset \left\{x \in \mathbb{R}^d | \ h(x) = 0\right\}.$$

We have proved (12) and with it the theorem. □

We conclude the section with the geometric interpretation of range ∇f. For a closed convex function $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$, we denote by epi f its epigraph:

$$\operatorname{epi} f := \left\{(x,q) \in \mathbb{R}^d \times \mathbb{R} |\ f(x) \leq q\right\}.$$

Let E be a closed convex set in $\mathbb{R}^d$. A point $x_0 \in E$ is called *exposed* if there is a hyperplane intersecting E only at x_0. In other words, there is $y_0 \in \mathbb{R}^d$ such that

$$\langle x - x_0, y_0 \rangle > 0, \quad x \in E \setminus \{x_0\}.$$

Proposition 3.5. *Let $f : \mathbb{R}^d \to \mathbb{R} \cup \{+\infty\}$ be a closed convex function such that* $\operatorname{int}\operatorname{dom} f \neq \emptyset$. *Then*

$$\{(y, f^*(y))|\ y \in \operatorname{range} \nabla f\} = \text{ \textit{exposed points of} } \operatorname{epi} f^*.$$

Proof. By definition, (y_0, r_0) is an exposed point of epi f^* if it belongs to epi f^* and

$$\langle y - y_0, x_0 \rangle + (r - r_0)q_0 > 0, \quad (y,r) \in \operatorname{epi} f^* \setminus \{(y_0, r_0)\},$$

for some $(x_0, q_0) \in \mathbb{R}^d \times \mathbb{R}$. The definition of epi f^* ensures that $q_0 > 0$ and $r_0 = f^*(y_0)$. Rescaling (x_0, q_0) so that $q_0 = 1$, we deduce that the function

$$y \to \langle x_0, y \rangle + f^*(y)$$

has the unique minimizer y_0. This is equivalent to y_0 being the only element of $\partial f(-x_0)$, which in turn is equivalent to $-x_0 \in \operatorname{dom} \nabla f$ and $y_0 = \nabla f(-x_0)$. □

4. Singular points of projections in S-spaces

Let $\mathcal{S}^d_m$ be the family of symmetric $d \times d$-matrices of full rank with $m \in \{0, 1, \dots, d\}$ positive eigenvalues. For $S \in \mathcal{S}^d_m$, the bilinear form

$$S(x,y) := \langle x, Sy \rangle = \sum_{i,j=1}^d x^i S_{ij} y^j, \quad x, y \in \mathbb{R}^d,$$

defines the scalar product on a pseudo-Euclidean space $\mathbb{R}^d_m$ with dimension d and index m, which we call the S-space. The quadratic form $S(x,x)$ is the *scalar square* on the S-space; its value may be negative.

For a closed set $G \subset \mathbb{R}^d$, we define the *Fitzpatrick-type* function

$$\psi_G(x) := \sup_{y \in G} \left(S(x,y) - \frac{1}{2} S(y,y) \right) \in \mathbb{R} \cup \{+\infty\}, \quad x \in \mathbb{R}^d,$$

and the projection multi-function

$$P_G(x) := \arg\min_{y \in G} S(x-y, x-y) = \arg\max_{y \in G} \left(S(x,y) - \frac{1}{2} S(y,y) \right), \quad x \in \mathbb{R}^d.$$

Clearly, ψ_G is a closed convex function and P_G takes values in the closed (possibly empty) subsets of G.

A closed set $G \subset \mathbb{R}^d$ is called S-*monotone* or S-*positive* if

$$S(x-y, x-y) \geq 0, \quad x, y \in G,$$

or, equivalently, if its projection multi-function has the natural fixed-point property:

$$x \in P_G(x), \quad x \in G.$$

We denote by $\mathcal{M}(S)$ the family of closed non-empty S-monotone sets in $\mathbb{R}^d$. We refer to [10], [21], [18], and [15, Appendix A], for the results on Fitzpatrick functions ψ_G associated with $G \in \mathcal{M}(S)$.

Example 4.1. (Standard form) If $d = 2m$ and

$$S(x,y) = \sum_{i=1}^{m} \left(x^i y^{m+i} + x^{m+i} y^i \right), \quad x, y \in \mathbb{R}^{2m},$$

then $S \in \mathcal{S}^{2m}_m$ and the S-monotonicity means the standard monotonicity in $\mathbb{R}^{2m} = \mathbb{R}^m \times \mathbb{R}^m$. For a *maximal* monotone set G, the function ψ_G becomes the classical Fitzpatrick function from [10].

Example 4.2. (Canonical form) If Λ is the *canonical* quadratic form in $\mathcal{S}^d_m$:

$$\Lambda(x,y) = \sum_{i=1}^{m} x^i y^i - \sum_{i=m+1}^{d} x^i y^i, \quad x, y \in \mathbb{R}^d,$$

then a closed set G is Λ-monotone if and only if

$$G = \operatorname{graph} f := \{(u, f(u)) |\, u \in D\},$$

where D is a closed set in $\mathbb{R}^m$ and $f: \ D \to \mathbb{R}^{d-m}$ is a 1-Lipschitz function:

$$|f(u) - f(v)| \leq |u - v|, \quad u, v \in D.$$

In view of the Kirszbraun Theorem, [9, 2.10.43], G is maximal Λ-monotone if and only if $D = \mathbb{R}^m$.

While working on the S-space, it is convenient to use appropriate versions of subdifferential and Clark-type subdifferential for a convex function f. For $x \in \operatorname{cl} \operatorname{dom} f$, they are defined as

$$\begin{aligned}
\partial^S f(x) &:= \left\{ y \in \mathbb{R}^d |\, S(z,y) \leq f(x+z) - f(x),\, z \in \mathbb{R}^d \right\} \\
&= \left\{ S^{-1} y \in \mathbb{R}^d |\, \langle z, y \rangle \leq f(x+z) - f(x),\, z \in \mathbb{R}^d \right\} \\
&= S^{-1} \partial f(x), \\
\overline{\partial}^S f(x) &:= S^{-1} \overline{\partial} f(x).
\end{aligned}$$

Lemma 4.3. *Let $S \in \mathcal{S}^d_m$, $G \in \mathcal{M}(S)$, and assume that $D := \operatorname{int} \operatorname{dom} \psi_G \neq \emptyset$.*

Then $$\operatorname{range} S^{-1} \nabla \psi_G \subset G, \quad \operatorname{dom} P_G = \operatorname{dom} \partial^S \psi_G = \operatorname{dom} \overline{\partial}^S \psi_G,$$

and $$\psi_G(x) = \psi^*_G(Sx) = \frac{1}{2} S(x,x), \quad x \in G,$$

$$P_G(x) = \partial^S \psi_G(x) \cap G, \quad x \in \mathbb{R}^d.$$

Proof. We write ψ_G as

$$\psi_G(x) = g^*(x) := \sup_{y\in\mathbb{R}^d} (\langle x, y\rangle - g(y)), \quad x \in \mathbb{R}^d,$$

where $g(y) = \frac{1}{2}S^{-1}(y, y)$ for $y \in SG$ and $g(y) = +\infty$ for $y \notin SG$. From the definition of ψ_G and the S-monotonicity of G we deduce that

$$\psi_G(x) = \frac{1}{2}S(x, x) - \frac{1}{2}\inf_{y\in G} S(x - y, x - y) = \frac{1}{2}S(x, x), \quad x \in G,$$

and then that
$$\psi_G(S^{-1}x) \le g(x), \quad x \in \mathbb{R}^d.$$

As $\psi_G^* = g^{**}$ and g^{**} is the largest closed convex function less than g, we have that

$$\psi_G(S^{-1}x) \le \psi_G^*(x) \le g(x), \quad x \in \mathbb{R}^d.$$

Putting together the relations above, we obtain that

$$\frac{1}{2}S(x, x) = \psi_G(x) \le \psi_G^*(Sx) \le g(Sx) = \frac{1}{2}S(x, x), \quad x \in G.$$

For every $x \in \mathbb{R}^d$, the values of the Fitzpatrick function ψ_G and of the projection multi-function P_G can now be written as

$$\psi_G(x) = \sup_{y\in G} (S(x, y) - \psi_G^*(Sy)) = \sup_{z\in SG} (\langle x, z\rangle - \psi_G^*(z)),$$

$$P_G(x) = \arg\max_{y\in G} \Big(S(x, y) - \frac{1}{2}S(y, y)\Big) = S^{-1} \arg\max_{z\in SG} (\langle x, z\rangle - \psi_G^*(z)).$$

The stated relations between P_G and $\partial^S\psi_G$ follow from Lemma 3.4 as soon as we observe that $P_G(x) = \partial^S\psi_G(x) = \emptyset$ for $x \notin \operatorname{cl} D$. The equality of the domains of $\partial^S\psi_G$ and $\overline{\partial}^S\psi_G$ is just a restatement of (7). Finally, Lemma 3.3 yields the inclusion of range $\nabla\psi_G$ into SG. □

To facilitate geometric interpretations, we also adapt the concept of a normal vector to the product structure of the S-space. Let H be a closed set in $\mathbb{R}^d$. A vector $w \in \mathbb{R}^d$ is called *S-regular normal to H at $x \in H$* if

$$\limsup_{\substack{H\ni y\to x\\ y\ne x}} \frac{S(w, y - x)}{|y - x|} = \limsup_{\substack{H\ni y\to x\\ y\ne x}} \frac{\langle Sw, y - x\rangle}{|y - x|} \le 0.$$

A vector $w \in \mathbb{R}^d$ is called *S-normal to H at x* if there exist $x_n \in H$ and an S-regular normal vector w_n to H at x_n, such that $x_n \to x$ and $w_n \to w$. In other words, w is S-(regular) normal to H at $x \in H$ if and only if Sw is (regular) normal to H at x in the classical Euclidean sense. It is easy to see that if $x \in P_H(z)$, then the vector $z - x$ is S-regular normal to H at x.

Lemma 4.4. *Let $S \in \mathcal{S}^d_m$, $j \in \{1, \ldots, d\}$, $C \in \mathcal{C}^j$, h be given by (1) for convex functions g_1 and g_2 with linear growth, and H be the zero-level set of the composition function $h \circ S$:*

$$H := \left\{x \in \mathbb{R}^d \,|\, h(Sx) = 0\right\}.$$

Then $h \in \mathcal{H}^j_C$ if and only if for every $x \in H$, there exists an S-normal vector $w \in C$ to H at x.

Proof. Let $z \in H$ and $w \in \mathbb{R}^d$. Setting

$$SH := \{Sx \mid x \in H\} = \left\{x \in \mathbb{R}^d \mid h(x) = 0\right\},$$

recalling that $S(w, y - z) = \langle w, Sy - Sz\rangle$, and using the trivial inequalities:

$$\frac{1}{\|S^{-1}\|} |y - z| \leq |Sy - Sz| \leq \|S\| \, |y - z|, \quad y \in \mathbb{R}^d,$$

where $\|A\| := \max_{|x|=1} |Ax|$ for a $d \times d$ matrix A, we deduce that w is normal to SH at Sz if and only if w is S-normal to H at z. Lemma 2.1 yields the result. $\square$

Theorem 4.5 and Lemma 4.4 show that the singular sets $\Sigma^j(P_G)$ and $\Sigma^j(\overline{\partial}^S \psi_G)$ can be covered by countable $c - c$ surfaces that have at each point an S-normal vector $w \in G - G$, with $w^j = 1$. By the S-monotonicity of G, such vector w points to the *non-negative* direction in the S-space: $S(w, w) \geq 0$.

Theorem 4.5. *Let $S \in \mathcal{S}^d_m$, $G \in \mathcal{M}(S)$, and assume that* $\operatorname{int} \operatorname{dom} \psi_G \neq \emptyset$. *Let $j \in \{1, \dots, d\}$. For every $n \geq 1$, there exist a compact set $C_n \subset G - G$ and a function $h_n \in \mathcal{H}^j_{\theta^j(C_n)}$ such that*

$$\Sigma^j(\overline{\partial}^S \psi_G) \subset \Sigma^j(P_G) \subset \bigcup_n \left\{x \in \mathbb{R}^d \mid h_n(Sx) = 0\right\}.$$

For any $\epsilon > 0$, all C_n can be chosen such that $\operatorname{diam} \theta^j(C_n) < \epsilon$.

Proof. Let $g(x) := \psi_G(S^{-1}x)$, $x \in \mathbb{R}^d$. Clearly,

$$\overline{\partial}^S \psi_G(x) := S^{-1}\overline{\partial}\psi_G(x) = \overline{\partial} g(Sx), \quad x \in \mathbb{R}^d.$$

Lemma 4.3 shows that $S^{-1} \operatorname{range} \nabla\psi_G \subset G$ and

$$P_G(x) = \partial^S \psi_G(x) \cap G = \left(S^{-1}\partial\psi_G(x)\right) \cap G = \partial g(Sx) \cap G, \quad x \in \mathbb{R}^d.$$

It follows that

$$\Sigma^j(\overline{\partial}^S \psi_G) = S^{-1}\Sigma^j(\overline{\partial} g), \quad \Sigma^j(P_G) = S^{-1}\Sigma^j_G(\partial g),$$

and a direct application of Theorem 3.1 yields the result. $\square$

A set $A \subset \mathbb{R}^d$ is called S-*isotropic* if

$$S(x - y, x - y) = 0, \quad x, y \in A.$$

We denote by $\mathcal{I}(S)$ the family of all closed S-isotropic subsets of $\mathbb{R}^d$.

Motivated by the study of existence and uniqueness of backward martingale transport maps in [16], we decompose the singular set of P_G as

$$\Sigma(P_G) := \{x \in \operatorname{dom} P_G \mid P_G(x) \text{ is not a point}\} = \Sigma_0(P_G) \cup \Sigma_1(P_G),$$

$$\Sigma_0(P_G) := \{x \in \Sigma(P_G) \mid P_G(x) \in \mathcal{I}(S)\},$$

$$\Sigma_1(P_G) := \{x \in \Sigma(P_G) \mid S(y_1 - y_2, y_1 - y_2) > 0 \text{ for } \textit{some } y_1, y_2 \in P_G(x)\}.$$

We further write $\Sigma_1(P_G)$ as

$$\Sigma_1(P_G) = \bigcup_{j=1}^{d} \Sigma_1^j(P_G),$$

where $\Sigma_1^j(P_G)$ consists of $x \in \Sigma(P_G)$ such that $S(y_1 - y_2, y_1 - y_2) > 0$ for some $y_1, y_2 \in P_G(x)$ with $y_1^j \neq y_2^j$.

Theorem 4.6 and Lemma 4.4 show that $\Sigma_1^j(P_G)$ can be covered by countable $c - c$ surfaces that have at each point a *strictly* S-positive S-normal vector w with $w^j = 1$.

Theorem 4.6. *Let $S \in \mathcal{S}_m^d$, $G \in \mathcal{M}(S)$, and assume that* $\operatorname{int}\operatorname{dom}\psi_G \neq \emptyset$. *Let $j \in \{1, \ldots, d\}$. For every $n \geq 1$, there exist a compact set $C_n \subset G - G$ with*

$$x^j > 0 \text{ and } S(x, x) > 0, \quad x \in C_n,$$

and a function $h_n \in \mathcal{H}^j_{\theta^j(C_n)}$ such that

$$\Sigma_1^j(P_G) \subset \bigcup_n \left\{x \in \mathbb{R}^d |\ h_n(Sx) = 0\right\}.$$

For any $\epsilon > 0$, all C_n can be chosen such that $\operatorname{diam}\theta^j(C_n) < \epsilon$.

Proof. Fix $\epsilon > 0$. Let (x_n) be a dense sequence in G and (r_n) be an enumeration of all positive rationals. Denote by $\alpha := (m, n, k, l)$ the indexes for which the compact sets

$$C_1^\alpha := \{x \in G |\ |x - x_m| \leq r_k\}, \quad C_2^\alpha := \{x \in G |\ |x - x_n| \leq r_l\},$$

satisfy the constraints:

$$\operatorname{diam}\theta^j(C_2^\alpha - C_1^\alpha) < \epsilon \text{ and } x^j > 0,\ S(x, x) > 0,\ x \in C_2^\alpha - C_1^\alpha.$$

We have that

$$\Sigma_1^j(P_G) = \bigcup_\alpha \Sigma_{C_1^\alpha, C_2^\alpha}(P_G) = \bigcup_\alpha \left\{x \in \mathbb{R}^d |\ P_G(x) \cap C_i^\alpha \neq \emptyset,\ i = 1, 2\right\}.$$

Let, again, $g(x) := \psi_G(S^{-1}x)$, $x \in \mathbb{R}^d$. From Lemma 4.3 we deduce that

$$P_G(x) = \left(S^{-1}\partial\psi_G(x)\right) \cap G = \partial g(Sx) \cap G, \quad x \in \mathbb{R}^d.$$

It follows that

$$x \in \Sigma_{C_1^\alpha, C_2^\alpha}(P_G) \iff Sx \in \Sigma_{C_1^\alpha, C_2^\alpha}(\partial g).$$

Applying Theorem 2.2 to each singular set $\Sigma_{C_1^\alpha, C_2^\alpha}(\partial g)$, we obtain the result. □

The singular set $\Sigma_0(P_G)$ is included into a larger set

$$\overline{\Sigma}_0(P_G) = \bigcup_{j=1}^{d} \overline{\Sigma}_0^j(P_G),$$

where $\overline{\Sigma}_0^j(P_G)$ consists of $x \in \operatorname{dom} P_G$ such that $S(y_1 - y_2, y_1 - y_2) = 0$ for *some* $y_1, y_2 \in P_G(x)$ with $y_1^j \neq y_2^j$.

Theorem 4.7 and Lemma 4.4 show that $\overline{\Sigma}_0^j(P_G)$ can be covered, for any $\delta > 0$, by countable $c - c$ surfaces that have at each point an S-normal vector w such that $w^j = 1$ and $0 \leq S(w, w) \leq \delta$.

Theorem 4.7. *Let $S \in \mathcal{S}_m^d$, $G \in \mathcal{M}(S)$, and assume that* $\operatorname{int}\operatorname{dom}\psi_G \neq \emptyset$. *Let $j \in \{1, \dots, d\}$ and $\delta > 0$. For every $n \geq 1$, there exist a compact set $C_n \subset G - G$ with*

$$x^j > 0, \ x \in C_n, \ \text{and } 0 \leq S(x, x) \leq \delta, \ x \in \theta^j(C_n),$$

and a function $h_n \in \mathcal{H}^j_{\theta^j(C_n)}$ such that

$$\overline{\Sigma}_0^j(P_G) \subset \bigcup_n \left\{x \in \mathbb{R}^d | \ h_n(Sx) = 0\right\}.$$

For any $\epsilon > 0$, all C_n can be chosen such that $\operatorname{diam}\theta^j(C_n) < \epsilon$.

Proof. The proof is almost identical to that of Theorem 4.6. We fix $\delta, \epsilon > 0$ and find a *countable* family (C_1^α, C_2^α) of pairs of compact sets $C_i^\alpha \subset G$, $i = 1, 2$, such that

$$x^j > 0, \ x \in C_2^\alpha - C_1^\alpha, \ \text{ and } \ \operatorname{diam}\theta^j(C_2^\alpha - C_1^\alpha) < \epsilon,$$

$$0 \leq S(x, x) < \delta, \ x \in \theta^j(C_2^\alpha - C_1^\alpha),$$

and every pair (y_1, y_2) of elements of G with $S(y_1 - y_2, y_1 - y_2) = 0$ and $y_1^j \neq y_2^j$ is contained in some (C_1^α, C_2^α). Then

$$\overline{\Sigma}_0^j(P_G) \subset \bigcup_\alpha \Sigma_{C_1^\alpha, C_2^\alpha}(P_G) = \bigcup_\alpha S^{-1}\Sigma_{C_1^\alpha, C_2^\alpha}(\partial g),$$

where $g(x) := \psi_G(S^{-1}x)$, $x \in \mathbb{R}^d$, and Theorem 2.2 applied to $\Sigma_{C_1^\alpha, C_2^\alpha}(\partial g)$ yields the result. □

The geometric description of the zero order singularities becomes especially simple if the index $m = 1$. In this case, $\overline{\Sigma}_0^j(P_G)$ can be covered by countable number of hyperplanes whose S-normal vectors are S-isotropic.

Theorem 4.8. *Let $S \in \mathcal{S}_1^d$, $G \in \mathcal{M}(S)$, and assume that* $\operatorname{int}\operatorname{dom}\psi_G \neq \emptyset$. *Let $j \in \{1, \dots, d\}$. For every $n \geq 1$, there exist $y_n, z_n \in G$ such that*

$$y_n^j - z_n^j > 0, \quad S(y_n - z_n, y_n - z_n) = 0, \tag{14}$$

and

$$\overline{\Sigma}_0^j(P_G) \subset \bigcup_n \left\{x \in \mathbb{R}^d | \ S(x - z_n, y_n - z_n) = 0\right\}.$$

Proof. By the law of inertia for quadratic forms, [12, Theorem 1, p. 297], there exists a $d \times d$ matrix V of full rank such that

$$S = V^T \Lambda V,$$

where V^T is the transpose of V and Λ is the diagonal matrix with diagonal entries $\{1, -1, \dots, -1\}$. In other words, Λ is the canonical quadratic form for $\mathcal{S}^d_1$ from Example 4.2. As $S(x,x) = \Lambda(Vx, Vx)$, we deduce that $F := VG \in \mathcal{M}(\Lambda)$ and

$$A \in \mathcal{I}(S) \iff VA \in \mathcal{I}(\Lambda).$$

As we pointed out in Example 4.2,

$$F = \operatorname{graph} f = \{(t, f(t)) |\, t \in D\},$$

for a 1-Lipschitz function $f:\ D \to \mathbb{R}^{d-1}$ defined a closed set $D \subset \mathbb{R}$.

Let $x = (s, f(s))$ and $y = (t, f(t))$ be distinct elements of F, where $s, t \in D$. We have that $\{x, y\} \in \mathcal{I}(\Lambda)$ if and only if $|f(t) - f(s)| = |t - s|$. The 1-Lipschitz property of f then implies that

$$f(r) = f(s) + \frac{r-s}{t-s}\left(f(t) - f(s)\right), \quad r \in D \cap [s,t].$$

It follows that the collection of all Λ-isotropic subsets of F can be decomposed into an intersection of F with at most countable union of line segments whose relative interiors do not intersect.

The same property then also holds for the S-isotropic subsets of G. Thus, there exist $y_n, z_n \in G$, $n \geq 1$, satisfying (14) and such that every S-isotropic subset of G having elements with distinct jth coordinates is a subset of some S-isotropic line $L_n := \{y_n + t(z_n - y_n) |\, t \in \mathbb{R}\}$. In particular, if $x \in \overline{\Sigma}^j_0(P_G)$, then $P_G(x)$ intersects some line L_n at distinct y and z. We have that

$$2S(x-z, y-z) = S(x-z, x-z) + S(y-z, y-z) - S(x-y, x-y) = 0.$$

As $y, z \in L_n \in \mathcal{I}(S)$, we obtain that $S(x - z_n, y_n - z_n) = 0$, as required. $\square$

Example 4.9. Let $d = 2$ and S be the standard bilinear form from Example 4.1:

$$S(x,y) = S((x_1, x_2), (y_1, y_2)) = x_1 y_2 + x_2 y_1.$$

Let $G \in \mathcal{M}(S)$. As $S(x,x) = 2x_1x_2$, we have that

$$\Sigma_1(P_G) = \Sigma^j_1(P_G), \quad j = 1, 2.$$

Theorem 4.6 yields convex functions $g_{1,n}$ and $g_{2,n}$ on $\mathbb{R}$ and constants $\epsilon_n > 0$, $n \geq 1$, such that ($g' = g'(t)$ is the derivative of $g = g(t)$)

$$\epsilon_n \leq g'_{1,n}(t) - g'_{2,n}(t) \leq \epsilon_n^{-1}, \quad t \in \operatorname{dom} g'_{1,n} \cap \operatorname{dom} g'_{2,n},$$

$$\Sigma_1(P_G) \subset \bigcup_n \left\{x \in \mathbb{R}^2 |\, x_2 = g_{2,n}(x_1) - g_{1,n}(x_1)\right\}.$$

Theorem 4.8 yields sequences (x^n_1) and (x^n_2) in $\mathbb{R}$ such that

$$\overline{\Sigma}^1_0(P_G) \subset \bigcup_n \left\{x \in \mathbb{R}^2 |\, x_2 = x^n_2\right\}, \quad \overline{\Sigma}^2_0(P_G) \subset \bigcup_n \left\{x \in \mathbb{R}^2 |\, x_1 = x^n_1\right\}.$$

These results improve [17, Theorem B.12], where G is a maximal monotone set and $g_n := g_{1,n} - g_{2,n}$ is a strictly increasing Lipschitz function such that $\epsilon_n \leq g'_n(t) \leq \epsilon_n^{-1}$, whenever it is differentiable.

Acknowledgments. We thank Giovanni Leoni for pointing out the references [23] and [11]. We also thank an anonymous referee for constructive remarks and for bringing to our attention the references [5] and [22].

References

[1] P. Albano, P. Cannarsa: *Structural properties of singularities of semiconcave functions*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) 28/4 (1999) 719–740.

[2] G. Alberti: *On the structure of singular sets of convex functions*, Calc. Var. Partial Diff. Equations 2 (1994) 17–27.

[3] L. Ambrosio, N. Gigli: *A user's guide to optimal transport*, in: *Modelling and Optimisation of Flows on Networks*, Lecture Notes in Mathematics 2062, Springer, Berlin (2013) 1–155.

[4] E. Asplund: *Čebyšev sets in Hilbert space*, Trans. Amer. Math. Soc. 144 (1969) 235–240.

[5] Y. Benyamini, J. Lindenstrauss: *Geometric Nonlinear Functional Analysis. Vol. 1*, AMS Colloquium Publications Volume 48, American Mathematical Society, Providence (2000).

[6] Y. Brenier: *Polar factorization and monotone rearrangement of vector-valued functions*, Comm. Pure Appl. Math. 44 (1991) 375–417.

[7] F. H. Clarke: *Generalized gradients and applications*, Trans. Amer. Math. Soc. 205 (1975) 247–262.

[8] P. Erdös: *On the Hausdorff dimension of some sets in Euclidean space*, Bull. Amer. Math. Soc. 52 (1946) 107–109.

[9] H. Federer: *Geometric Measure Theory*, Grundlehren der Mathematischen Wissenschaften 153, Springer, New York (1969).

[10] S. Fitzpatrick: *Representing monotone operators by convex functions*, in: *Workshop/Miniconference on Functional Analysis and Optimization*, Canberra 1988), Proc. Centre Math. Anal. Australian Nat. Univ. vol. 20, Australian National University, Canberra (1988) 59–65.

[11] T. Fowler, D. Preiss: *A simple proof of Zahorski's description of non-differentiability sets of Lipschitz functions*, Real Anal. Exchange 34 (2009) 127–138.

[12] F. R. Gantmacher: *The Theory of Matrices. Vol. 1*, translated from the Russian by K. A. Hirsch, reprint of the 1959 translation, AMS Chelsea Publishing, American Mathematical Society, Providence (1998).

[13] P. Hajłasz: *On an old theorem of Erdös about ambiguous locus*, Colloq. Math. 168 (2022) 249–256.

[14] J.-B. Hiriart-Urruty, C. Lemaréchal: *Fundamentals of Convex Analysis*, Grundlehren Text Editions, Springer, Berlin (2001).

[15] D. Kramkov, M. Sîrbu: *Backward martingale transport and Fitzpatrick functions in pseudo-Euclidean spaces*, Ann. Appl. Probab. 34/1B (2024) 1571–1599.

[16] D. Kramkov, M. Sîrbu: *Backward martingale transport maps in pseudo-Euclidean spaces*, arXiv: 2304.08290 (2023).

[17] D. Kramkov, Y. Xu: *An optimal transport problem with backward martingale constraints motivated by insider trading*, Ann. Appl. Probab. 32 (2022) 294–326.

[18] J.-P. Penot: *Positive sets, conservative sets and dissipative sets*, J. Convex Analysis 16 (2009) 973–986.

[19] R. T. Rockafellar: *Convex Analysis*, Princeton Mathematical Series No. 28, Princeton University Press, Princeton (1970).

[20] R. T. Rockafellar, R. J.-B. Wets: *Variational Analysis*, Grundlehren der Mathematischen Wissenschaften 317, Springer, Berlin (1998).

[21] S. Simons: *Positive sets and monotone sets*, J. Convex Analysis 14 (2007) 297–317.

[22] L. Thibault: *Unilateral Variational Analysis in Banach Spaces. Part II: Special Classes of Functions and Sets*, World Scientific, Singapore (2023).

[23] Z. Zahorski: *Sur la première dérivée*, Trans. Amer. Math. Soc. 69 (1950) 1–54.

[24] L. Zajíček: *On the differentiation of convex functions in finite and infinite dimensional spaces*, Czechoslovak Math. J. 29/104 (1979) 340–348.