PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 139, Number 7, July 2011, Pages 2519–2527
S 0002-9939(2010)10692-5
Article electronically published on December 21, 2010

GENERIC NONDEGENERACY IN CONVEX OPTIMIZATION

DMITRIY DRUSVYATSKIY AND ADRIAN S. LEWIS

(Communicated by Tatiana Toro)

ABSTRACT. We show that minimizers of convex functions subject to almost all linear perturbations are nondegenerate. An analogous result holds more generally for *lower*-$\mathbf{C}^2$ functions.

1. INTRODUCTION

In this work we study the nature of minimizers of "typical" convex functions. We model this question by considering a fixed extended-real-valued convex function f and then studying properties of minimizers of the perturbed function $x \mapsto f_v(x) = f(x) - v^T x$ that hold "generically", by which we mean for *almost all* values of the data vector v in $\mathbf{R}^n$ (in the sense of Lebesgue measure).

Classical theory shows that, given a proper convex function f, the perturbed function f_v typically has at most one minimizer. To see this, note first that we may assume f is closed, since any minimizer of f also minimizes its closure. Now we observe that the Fenchel conjugate f^* is differentiable almost everywhere on the interior of its domain, by Rademacher's theorem (see for example [9, Theorem 9.60]), so for almost all vectors v, the subdifferential $\partial f^*(v)$ is either single-valued or empty. The result now follows, since this subdifferential coincides with the set $(\partial f)^{-1}(v)$, which is exactly the set of minimizers of f_v.

Our aim is to strengthen this classical result. Minimizers x of the perturbed function f_v are characterized by the property that the vector zero lies in the subdifferential $\partial f_v(x)$. We prove, for almost all vectors v, that the minimizer x is not only unique but also *nondegenerate*, by which we mean that zero lies in the relative interior of the subdifferential: $0 \in \mathrm{ri}\partial f_v(x)$ (or equivalently, the positive span $\mathbf{R}_+\partial f_v(x)$ is a subspace). The proof, following an idea of [7], uses a result in geometric measure theory due to Larman [5].

As an example, consider the standard linear programming problem

$$\max_{x\in\mathbf{R}^n} \Big\{ v^T x : a_i^T x \leq b_i \ (i = 1, 2, \ldots, m) \Big\},$$

Received by the editors May 5, 2010 and, in revised form, July 8, 2010.

2010 *Mathematics Subject Classification.* Primary 49J53, 32F32; Secondary 47H05.

Key words and phrases. Convex functions, normal cone, subdifferential, Hausdorff measure, lower-$\mathbf{C}^2$ functions.

The work of the first author was supported in part by the NDSEG grant from the Department of Defense.

The work of the second author was supported in part by National Science Foundation Grant DMS-0806057.

for given vectors $a_i \in \mathbf{R}^n$ and scalars $b_i \in \mathbf{R}$. We can restate this problem as minimizing the perturbed function f_v corresponding to the original function f that takes the value zero on the feasible region and $+\infty$ elsewhere. Consider an optimal solution $\bar{x}$ and the corresponding index set of active constraints, $I = \{i : a_i^T \bar{x} = b_i\}$. Then we have

$$\partial f_v(\bar{x}) = -v + \{\sum_{i \in I} \lambda_i a_i : \lambda_i \geq 0\},$$

$$\mathrm{ri}\partial f_v(\bar{x}) = -v + \{\sum_{i \in I} \lambda_i a_i : \lambda_i > 0\}.$$

Thus the minimizer $\bar{x}$ of f_v is nondegenerate exactly when there exists a dual-feasible solution $\lambda \in \mathbf{R}^m$ satisfying strict complementary slackness. We hence recover the well-known fact that, for almost all objective functions, if a linear program has an optimal solution, then that solution is unique and furthermore corresponds to a strictly complementary-slack dual solution.

For convex functions, critical points (those at which zero is a subgradient) coincide with minimizers. For nonconvex functions, we can more generally consider nondegeneracy of critical points. It transpires that our result on typical nondegeneracy extends in particular to all *lower*-$\mathbf{C^2}$ functions (those functions locally representable as differences of convex functions and convex quadratics). However, in more general contexts the result may fail. The classical generalization of the subdifferential of a convex function is the Clarke generalized gradient [4], but [3] presents a locally Lipschitz function $f\colon \mathbf{R} \to \mathbf{R}$, whose Clarke generalized gradient $\partial_c f$ at any point $x \in \mathbf{R}$ is the interval $[-x, x]$. In this case, the perturbed function f_v has a degenerate critical point for *every* nonzero value of v.

2. Preliminaries

2.1. **Variational analysis.** We recall some standard notions from variational analysis (see for example [9]). Consider the extended real line $\overline{\mathbf{R}} := \mathbf{R} \cup \{-\infty\} \cup \{+\infty\}$. We say that an extended real-valued function is *proper* if it is never $\{-\infty\}$ and is not always $\{+\infty\}$.

For a function $f\colon \mathbf{R}^n \to \overline{\mathbf{R}}$, we define the *domain* of f to be

$$\mathrm{dom}\, f = \{x \in \mathbf{R}^n : f(x) < +\infty\},$$

and we define the *epigraph* of f to be

$$\mathrm{epi}\, f = \{(x, r) \in \mathbf{R}^n \times \mathbf{R} : r \geq f(x)\}.$$

A function is *convex* when its epigraph is convex and is *closed* when its epigraph is closed. Throughout, we will use $|\cdot|$ to denote the standard Euclidean norm.

Definition 2.1. Consider a set $S \subset \mathbf{R}^n$ and a point $\bar{x} \in S$. The *regular normal cone* to S at $\bar{x}$, denoted $\hat{N}_S(\bar{x})$, consists of all vectors $v \in \mathbf{R}^n$ such that

$$\langle v, x - \bar{x} \rangle \leq o(|x - \bar{x}|) \text{ for } x \in S,$$

where we denote by $o(|x - \bar{x}|)$ for $x \in S$ a term with the property that

$$\frac{o(|x - \bar{x}|)}{|x - \bar{x}|} \to 0$$

when $x \xrightarrow{S} \bar{x}$ with $x \neq \bar{x}$.

Definition 2.2. Consider a set $S \subset \mathbf{R}^n$ and a point $\bar{x} \in S$. The *limiting normal cone* to S at $\bar{x}$, denoted $N_S(\bar{x})$, consists of all vectors $v \in \mathbf{R}^n$ such that there are sequences $x_r \xrightarrow{S} \bar{x}$ and $v_r \to v$ with $v_r \in \hat{N}_S(x_r)$.

In the presence of convexity, normal cones have a much simpler form.

Theorem 2.3 ([9, Theorem 6.9])**.** *For a convex set $S \subset \mathbf{R}^n$ and a point $\bar{x} \in S$, the regular and the limiting normal cones coincide and consist of all vectors $v \in \mathbf{R}^n$ such that*

$$\langle v, x - \bar{x} \rangle \leq 0 \text{ for all } x \in S.$$

Normal cones allow us to study geometric objects. We now define subdifferentials, which allow us to analyze behavior of functions.

Definition 2.4. Consider a function $f \colon \mathbf{R}^n \to \overline{\mathbf{R}}$ and a point $\bar{x} \in \mathbf{R}^n$ where f is finite. The *regular* and the *limiting subdifferentials* of f at $\bar{x}$, respectively, are defined by

$$\begin{aligned} \hat{\partial} f(\bar{x}) &= \big\{ v \in \mathbf{R}^n : (v, -1) \in \hat{N}_{\operatorname{epi} f}(\bar{x}, f(\bar{x})) \big\}, \\ \partial f(\bar{x}) &= \big\{ v \in \mathbf{R}^n : (v, -1) \in N_{\operatorname{epi} f}(\bar{x}, f(\bar{x})) \big\}. \end{aligned}$$

If the function f is convex, both subdifferentials reduce to the classical convex subdifferential

$$\big\{ v \in \mathbf{R}^n : \langle v, x - \bar{x} \rangle \leq f(x) - f(\bar{x}) \text{ for all } x \in \mathbf{R}^n \big\}.$$

Remark 2.5. For $x \in \mathbf{R}^n$ where $f(x)$ is not finite, we follow the convention that $\hat{\partial} f(x) = \partial f(x) = \emptyset$. The regular and the limiting subdifferentials are always closed sets, and the regular subdifferential is convex.

Subdifferentials play the role of generalized gradients in the following sense.

Theorem 2.6 ([9, Exercise 8.8])**.** *Consider a function $f \colon \mathbf{R}^n \to \overline{\mathbf{R}}$ and a point $\bar{x} \in \mathbf{R}^n$. If f can be written as $f = g + h$, where g is finite at $\bar{x}$ and h is $\mathbf{C^1}$ smooth on a neighborhood of $\bar{x}$, then*

$$\begin{aligned} \partial f(\bar{x}) &= \partial g(\bar{x}) + \nabla h(\bar{x}), \\ \hat{\partial} f(\bar{x}) &= \hat{\partial} g(\bar{x}) + \nabla h(\bar{x}). \end{aligned}$$

Theorem 2.7 ([9, Theorems 12.12, 12.17])**.** *Let $f \colon \mathbf{R}^n \to \overline{\mathbf{R}}$ be a proper, convex function. Then on the set where the set-valued mapping $(I + \partial f)^{-1}$ takes nonempty values, it is single-valued and Lipschitz continuous with constant* 1.

Remark 2.8. Theorem 2.7 is a special case of the celebrated theorem of Minty. See [6] or [9, Section 12.B] for more details.

We now define a large and robust class of functions that includes both $\mathbf{C^2}$ smooth functions and finite convex functions.

Definition 2.9 ([9, Theorem 10.33])**.** A function $f \colon O \to \mathbf{R}$, where O is an open set in $\mathbf{R}^n$, is said to be *lower-*$\mathbf{C^2}$ on O if for each point $\bar{x} \in O$ there is a neighborhood around $\bar{x}$ and a scalar ρ such that on this neighborhood $f + \rho|\cdot|^2$ is a finite convex function.

By Theorem 2.6, the regular and limiting subdifferentials coincide for lower-$\mathbf{C^2}$ functions.

Remark 2.10. To illustrate the abundance of lower-$\mathbf{C^2}$ functions, consider the following example. Given $\mathbf{C^2}$ functions $f_i\colon O \to \mathbf{R}$ on an open set $O \subset \mathbf{R}^n$ $(i = 1, \ldots, m)$, the function $f = \max\{f_1, \ldots, f_m\}$ is lower-$\mathbf{C^2}$ on O. For more details see [9, Chapter 10.F].

2.2. **Hausdorff measures.** For a set $U \subset \mathbf{R}^n$, let diam U denote its diameter; that is,

$$\operatorname{diam}(U) = \sup_{x,y \in U} |x - y|.$$

Definition 2.11. Consider a set $S \subset \mathbf{R}^n$ and real numbers $\delta, d > 0$. We define

$$\lambda_d^\delta(S) = \inf \Big\{ \sum_{i=1}^{\infty} \operatorname{diam}(U_i)^d : S \subset \bigcup_{i=1}^{\infty} U_i,\ \operatorname{diam}(U_i) < \delta \Big\}.$$

Observe that the infimum in the definition above is taken over all countable covers $\{U_i\}$ of S such that $\operatorname{diam}(U_i) < \delta$ for each i.

Definition 2.12. For a set $S \subset \mathbf{R}^n$, define the *d-dimensional Hausdorff measure* of S to be

$$\lambda_d(S) = \lim_{\delta \to 0} \lambda_d^\delta(S).$$

It can be shown that for each $d > 0$, the set function λ_d is an outer measure on $\mathbf{R}^n$. Furthermore, if d is a positive integer, then on Lebesgue measurable sets in $\mathbf{R}^d$ the d-dimensional Hausdorff measure is a rescaling of the d-dimensional Lebesgue measure. For more details, see [10]. The following is an easy consequence of the definition of Hausdorff measure.

Proposition 2.13. *Consider a set $S \subset \mathbf{R}^n$ and let $f\colon S \to \mathbf{R}^m$ be a Lipschitz continuous mapping with Lipschitz constant κ. Then for any real number $d > 0$, we have $\lambda_d(f(S)) \leq \kappa^d \lambda_d(S)$.*

Corollary 2.14. *Consider a set $S \subset \mathbf{R}^n$ and let $f\colon S \to \mathbf{R}^m$ be a locally Lipschitz mapping. Then for any real number $d > 0$, if $\lambda_d(S) = 0$, then $\lambda_d(f(S)) = 0$.*

Proof. Around each point $x \in S$, consider a neighborhood in S on which f is Lipschitz continuous. This collection of neighborhoods forms a cover of S, and hence there is a countable subcover, say $\{V_i\}$. By Proposition 2.13, for each index i we have $\lambda_d(f(V_i)) = 0$, and hence

$$\lambda_d(f(S)) = \lambda_d\Big(\bigcup_{i=1}^{\infty} f(V_i)\Big) \leq \liminf_{n \to \infty} \sum_{i=1}^{n} \lambda_d(V_i) = 0,$$

as claimed. □

We note that for $d = n$, Corollary 2.14 appears as Lemma 2.5 in [1].

Definition 2.15. Consider a compact, convex set $F \subset \mathbf{R}^n$. The set of maximizers $\operatorname{argmax}_{x \in F} \langle c, x \rangle$ is called the *exposed face* of the set F corresponding to the vector c. In particular, the set F is itself an exposed face (corresponding to $c = 0$). All other exposed faces are said to be *proper*.

For a convex set $S \subset \mathbf{R}^n$, we will denote its closure, relative interior, and relative boundary by $\mathrm{cl}S$, $\mathrm{ri}S$, and $\mathrm{rb}S$, respectively. To prove the main result, we will need the following two theorems.

Theorem 2.16 (Larman [5]). *Let $S \subset \mathbf{R}^n$ be a compact convex set. Let N be the union of the relative boundaries of all the proper exposed faces. Then $\lambda_{n-1}(N) = 0$.*

Theorem 2.17 ([2, Proposition 3]). *Suppose zero lies in the interior of the compact convex set $F \subset \mathbf{R}^n$. Then the proper exposed faces of the polar set F° are those sets of the form*

$$G = \{c \in N_F(x) : \langle c, x\rangle = 1\},$$

for points x on the boundary of F. Furthermore, any such exposed face has relative interior given by

$$\mathrm{ri}G = \{c \in \mathrm{ri}N_F(x) : \langle c, x\rangle = 1\}.$$

3. Main result

3.1. **Subdifferentials of convex functions.** The unit sphere in $\mathbf{R}^n$ will be denoted by $\mathbb{S}^{n-1}$, and an open ball of radius r around a point $x \in \mathbf{R}^n$ will be denoted by $B(x, r)$.

Lemma 3.1. *Let $F \subset \mathbf{R}^n$ be a convex set. Then*

$$\lambda_{n-1}\Big(\big(\bigcup_{x\in F} \mathrm{rb}N_F(x)\big) \cap \mathbb{S}^{n-1}\Big) = 0.$$

Proof. Observe that $N_F(x) = N_{\mathrm{cl}\,F}(x)$ for $x \in F$, so it is sufficient to show that the statement of the lemma holds for a closed convex set F. First, let us consider the case when F is a compact convex set. Without loss of generality, we can assume that zero is in the interior of F, since otherwise we can translate F, so as to have $0 \in \mathrm{ri}F$, and then consider $\mathbf{R}^n$ as the direct sum of the span of F and its orthogonal complement. Define

$$G := \bigcup_{x\in F} \{c \in \mathrm{rb}N_F(x) : \langle c, x\rangle = 1\}.$$

Combining Theorems 2.16 and 2.17 , we deduce $\lambda_{n-1}(G) = 0$. Observe that G is contained in $\mathbf{R}^n \setminus \{0\}$. Now consider the mapping

$$f : \mathbf{R}^n \setminus \{0\} \to \mathbb{S}^{n-1},$$
$$x \mapsto |x|^{-1}x.$$

The mapping f is locally Lipschitz. Consequently, by Corollary 2.14, we have $\lambda_{n-1}(f(G)) = 0$. Observe that the image set $f(G)$ is contained in $(\bigcup_{x\in F} \mathrm{rb}N_F(x))\cap \mathbb{S}^{n-1}$, since f simply scales each element of G. Now, to see the reverse inclusion, consider a vector $c \in (\mathrm{rb}N_F(\bar{x})) \cap \mathbb{S}^{n-1}$ for some vector $\bar{x} \in F$. By definition of the normal cone, we have

$$\langle c, \bar{x} - x\rangle \geq 0, \text{ for all } x \in F.$$

In particular, since 0 lies in the interior of F, we have $\langle c, \bar{x}\rangle > 0$. So we deduce $\widehat{c} := |\langle c, \bar{x}\rangle|^{-1}c \in G$ and $f(\widehat{c}) = c$. Thus we have shown

$$f(G) = \big(\bigcup_{x\in F} \mathrm{rb}N_F(x)\big) \cap \mathbb{S}^{n-1},$$

and consequently

$$\lambda_{n-1}\Big(\big(\bigcup_{x\in F} \mathrm{rb}N_F(x)\big) \cap \mathbb{S}^{n-1}\Big) = 0,$$

as we claimed.

To get rid of the boundedness assumption on F, we will use a standard limiting argument. Assume that F is a closed convex set that is not necessarily bounded. For a positive integer k, let $F_k = F \cap B(0,k)$. Observe

$$F_k \uparrow F,$$
$$\Big(\bigcup_{x \in B(0,k)\cap F} \mathrm{rb} N_F(x) \Big) \uparrow \Big(\bigcup_{x\in F} \mathrm{rb} N_F(x) \Big).$$

Thus we have

$$\begin{aligned}
\lambda_{n-1}\Big((\bigcup_{x\in F} \mathrm{rb} N_F(x)) \cap \mathbb{S}^{n-1} \Big) &= \lim_{k\to\infty} \lambda_{n-1}\Big((\bigcup_{x\in B(0,k)\cap F} \mathrm{rb} N_F(x)) \cap \mathbb{S}^{n-1} \Big) \\
&= \lim_{k\to\infty} \lambda_{n-1}\Big((\bigcup_{x\in B(0,k)\cap F} \mathrm{rb} N_{\overline{B(0,k)}\cap F}(x)) \cap \mathbb{S}^{n-1} \Big) \\
&\le \lim_{k\to\infty} \lambda_{n-1}\Big((\bigcup_{x\in \overline{B(0,k)}\cap F} \mathrm{rb} N_{\overline{B(0,k)}\cap F}(x)) \cap \mathbb{S}^{n-1} \Big) \\
&= 0,
\end{aligned}$$

where the final equality follows since $\overline{B(0,k)} \cap F$ is a compact convex set. □

We need the following simple proposition. For future reference, let $\pi\colon \mathbf{R}^{n+1} \to \mathbf{R}^n$ be the canonical projection onto the first n coordinates.

Proposition 3.2. *Consider a convex function $f\colon \mathbf{R}^n \to \overline{\mathbf{R}}$ and a point $x \in \mathbf{R}^n$. Then we have the relation,*

$$v \in \mathrm{rb}\partial f(x) \Leftrightarrow (v,-1) \in \mathrm{rb} N_{\mathrm{epi}\, f}(x, f(x)).$$

Proof. Let K denote the normal cone, $N_{\mathrm{epi}\, f}(x,f(x))$. If $\partial f(x) = \emptyset$, then there is no $v \in \mathbf{R}^n$ such that $(v,-1) \in \mathrm{rb}K$, and hence the result holds trivially. Assume that $\partial f(x)$ is nonempty. Observe

$$\mathrm{ri}\, K \not\subset \{y \in \mathbf{R}^{n+1} : y_{n+1} \ge 0\},$$

since otherwise taking closures gives $y_{n+1} \ge 0$ for all $y \in K$, and hence we have $\partial f(x) = \emptyset$, which is a contradiction. Thus there exists a point $y \in \mathrm{ri}\, K$ with $y_{n+1} < 0$. Since K is a cone, we can rescale to get $\hat{y} \in \mathrm{ri}K$ with $\hat{y}_{n+1} = -1$. Hence

$$\mathrm{ri}K \cap \{y \in \mathbf{R}^{n+1} : y_{k+1} = -1\} \neq \emptyset.$$

Using [9, Proposition 2.42], we deduce that

$$\mathrm{ri}(K \cap \{y \in \mathbf{R}^{n+1} : y_{k+1} = -1\}) = \mathrm{ri}K \cap \{y \in \mathbf{R}^{n+1} : y_{k+1} = -1\}. \tag{3.1}$$

Finally, we have

$$\mathrm{ri}\partial f(x) = \pi\Big(\mathrm{ri}(K \cap \{y \in \mathbf{R}^{n+1} : y_{k+1} = -1\})\Big) = \{v : (v,-1) \in \mathrm{ri}K\},$$

where the last equality follows from (3.1). Taking complements, the result follows. □

Theorem 3.3. *Let $f : \mathbf{R}^n \to \overline{\mathbf{R}}$ be a convex function. Then the set*

$$\bigcup_{x\in\mathbf{R}^n} \mathrm{rb}\partial f(x)$$

is Lebesgue null.

Proof. Let

$$H_{-1} := \{x \in \mathbf{R}^{n+1} : x_{n+1} = -1\},$$
$$H_{<} := \{x \in \mathbf{R}^{n+1} : x_{n+1} < 0\},$$
$$K := \Big(\bigcup_{x \in \operatorname{dom} f} \operatorname{rb} N_{\operatorname{epi} f}(x, f(x)) \Big) \cap \mathbb{S}^n \cap H_{<}.$$

Applying Lemma 3.1 to epi$f$, we deduce $\lambda_n(K) = 0$. Consider the mapping

$$\phi : H_{<} \to H_{-1},\ c \mapsto |c_{n+1}|^{-1} c.$$

Observe that ϕ is locally Lipschitz, and therefore by Corollary 2.14, we have $\lambda_n(\phi(K)) = 0$. From Proposition 3.2, we have

$$\pi \circ \phi(K) = \bigcup_{x \in \operatorname{dom} f} \operatorname{rb}\partial f(x).$$

Since π is Lipschitz as well, we deduce $\lambda_n(\bigcup_{x\in\mathbf{R}^n} \operatorname{rb}\partial f(x)) = 0$. A routine argument shows that a set has n-dimensional Hausdorff measure zero if and only if it is Lebesgue null. Hence, the set $\bigcup_{x\in\mathbf{R}^n} \operatorname{rb}\partial f(x)$ is Lebesgue measurable and has Lebesgue measure zero. □

Definition 3.4. Consider a convex function $f\colon \mathbf{R}^n \to \overline{\mathbf{R}}$. A minimizer $x \in \mathbf{R}^n$ of f is said to be *nondegenerate* if it satisfies the property $0 \in \operatorname{ri}\partial f(x)$.

Corollary 3.5. *Let $f\colon \mathbf{R}^n \to \overline{\mathbf{R}}$ be a proper convex function. Consider the collection of perturbed functions $f_v(x) = f(x) - \langle v, x\rangle$, indexed by vectors $v \in \mathbf{R}^n$. Then for a full measure set of vectors $v \in \mathbf{R}^n$, the function f_v has at most one minimizer, which furthermore is nondegenerate.*

Proof. The uniqueness part of the claim is classical, as discussed in the introduction. Thus it is sufficient to show that for a full measure set of vectors $v \in \mathbf{R}^n$, every critical point of f_v is nondegenerate. Indeed, we have $0 \in \operatorname{rb}\partial f_v(x) \Leftrightarrow v \in \operatorname{rb}\partial f(x)$. By Theorem 3.3, the set of vectors v for which $v \in \operatorname{rb}\partial f(x)$ for some $x \in \mathbf{R}^n$ has Lebesgue measure zero, and so the result follows. □

We remark that there are proper convex functions $f\colon \mathbf{R}^n \to \overline{\mathbf{R}}$ with the property that for a full measure set of vectors $v \in \mathbf{R}^n$, the function $f_v(x) = f(x) - \langle v, x\rangle$ has no minimizers, a simple example being $f = \langle a, x\rangle$ for any vector $a \in \mathbf{R}^n$.

3.2. **Extension to lower-$\mathbf{C}^2$ functions.** Having proved Theorem 3.3, we can now easily extend this theorem to a nonconvex situation. In particular, shortly we will show that an analogous statement holds for all lower-$\mathbf{C^2}$ functions.

Theorem 3.6. *Consider a proper function $f\colon \mathbf{R}^n \to \overline{\mathbf{R}}$ with the property that for any point $\bar{x}$ in its domain, there is a neighborhood V around $\bar{x}$ such that on V the function f admits the representation $f = g - \frac{1}{2}\rho|\cdot|^2$, where g is a convex function and ρ is a positive real number. Then the set*

$$\bigcup_{x\in\mathbf{R}^n} \operatorname{rb}\partial f(x)$$

is Lebesgue null.

Remark 3.7. In Theorem 3.6, unlike in the definition of lower-$\mathbf{C^2}$ functions, the domain of f is not required to be an open set and the convex function g in the local representation of f is not required to be finite.

Proof. For each point $x \in \mathrm{dom} f$, consider the neighborhood guaranteed to exist by our assumption on f. This collection of neighborhoods is an open cover of the domain of f and hence has a countable subcover, say $\{V_i\}$. Consider an arbitrary set V_i from this cover. On V_i, we have $f = g - \frac{1}{2}\rho|\cdot|^2$, and hence

$$\bigcup_{x\in V_i} \mathrm{rb}\partial f(x) = \bigcup_{x\in V_i\cap \mathrm{dom}\, f} \mathrm{rb}\partial g(x) - \rho x = \bigcup_{x\in V_i\cap \mathrm{dom}\, f} \mathrm{rb}(\partial g(x) + x) - (\rho+1)x. \tag{3.2}$$

Consider the map

$$H\colon \bigcup_{x\in V_i\cap \mathrm{dom}\, f} \mathrm{rb}(\partial g(x) + x) \to \bigcup_{x\in V_i} \mathrm{rb}\partial f(x),$$

$$c \mapsto c - (\rho+1)(\partial g + I)^{-1}(c).$$

In light of (3.2) and Theorem 2.7, the mapping H is well-defined, surjective, and Lipschitz continuous. Observe that

$$\lambda_n\Big(\bigcup_{x\in V_i\cap \mathrm{dom}\, f} \mathrm{rb}(\partial g(x) + x)\Big) = \lambda_n\Big(\bigcup_{x\in V_i\cap \mathrm{dom}\, f} \mathrm{rb}\partial(g(\cdot) + \frac{1}{2}|\cdot|^2)(x)\Big) = 0,$$

where the last equality follows from convexity of $g+\frac{1}{2}|\cdot|^2$ and Theorem 3.3. From the equation above and Corollary 2.14, we deduce $\lambda_n\Big(\bigcup_{x\in V_i} \mathrm{rb}\partial f(x)\Big) = 0$. Hence, the set $\bigcup_{x\in V_i} \mathrm{rb}\partial f(x)$ is Lebesgue measurable and has Lebesgue measure zero. Finally, since $\{V_i\}$ is a countable cover of $\mathrm{dom} f$, it easily follows from a limiting argument that $\bigcup_{x\in \mathbf{R}^n} \mathrm{rb}\partial f(x)$ is a Lebesgue null set, as was claimed. □

Corollary 3.8. *Let $f\colon O \to \overline{\mathbf{R}}$ be a lower-$\mathbf{C^2}$ function on an open set $O \subset \mathbf{R}^n$. Then the set*

$$\bigcup_{x\in \mathbf{R}^n} \mathrm{rb}\partial f(x).$$

is Lebesgue null.

Proof. From Definition 2.9, f satisfies the conditions of Theorem 3.6, and hence the result follows. □

Definition 3.9. Let $f\colon O \to \mathbf{R}$ be a lower-$\mathbf{C^2}$ function on an open set $O \subset \mathbf{R}^n$. We say that a point $x \in \mathbf{R}^n$ is *critical* for the function f if $0 \in \partial f(x)$, and we call such a critical point x *nondegenerate* if the stronger property $0 \in \mathrm{ri}\, \partial f(x)$ holds.

Corollary 3.10. *Let $f\colon O \to \mathbf{R}$ be a lower-$\mathbf{C^2}$ function on an open set $O \subset \mathbf{R}^n$. Consider the collection of perturbed functions $f_v(x) = f(x) - \langle v, x\rangle$, indexed by vectors $v \in \mathbf{R}^n$. Then for a full measure set of vectors $v \in \mathbf{R}^n$, every critical point of the function f_v is nondegenerate.*

Proof. We have $0 \in \mathrm{rb}\partial f_v(x) \Leftrightarrow v \in \mathrm{rb}\partial f(x)$. By Corollary 3.8, the set of vectors v for which $v \in \mathrm{rb}\partial f(x)$ for some $x \in \mathbf{R}^n$ has Lebesgue measure zero, and so the result follows. □

4. A conjecture

We can formulate Theorem 3.3 in terms of monotone set-valued mappings. See [9, Chapter 12] for the definitions. If we restrict our attention in the theorem to closed proper convex functions f, then Theorem 3.3 is equivalent to the statement that for a maximal cyclically monotone mapping $F\colon \mathbf{R}^n \rightrightarrows \mathbf{R}^n$, the image of the set-valued map $x \mapsto \mathrm{rb}F(x)$ has Lebesgue measure zero (see [9, Theorem 12.25]). We make the following related conjecture.

Conjecture 4.1. *Let $F\colon \mathbf{R}^n \rightrightarrows \mathbf{R}^n$ be a maximal monotone mapping. Then the image of the map $x \mapsto \mathrm{rb}F(x)$ has Lebesgue measure zero; that is, the set*

$$\bigcup_{x\in\mathbf{R}^n} \mathrm{rb}F(x)$$

is Lebesgue null.

A proof of Conjecture 4.1, along with the techniques presented in this paper, might extend the result of Corollary 3.8 to the class of "prox-regular" functions [8].

References

1. H.H. Bauschke, M.S. Macklem, and X. Wang, *Chebyshev sets, Klee sets, and Chebyshev centers with respect to Bregman distances: Recent results and open problems*, arXiv:1003.3127v1 [math.OC] (2010).
2. J. Bolte, A. Daniilidis, and A.S. Lewis, *Generic optimality conditions for semi-algebraic convex programs*, submitted to Mathematics of Operations Research, 2009.
3. J.M. Borwein, W.B. Moors, and X. Wang, *Lipschitz functions with prescribed derivatives and subderivatives*, Nonlinear Analysis **29** (1997), no. 1, 53 – 63. MR1447569 (98j:49019)
4. F.H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley, New York, 1983. MR709590 (85m:49002)
5. D.G. Larman, *On a conjecture of Klee and Martin for convex bodies*, Proc. London Math. Soc. (3) **23** (1971), no. 4, 668–682. MR0293498 (45:2575)
6. G.J. Minty, *Monotone (nonlinear) operators in Hilbert space*, Duke Math. J. **29** (1962), 341–346. MR0169064 (29:6319)
7. G. Pataki and L. Tunçel, *On the generic properties of convex optimization problems in conic form*, Math. Programming **89** (2001), 449–457. MR1814549 (2001k:90049)
8. R.A. Poliquin and R.T. Rockafellar, *Prox-regular functions in variational analysis*, Trans. Amer. Math. Soc. **348** (1996), no. 5, 1805–1838. MR1333397 (96h:49039)
9. R.T. Rockafellar and R. J-B. Wets, *Variational Analysis*, Springer, Berlin, 1998. MR1491362 (98m:49001)
10. C.A. Rogers, *Hausdorff Measures*, second ed., Cambridge University Press, Cambridge, 1998. MR1692618 (2000b:28009)

School of Operations Research and Information Engineering, Cornell University, Ithaca, New York 14853

E-mail address: dd379@cornell.edu

School of Operations Research and Information Engineering, Cornell University, Ithaca, New York 14853

URL: ttp://people.orie.cornell.edu/~aslewis/