Quadratic growth and critical point stability of semi-algebraic functions

D. Drusvyatskiy · A.D. Ioffe

Received: date / Accepted: date

Abstract We show that quadratic growth of a semi-algebraic function is equivalent to strong metric subregularity of the subdifferential — a kind of stability of generalized critical points. In contrast, this equivalence can easily fail outside of the semi-algebraic setting.

Keywords subdifferentials · quadratic growth · strong metric subregularity · semi-algebraic

Mathematics Subject Classification (2010) 49J53 · 14P10 · 54C60 · 65K10 · 90C31 · 49J52

1 Introduction

Quadratic growth conditions of extended-real-valued functions play a central role in nonlinear optimization, both for convergence analysis of algorithms and for perturbation theory. See for example [6, 15, 21]. Classically, a point \( \bar{x} \) is a
strong local minimizer of a function $f$ on $\mathbb{R}^n$ if there exists a constant $\alpha > 0$ and a neighborhood $U$ of $\bar{x}$ such that the inequality

$$f(x) \geq f(\bar{x}) + \alpha \|x - \bar{x}\|^2 \quad \text{holds for all } x \in U. \quad (1.1)$$

Here $\| \cdot \|$ denotes the Euclidean norm on $\mathbb{R}^n$. For $C^2$-smooth functions $f$ this condition simply amounts to requiring $\bar{x}$ to be a critical point with the Hessian $\nabla^2 f(\bar{x})$ being positive definite.

What condition would then entail second order growth of the function near a minimizer, when derivatives do not exist? In a seminal paper [1], Aragón Artacho and Geoffroy gave an elegant and rather unexpected answer to this question for convex functions: a minimizer $\bar{x}$ of a convex function $f$ on a Hilbert space is a strong local minimizer if and only if the convex subdifferential $\partial f$ is strongly subregular at $(0, 0)$. The latter means that there exists a constant $\kappa \geq 0$ and a neighborhood $V$ of $\bar{x}$ so that the inequality

$$\|x - \bar{x}\| \leq \kappa \cdot d(0; \partial f(x)) \quad \text{holds for all } x \in V, \quad (1.2)$$

where $d(0; \partial f(x))$ denotes the minimal norm of subgradients $v \in \partial f(x)$. Hence $\bar{x}$ being a strong minimizer amounts to requiring existence of a linear bound on the distance of a putative solution $x$ to $\bar{x}$ in terms of the “distance to criticality” $d(0; \partial f(x))$ — an appealing property. An entirely analogous characterization is valid for $C^2$-smooth functions.

In the current work, we aim to address the question: to what extent is such a transparent and mathematically rigorous interpretation of quadratic growth true more generally? A partial answer was given in [2] and [11]. Namely, in [2] the authors extended the aforementioned equivalence to convex functions on all Banach spaces. More importantly for the purposes of our paper, pushing beyond convexity, the authors observed that for a local minimizer $\bar{x}$ of a lower semi-continuous function (not necessarily convex) on an Asplund space, strong subregularity now of the limiting Fréchet subdifferential always entails that $\bar{x}$ is a strong local minimizer. A sharper version of the latter result was also proved independently in [11]. In Section 4 we show that this implication holds practically unconditionally — for any lower-semicontinuous function on any Banach space and any subdifferential which is “trusted” on the space.

On the other hand, since we are interested in an interpretation of quadratic growth, it is the converse implication that we are after. Evidently, the subdifferential may fail to be strongly subregular at a strong local minimizer; case in point, for $f(x) := 2x^2 + \frac{1}{2}x^2 \sin \left(\frac{1}{x}\right)$. See Example 3.3 for more details. Such badly behaved functions, however, rarely appear in practice. The usual way to eliminate such functions from consideration is to restrict attention to favorable function classes, such as to those that are amenable [17] or prox-regular [18].

In the current work, we take a different approach. We consider prototypical nonpathological functions, namely those that are semi-algebraic — meaning that their epigraphs are composed of finitely many sets, each defined by finitely many polynomial inequalities. See [4][10] for more on semi-algebraic
geometry. This class subsumes neither the case of convex nor the case of $C^2$-smooth functions. Nevertheless, it has great appeal. Semi-algebraic functions are common, they are easy to recognize (via the Tarski-Seidenberg Theorem on quantifier elimination), and in sharp contrast to the usual favorable function classes of variational analysis, semi-algebraic functions do not need to be Clarke-regular \cite{20}. For a discussion of the role of semi-algebraic functions in nonsmooth optimization, see \cite{12}. We will show that for semi-algebraic functions, satisfying a minimal continuity condition, strong local minimality and strong subregularity of the subdifferential are equivalent properties (Theorem \ref{thm:main}) — the principle result of the manuscript. Our argument, is entirely geometric and uses only the most basic tools of semi-algebraic geometry. To be concrete, we state our results for semi-algebraic functions. Analogous results, with essentially identical proofs, hold for functions definable in an “o-minimal structure” and, more generally, for “tame” functions. See \cite{24} for more on the subject.

The outline of the manuscript is as follows. In Section 2 we record basic elements of variational analysis and semi-algebraic geometry that we will need throughout the manuscript. In Section 3 we prove that in the semi-algebraic setting strong metric subregularity of the subdifferential at a local minimizer is equivalent to strong local minimality. In Section 4 we show that strong metric subregularity of the subdifferential at a local minimizer, for virtually any subdifferential of interest, implies a quadratic growth condition.

2 Preliminaries

2.1 Some elements of Variational Analysis in $\mathbb{R}^n$

In this subsection, we summarize some of the fundamental tools used in variational analysis and nonsmooth optimization. We refer the reader to the monographs \cite{7,8,14,16,20} for more details. Unless otherwise stated, we follow the terminology and notation of \cite{20}.

Throughout $\mathbb{R}^n$ will denote the $n$-dimensional Euclidean space, with the inner-product written as $\langle \cdot, \cdot \rangle$. We will denote the induced norm by $\| \cdot \|$. The closed ball centered at $x \in \mathbb{R}^n$ of radius $r$ will be denoted by $B_r(x)$ while the closed unit ball will be denoted by $B$. The distance of a point $x \in \mathbb{R}^n$ to a set $Q \subset \mathbb{R}^n$ is the quantity

$$d(x; Q) := \inf_{y \in Q} \| x - y \|.$$ 

\footnote{This theorem was proved in September 2012 but we had to delay further work over this subject as there were two more manuscripts (in a more advanced state) that demanded our attention. Shortly after that, the first author in conversations with Nghia became aware of the very recent results of Nghia-Mordukhovich leading to \cite{11}.}
A set-valued mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is a mapping assigning to each point $x \in \mathbb{R}^n$ a subset $F(x) \subset \mathbb{R}^m$. The \textit{domain} and \textit{graph} of $F$ are defined by
\[
\text{dom } F := \{x \in \mathbb{R}^n : F(x) \neq \emptyset \}, \quad \text{gph } F := \{(x,y) \in \mathbb{R}^n \times \mathbb{R}^m : y \in F(x)\},
\]
respectively.

**Definition 2.1 (Strong metric subregularity)**
Consider a mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ and a pair $(\bar{x}, \bar{y}) \in \text{gph } F$. We say that $F$ is \textit{strongly metrically subregular} at $(\bar{x}, \bar{y})$ with modulus $\kappa \geq 0$ if there is a real number $\varepsilon > 0$ such that the inequality
\[
\|x - \bar{x}\| \leq \kappa d(\bar{y}; F(x)) \quad \text{holds for all } x \in B_\varepsilon(\bar{x}).
\]

Observe that in particular, strong metric subregularity of $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ at $(\bar{x}, \bar{y})$ implies that $\bar{x}$ is a locally isolated point of $F^{-1}(\bar{y})$. Given a set $Q \subset \mathbb{R}^n$ and a point $\bar{x} \in Q$, we will denote by $\omega(\|x - \bar{x}\|)$ for $x \in Q$ a term with the property that
\[
\frac{\omega(\|x - \bar{x}\|)}{|x - \bar{x}|} \to 0 \quad \text{when } x \not\in Q \text{ with } x \neq \bar{x}.
\]

**Definition 2.2 (Normal cones)**
Consider a set $Q \subset \mathbb{R}^n$ and a point $\bar{x} \in Q$.

- The \textit{Fréchet normal cone} to $Q$ at $\bar{x}$, denoted $\tilde{N}_Q(\bar{x})$, consists of all vectors $v \in \mathbb{R}^n$ satisfying
  \[
  \langle v, x - \bar{x} \rangle \leq \omega(\|x - \bar{x}\|) \quad \text{for } x \in Q.
  \]

- The \textit{limiting normal cone} to $Q$ at $\bar{x}$, denoted $N_Q(\bar{x})$, consists of all vectors $v \in \mathbb{R}^n$ such that there are sequences $x_i \not\in Q$, $\bar{x}$ and $v_i \to v$ with $v_i \in \tilde{N}_Q(x_i)$.

Given any set $Q \subset \mathbb{R}^n$ and a mapping $f : Q \to \tilde{Q}$, where $\tilde{Q} \subset \mathbb{R}^m$, we say that $f$ is $C^p$-\textit{smooth} (for $p = 1, \ldots, \infty$) if for each point $\bar{x} \in Q$, there is a neighbourhood $U$ of $\bar{x}$ and a $C^p$-smooth mapping $\hat{f} : \mathbb{R}^n \to \mathbb{R}^m$ that agrees with $f$ on $Q \cap U$. If a $C^p$-smooth function $f$ is bijective and its inverse is also $C^p$-smooth, then we say that $f$ is a $C^p$-\textit{diffeomorphism}.

**Definition 2.3 (Manifolds)** Consider a set $M \subset \mathbb{R}^n$. We say that $M$ is a $C^p$ \textit{manifold} (or “embedded submanifold”) of dimension $r$ if for each point $\bar{x} \in M$, there is an open neighbourhood $U$ around $\bar{x}$ such that $M \cap U = F^{-1}(0)$, where $F : U \to \mathbb{R}^{n-r}$ is a $C^p$-smooth mapping with $\nabla F(\bar{x})$ of full rank.

If $M$ is a $C^1$ manifold, then for every point $\bar{x} \in M$, the normal cones $\tilde{N}_M(\bar{x})$ and $N_M(\bar{x})$ are equal to the normal space to $M$ at $\bar{x}$, in the sense of differential geometry [20, Example 6.8].
Consider the extended real line $\mathbb{R} := \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$. We say that an extended-real-valued function is proper if it is never $\{-\infty\}$ and is not always $\{+\infty\}$. For a function $f: \mathbb{R}^n \to \mathbb{R}$, we define the domain of $f$ to be
\[ \text{dom } f := \{ x \in \mathbb{R}^n : f(x) < +\infty \}, \]
and we define the epigraph of $f$ to be
\[ \text{epi } f := \{ (x, r) \in \mathbb{R}^n \times \mathbb{R} : r \geq f(x) \}. \]
A function $f: \mathbb{R}^n \to \mathbb{R}$ is lower-semicontinuous (or lsc for short) at $\bar{x}$ if the inequality $\liminf_{x \to \bar{x}} f(x) \geq f(\bar{x})$ holds.

The following is the key notion we study in the current work.

**Definition 2.4 (Strong local minimizer)** A point $\bar{x}$ is a strong local minimizer of a function $f: \mathbb{R}^n \to \mathbb{R}$ if there exists a constant $\alpha > 0$ and a neighborhood $U$ of $\bar{x}$ such that the inequality $f(x) \geq f(\bar{x}) + \alpha \|x - \bar{x}\|^2$ holds for all $x \in U$.

Functional counterparts of normal cones are subdifferentials.

**Definition 2.5 (Subdifferentials)** Consider a function $f: \mathbb{R}^n \to \mathbb{R}$ and a point $\bar{x} \in \mathbb{R}^n$ where $f$ is finite. The regular and limiting subdifferentials of $f$ at $\bar{x}$, respectively, are defined by
\[ \hat{\partial} f(\bar{x}) = \{ v \in \mathbb{R}^n : (v, -1) \in \hat{N}_{\text{epi } f}(\bar{x}, f(\bar{x})) \}, \]
\[ \partial f(\bar{x}) = \{ v \in \mathbb{R}^n : (v, -1) \in N_{\text{epi } f}(\bar{x}, f(\bar{x})) \}. \]
The horizon subdifferential of $f$ at $\bar{x}$ is the set
\[ \partial^\infty f(\bar{x}) = \{ v \in \mathbb{R}^n : (v, 0) \in N_{\text{epi } f}(\bar{x}, f(\bar{x})) \}. \]
For $x$ such that $f(x)$ is not finite, we follow the convention that $\hat{\partial} f(x) = \partial f(x) = \partial^\infty f(x) = \emptyset$. The subdifferentials $\hat{\partial} f(\bar{x})$ and $\partial f(\bar{x})$ generalize the classical notion of gradient. In particular, for $C^1$-smooth functions $f$ on $\mathbb{R}^n$, these three subdifferentials consist only of the gradient $\nabla f(x)$ for each $x \in \mathbb{R}^n$.

For convex $f$, these subdifferentials coincide with the convex subdifferential. The horizon subdifferential $\partial^\infty f(\bar{x})$ plays an entirely different role — it detects horizontal “normals” to the epigraph — and it plays a decisive role in subdifferential calculus. See [20, Corollary 10.9] or [19] for more details.

**Theorem 2.6 (Fermat & sum rules)** Consider an lsc function $f: \mathbb{R}^n \to \mathbb{R}$ and a closed set $Q \subset \mathbb{R}^n$. If $\bar{x}$ is a local minimizer of $f$ on $Q$ and the qualification condition
\[ \partial^\infty f(\bar{x}) \cap N_Q(\bar{x}) = \{0\} \]
is valid, then the inclusion $0 \in \partial f(\bar{x}) + N_Q(\bar{x})$ holds.
We will also need the following definition in order to guarantee that the subdifferential $\partial f$ adequately reflects properties of the function $f$ itself.

**Definition 2.7 (Subdifferential continuity)** A function $f: \mathbb{R}^n \to \mathbb{R}$ is subdifferentially continuous at $\bar{x}$ for $\bar{v} \in \partial f(\bar{x})$ if for any sequences $x_i \to \bar{x}$ and $v_i \to \bar{v}$, with $v_i \in \partial f(x_i)$, it must be the case that $f(x_i) \to f(\bar{x})$.

Subdifferential continuity of a function $f$ at $\bar{x}$ for $\bar{v}$ was introduced in [18, Definition 1.14], and it amounts to requiring the function $(x, v) \mapsto f(x)$, restricted to $\text{gph} \partial f$, to be continuous in the usual sense at the point $(\bar{x}, \bar{v})$.

For example, any strongly amenable (in particular, lsc and convex) function is subdifferentially continuous [18, Proposition 2.5].

### 2.2 Semi-algebraic geometry

A semi-algebraic set $S \subset \mathbb{R}^n$ is a finite union of sets of the form

$$\{ x \in \mathbb{R}^n : P_1(x) = 0, \ldots, P_k(x) = 0, Q_1(x) < 0, \ldots, Q_l(x) < 0 \},$$

where $P_1, \ldots, P_k$ and $Q_1, \ldots, Q_l$ are polynomials in $n$ variables. In other words, $S$ is a union of finitely many sets, each defined by finitely many polynomial equalities and inequalities. A function $f: \mathbb{R}^n \to \mathbb{R}$ is semi-algebraic if $\text{epi} f \subset \mathbb{R}^{n+1}$ is a semi-algebraic set. For an extensive discussion on semi-algebraic geometry, see the monographs of Basu-Pollack-Roy [3], Lou van den Dries [23], and Shiota [22]. For a quick survey, see the article of van den Dries-Miller [24] and the surveys of Coste [9,10]. Unless otherwise stated, we follow the notation of [24] and [10].

A fundamental fact about semi-algebraic sets is provided by the Tarski-Seidenberg Theorem [10, Theorem 2.3]. It states that the image of any semi-algebraic set $Q \subset \mathbb{R}^n$, under a projection to any linear subspace of $\mathbb{R}^n$, is a semi-algebraic set. From this result, it follows that a great many constructions preserve semi-algebraicity. In particular, for a semi-algebraic function $f: \mathbb{R}^n \to \mathbb{R}$, it is easy to see that all the subdifferential set-valued mappings are semi-algebraic. See for example [12, Proposition 3.1].

The following two well-known theorems will be of great use for us.

**Theorem 2.8 (Semi-algebraic monotonicity)** Consider a semi-algebraic function $f: (a, b) \to \mathbb{R}$. Then there are finitely many points $a = t_0 < t_1 < \ldots < t_k = b$ such that the restriction of $f$ to each interval $(t_i, t_{i+1})$ is $C^2$-smooth and either strictly monotone or constant.

**Theorem 2.9 (Semi-algebraic selection)** Consider a semi-algebraic set-valued mapping $F: \mathbb{R}^n \rightrightarrows \mathbb{R}^m$. Then there is a single-valued semi-algebraic mapping $f: \text{dom } F \to \mathbb{R}^m$ satisfying $f(x) \in F(x)$ for all $x \in \text{dom } F$.

The proof of the following result appears implicitly in [5, Proposition 4]. We outline an argument for completeness.
Lemma 2.10 (Semi-algebraic chain rule)
Consider a semi-algebraic function $f: \mathbb{R}^n \to \mathbb{R}$ and a semi-algebraic curve $x: [0,T) \to \operatorname{dom} f$. Then there exists $\varepsilon > 0$ so that both $x$ and $f \circ x$ are $C^2$-smooth on $(0, \varepsilon)$ and for any $t \in (0, \varepsilon)$ we have
\[
\begin{align*}
v \in \partial f(x(t)) & \quad \implies \quad \langle v, \dot{x}(t) \rangle = (f \circ x)'(t), \\
v \in \partial^\infty f(x(t)) & \quad \implies \quad \langle v, \ddot{x}(t) \rangle = 0.
\end{align*}
\]

Proof By Theorem 2.8, there exists $\varepsilon > 0$ such that both $x$ and $f \circ x$ are $C^2$-smooth on $(0, \varepsilon)$. Let $\phi := f \circ x$ and define
\[
\mathcal{M} := \{(x(t), \phi(t)) : t \in (0, \varepsilon)\}.
\]

Clearly $\mathcal{M}$ is a semi-algebraic $C^2$-submanifold of the epigraph $\operatorname{epi} f$. Taking if necessary a smaller $\varepsilon$, we can be sure that there is a Whitney $C^2$-stratification of $\operatorname{epi} f$ such that $\mathcal{M}$ is a stratum; see for example [24, Theorem 4.8]. Then for any real $t \in (0, \varepsilon)$, the inclusion
\[
N_{\operatorname{epi} f}(x(t), \phi(t)) \subset N_{\mathcal{M}}(x(t), \phi(t))
\]
holds. On the other hand, we have the representation
\[
N_{\mathcal{M}}(x(t), \phi(t)) = \{(v, \alpha) : \langle v, \ddot{x}(t) \rangle + \alpha \phi'(t) = 0\}.
\]
The result follows immediately. \qed

3 Strong subregularity and growth of semi-algebraic functions

We are now ready to prove the main result of the current work. We should note that the implication $1 \Rightarrow 2$ in the theorem below is true without semi-algebraicity (see Theorem [4.3], [11, Theorem 3.1], [2, Theorem 2.1]). The semi-algebraic setting, on the other hand, allows us to provide an appealing geometric argument of this result. The implication $2 \Rightarrow 1$, in contrast, may easily fail when the function in question is not semi-algebraic; see Example 3.3.

Theorem 3.1 (Strong metric subregularity and quadratic growth)
Consider an lsc, semi-algebraic function $f: \mathbb{R}^n \to \mathbb{R}$ and a point $\bar{x} \in \mathbb{R}^n$ that is a local minimizer of $f$. Consider the following two statements:

1. Subdifferential $\partial f$ is strongly metrically subregular at $(\bar{x}, 0)$ with modulus $\kappa$,
2. There exist real numbers $\alpha > 0$ and $\varepsilon > 0$ such that the inequality
\[
f(x) \geq f(\bar{x}) + \frac{\alpha}{2} \|x - \bar{x}\|^2
\]
holds for all $x \in B_\varepsilon(\bar{x})$.

Then the implication $1 \Rightarrow 2$ holds where we can choose $\alpha$ arbitrarily in $\in (0, \kappa^{-1})$. The converse implication $2 \Rightarrow 1$ holds provided that $f$ is subdifferentially continuous at $\bar{x}$ for $\bar{v} = 0$. 

Proof Without loss of generality, assume \( \bar{x} = 0 \) and \( f(\bar{x}) = 0 \).

1 \( \Rightarrow \) 2: Suppose that the subdifferential \( \partial f \) is strongly metrically subregular at \((\bar{x}, 0)\) with modulus \( \kappa \) and define the function

\[
\varphi(t) := \inf\{ f(x) : \|x\| = t \}.
\]

Standard arguments using quantifier elimination show that \( \varphi \) is semi-algebraic. It follows from Theorem 2.3 that \( \varphi \) is \( C^2 \)-smooth for all sufficiently small \( t \). If \( \lim_{t \to 0} \varphi(t) > 0 \), or \( \varphi(t) \to 0 \) but \( \lim_{t \to 0} \varphi(t) \) is positive, then the theorem obviously holds. So we assume \( \varphi(t) \to 0 \) and \( \varphi(t) \to 0 \) as \( t \to 0 \).

Note that the infimum in the definition of \( \varphi \) is attained if \( \varphi(t) \) is finite. By Theorem 2.3, there is a semi-algebraic mapping \( x(t) \) such that \( f(x(t)) = \varphi(t) \) for sufficiently small \( t \).

Applying Lemma 2.10, we deduce that there is a real \( \varepsilon > 0 \) so that both \( x \) and \( \varphi \) are \( C^2 \)-smooth on \((0, \varepsilon)\) and for any \( t \in (0, \varepsilon) \) we have

\[
v \in \partial f(x(t)) \implies \langle v, \dot{x}(t) \rangle = \varphi'(t), \quad (3.1)
\]

Observe

\[
\langle x(t), \dot{x}(t) \rangle = \frac{1}{2} \frac{d}{dt} \|x(t)\|^2 = t,
\]

and hence the qualification condition

\[
\partial^\infty f(x(t)) \cap N_{\{x : \|x\| = t\}}(x(t)) = \{0\}
\]

holds for all \( t \in (0, \varepsilon) \). Consequently, since \( x(t) \) minimizes \( f \) subject to \( \|x\| = t \), applying Theorem 2.6, we deduce that there is a real number \( \lambda(t) \) satisfying

\[
\lambda(t)x(t) \in \partial f(x(t)). \quad (3.2)
\]

By strong subregularity, we have \( \|\lambda(t)x(t)\| \geq \kappa^{-1}\|x(t)\| \), that is \( |\lambda(t)| \geq \kappa^{-1} \). Finally, combining \((3.1)\) and \((3.2)\), we get

\[
\dot{\varphi}(t) = \lambda(t)\langle x(t), \dot{x}(t) \rangle = \lambda(t)\frac{1}{2} \frac{d}{dt} \|x(t)\|^2 = \lambda(t)t.
\]

Since \( \bar{x} \) is a local minimizer, we have \( \dot{\varphi}(t) \geq 0 \). Consequently, we obtain \( \lambda(t) \geq 0 \) and \( \lambda(t) \geq \kappa^{-1} \), and hence \( \varphi(t) \geq \frac{\kappa^{-1}}{2}t^2 \).

Suppose now that \( f \) is subdifferentially continuous at \( \bar{x} \) for \( \bar{v} = 0 \).

2 \( \Rightarrow \) 1: Assume that 2 holds. Suppose also for the sake of contradiction that \( \partial f \) is not strongly metrically subregular at \((\bar{x}, 0)\). Define the function

\[
H(t) := \arg\min \left\{ \frac{d(0, \partial f(x))}{t} : \|x\| = t \right\}.
\]

Then an application of Theorem 2.4 yields a semi-algebraic path \( x : (0, \varepsilon) \to \mathbb{R}^n \) satisfying \( x(t) \in H(t) \) for all \( t \). Applying Theorem 2.4 to the mapping

\[
t \mapsto \arg\min \{ \|v\| : v \in \partial f(x(t)) \},
\]
yields a semi-algebraic path \( v: (0, \varepsilon) \to \mathbb{R}^n \) satisfying \( v(t) \in \partial f(x(t)) \).

Observe
\[
\lim_{t \searrow 0} x(t) = 0, \quad \lim_{t \searrow 0} \|v(t)\| = +\infty, \quad \lim_{t \searrow 0} v(t) = 0.
\]
Clearly we may extend \( x \) and \( v \) continuously to \([0, \varepsilon)\). Decreasing \( \varepsilon \), we may assume that on the interval \((0, \varepsilon)\) the composition \( f \circ x \) is \( C^2 \)-smooth, \(|v(t)| \) is non-increasing, and that \( \dot{x}(t) \) is nonzero. Moreover since \( f \) is subdifferentially continuous at \( \bar{x} \) for 0, we have \( \lim_{t \searrow 0} f(x(t)) = 0 \). Hence the composition \( f \circ x \) is continuous on \([0, \varepsilon)\).

We now reparametrize \( x \) by arclength. Namely, since \( x \) is semi-algebraic and bounded, an application of Theorem 2.8 implies that \( x \) has finite length
\[
L := \int_0^\varepsilon \|\dot{x}(t)\| \, dt.
\]
Define the function \( s: [0, \varepsilon] \to [0, L] \) by setting
\[
s(t) := \int_0^t \|\dot{x}(t)\| \, dt.
\]
Let \( y(\tau) := x(s^{-1}(\tau)) \) and \( \omega(\tau) := v(s^{-1}(\tau)) \). Clearly \( y \) is \( C^2 \)-smooth on \((0, L)\) and satisfies \( \|\dot{y}(\tau)\| = 1 \) for all \( \tau \in (0, L) \).

We successively conclude, in part using Theorem 2.9,
\[
\alpha |y(\tau)|^2 \leq f(y(\tau)) = \int_0^{\tau} \frac{d}{dr} (f \circ y)(r) \, dr = \int_0^{\tau} \langle \dot{y}(r), \omega(r) \rangle \, dr \\
\leq \int_0^{\tau} \|\dot{y}(r)\| \cdot \|\omega(r)\| \, dr = \int_0^{\tau} \|\omega(r)\| \, dr \\
\leq \tau \|\omega(\tau)\|.
\]
We deduce
\[
0 < \alpha \leq \frac{\tau}{\|y(\tau)\|} \|\omega(\tau)\|. \tag{3.3}
\]
Now the mean value theorem for vector-valued functions implies
\[
y(\tau) = \left( \int_0^1 \dot{y}(h\tau) \, dh \right) \tau.
\]
Hence we have
\[
\sqrt{n} \cdot \frac{\|y(\tau)\|}{\tau} \geq \frac{\|y(\tau)\|}{\tau} = \sum_{i=1}^{n} \left| \int_0^1 \dot{y}_i(h\tau) \, dh \right|.
\]
Since \( x \) is semi-algebraic, so is the derivative \( \dot{x} \). Applying Theorem 2.8, we deduce that there exists \( \delta > 0 \) such that each function \( \dot{y}_i \) has a constant sign on \((0, \delta)\). Hence for \( \tau \in (0, \delta) \), we obtain
\[
\sum_{i=1}^{n} \left| \int_0^1 \dot{y}_i(h\tau) \, dh \right| = \int_0^{1} \sum_{i=1}^{n} |\dot{y}_i(h\tau)| \, dh \geq 1.
\]
We conclude that the quantity $\frac{\tau}{\|w(\tau)\|}$ is bounded for small $\tau$. Letting $\tau$ tend to zero in $(3.3)$, we arrive at a contradiction. \hfill \Box

The following examples show that the implication $2 \Rightarrow 1$ of Theorem 3.1 may easily fail in absence of subdifferential continuity or semi-algebraicity.

**Example 3.2 (Equivalence fails without subdifferential continuity)**

Consider the lsc, semi-algebraic function $f: \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} 
1 + x^4, & x < 0, \\
0, & x \geq 0.
\end{cases}$$

Observe that $f$ is not subdifferentially continuous at $\bar{x} = 0$ for $\bar{v} = 0$. Clearly 0 is a strong local minimizer of $f$. However, one can easily check that $\partial f$ is not strongly metrically subregular at $(0, 0)$.

**Example 3.3 (Equivalence fails without semi-algebraicity)**

Consider the function $f: \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) := 2x^2 + \frac{1}{2} x^2 \sin \left( \frac{1}{x^2} \right).$$

It is easy to check that the origin is a strong local minimizer of $f$, while $\partial f$ is not strongly metrically subregular at $(0, 0)$.

### 4 Strong metric subregularity entails quadratic growth

In this section, we show that strong subregularity of a subdifferential, for virtually any subdifferential of interest, at a local minimizer implies strong local minimality. We begin with some notation. In this section $\mathcal{X}$ will denote a Banach space with norm $\| \cdot \|$, while the dual space of $\mathcal{X}$ will be written as $\mathcal{X}^\ast$. The closed unit ball in $\mathcal{X}$ will be denoted by $B$, while the closed unit ball of $\mathcal{X}^\ast$ will be written as $B_{\mathcal{X}^\ast}$.

There are several types of subdifferentials used in variational analysis in Banach spaces (proximal, Fréchet, limiting Fréchet, Dini-Hadamard, G-subdifferential, generalized gradient). We do not need a general definition here: an interested reader may look into [13]. The important point is that most of them can be effectively used only in certain classes of Banach spaces. This observation is formally represented by the following definition playing an important role in the general theory.

**Definition 4.1 (Trustworthy spaces)** A subdifferential $\partial$ can be trusted on a Banach space $\mathcal{X}$ if for any lsc function $f: \mathcal{X} \to \overline{\mathbb{R}}$, any point $\bar{x} \in \text{dom } f$ and any function $g: \mathcal{X} \to \overline{\mathbb{R}}$ which is convex continuous near $\bar{x}$ the following holds: if $f + g$ attains a local minimum at $\bar{x}$, then for any $\varepsilon > 0$ there are $x, u, x^\ast, u^\ast$ such that both $x$ and $u$ are $\varepsilon$-close to $\bar{x}$, $f(x)$ is $\varepsilon$-close to $f(\bar{x})$, $x^\ast \in \partial f(x)$, $u^\ast \in \partial g(u)$ and $\|x^\ast + u^\ast\| < \varepsilon$. 


The Fréchet and the limiting Fréchet subdifferential can be trusted on Asplund spaces and only on them; the Dini-Hadamard subdifferential can be trusted on Gâteaux smooth spaces; the G-subdifferential and the generalized gradient of Clarke can be trusted on all Banach spaces. See [13] for details.

Let us finally agree to say that the exact calculus holds for a given subdifferential \( \partial \) in a given Banach space \( X \) if the inclusion

\[
\partial(f + g)(x) \subset \partial f(x) + \partial g(x)
\]

holds whenever \( f: X \to \mathbb{R} \) is lsc and \( g: X \to \mathbb{R} \) is Lipschitz continuous near \( x \). This is the case when \( \partial \) is a "robust" subdifferential: limiting Fréchet subdifferential in an Asplund space, G-subdifferential or generalized gradient on any Banach space.

The notion of strong metric subregularity trivially generalizes to set-valued mapping between metric spaces. We will need the following simple result asserting that small Lipschitz perturbations do not destroy strong metric subregularity.

The following is the main result of the current section. We should mention that the first part of the argument is an elaboration of the proof of [2, Theorem 2.1], while the second part is to a large extent a clarification of the proof of [11, Theorem 3.1] for the case of strong subregularity, making it shorter and technically simpler.

**Theorem 4.3 (Strong metric subregularity of subdifferentials)**

Consider an lsc function \( f: X \to \mathbb{R} \) defined on a Banach space \( X \), which has a local minimum at \( \bar{x} \in \text{dom } f \). Let \( \partial \) be a subdifferential trusted on \( X \) and
suppose that \( \partial f \) is strongly subregular at \((\bar{x}, 0)\) with modulus \( \kappa > 0 \). Then for any real number \( \alpha \in (0, \frac{1}{\kappa}) \) there exists \( \varepsilon > 0 \) such that the inequality

\[
f(x) \geq f(\bar{x}) + \frac{\alpha}{2} \|x - \bar{x}\|^2 \quad \text{holds for all } x \in B_{\varepsilon}(\bar{x}).
\]

Moreover, if the exact calculus holds, then the conclusion is valid for any \( \alpha \in (0, \frac{1}{\kappa}) \).

**Proof** We assume for simplicity that \( \bar{x} = 0 \) and \( f(\bar{x}) = 0 \). We must estimate the lower bound of \( \alpha > 0 \) with the property that there is a sequence \( u_k \) satisfying

\[
u_k \to 0, \quad u_k \neq 0, \quad f(u_k) \leq \frac{\alpha}{2} \|u_k\|^2
\]  

(4.1)

So suppose there is such a sequence and let \( \frac{\kappa}{\alpha} = \frac{\lambda}{\kappa} \) for some \( \lambda \in (0, 1) \). By Ekeland’s principle for any \( k \) there exists \( w_k \) such that \( \|w_k - u_k\| \leq \lambda \|u_k\| \), \( f(w_k) \leq f(u_k) \) and \( q_k(x) = f(x) + \frac{\lambda}{\kappa} \|u_k\| \|x - w_k\| \) attains a minimum at \( w_k \).

We see that \( g_k \) is a sum of an lsc and a convex continuous function. As \( \partial \) is trusted on \( X \), we can find for any \( k \) and any \( \varepsilon_k > 0 \) a point \( x_k \) such that \( |f(x_k) - f(w_k)| < \frac{\lambda}{\kappa} \|u_k\| \), \( \|x_k - w_k\| < \frac{\lambda}{\kappa} \|u_k\| \) and \( 0 \in \partial f(x_k) + \frac{\lambda + \varepsilon_k}{\kappa} \|u_k\| B_{X^*} \).

We deduce

\[
d(0, \partial f(x_k)) \leq \frac{\lambda + \varepsilon_k}{\kappa} \|u_k\|.
\]

Observe \( \|x_k\| \geq \|u_k\| - \|x_k - u_k\| \geq (1 - (\frac{\lambda}{\kappa}) \|u_k\| \), that is

\[
d(0, \partial f(x_k)) \leq \frac{\lambda + \varepsilon_k}{\kappa - (\kappa \lambda + \varepsilon_k)} \|x_k\|.
\]

This inequality would be in contradiction with the assumed strong subregularity of \( \partial f \) if \( (\lambda + \varepsilon_k)(\kappa - (\kappa \lambda + \varepsilon_k))^{-1} < \kappa^{-1} \). The latter would definitely happen if \( \varepsilon_k \to 0 \) and \( \lambda < 1 - \lambda \). The latter gives \( \lambda < 1/2 \). Thus, if \( \alpha < \frac{1}{2\kappa} \), there cannot be a sequence \( u_k \) satisfying (4.1).

Suppose now that exact calculus holds for the subdifferential \( \partial \). Set \( a_0 := \frac{1}{2\kappa} \) and define

\[
f_1(x) := f(x) - \frac{\kappa}{2} \|x\|^2.
\]

Then \( \partial f_1(x) \subset \partial f(x) + a_0 \|x\| \|\partial f\| \|x\| \). By Lemma 4.2 the subdifferential \( \partial f_1 \) is strongly subregular at \((0, 0)\) with modulus \((\kappa^{-1} - a_0)^{-1} \). It follows according to what has just been proved that, for \( \alpha = (\kappa^{-1} - a_0) \) there exists \( \tau_1 > 0 \) satisfying \( f_1(x) \geq \frac{\kappa}{2} \|x\|^2 \) for all \( x \in \tau_1 B \).

Hence for such \( x \), we have \( f(x) \geq \left( \frac{\kappa}{2} + \frac{1}{2} \right) \frac{\kappa}{2} \|x\|^2 \). Repeating this procedure inductively, we obtain the result.

\[\square\]

**Remark 4.4** The argument of the theorem above easily extends to the case when \( \partial \) is a Dini-Hadamard (respectively Fréchet) subdifferential and the norm in \( X \) is Gâteaux (respectively Fréchet) smooth.

**Acknowledgements** Much of the current manuscript was written while the first author was visiting the ECE department at Northwestern University. The first author thanks the department, and Jorge Nocedal in particular, for the hospitality.
References

1. F.J. Aragón Artacho and M.H. Geoffroy. Characterization of metric regularity of subdifferentials. *J. Convex Anal.*, 15(2):365–380, 2008.
2. F.J. Aragón Artacho and M.H. Geoffroy. Metric subregularity of the convex subdifferential in Banach spaces. *Preprint*, arXiv:1303.3654, 2013.
3. S. Basu, R. Pollack, and M. Roy. *Algorithms in Real Algebraic Geometry* (Algorithms and Computation in Mathematics). Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.
4. J. Bochnak, M. Coste, and M.-F. Roy. *Real algebraic geometry*, volume 36 of Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]. Springer-Verlag, Berlin, 1998. Translated from the 1987 French original, Revised by the authors.
5. J. Bolte, A. Daniilidis, A.S. Lewis, and M. Shiota. Clarke subgradients of stratifiable functions. *SIAM J. Optim.*, 18(2):556–572, 2007.
6. J.F. Bonnans and A. Shapiro. *Perturbation Analysis of Optimization Problems*. Springer, New York, 2000.
7. J.M. Borwein and Q.J. Zhu. *Techniques of Variational Analysis*. Springer Verlag, New York, 2005.
8. P.H. Clarke, Yu. Ledyaev, R.I. Stern, and P.R. Wolenski. *Nonsmooth Analysis and Control Theory*. Texts in Math. 178, Springer, New York, 1998.
9. M. Coste. *An introduction to o-minimal geometry*. RAAG Notes, 81 pages, Institut de Recherche Mathématiques de Rennes, November 1999.
10. M. Coste. *An introduction to semialgebraic geometry*. RAAG Notes, 78 pages, Institut de Recherche Mathématiques de Rennes, October 2002.
11. T.T.A. Nghia D. Drusvyatskiy, B.S. Mordukhovich. Second-order growth, tilt-stability, and metric regularity of the subdifferential. *Preprint*, arXiv:1304.7385, 2013.
12. A.D. Ioffe. An invitation to tame optimization. *SIAM Journal on Optimization*, 19(4):1894–1917, 2009.
13. A.D. Ioffe. On the theory of subdifferentials. *Adv. Nonlinear Anal.*, 1(1):47–120, 2012.
14. B.S. Mordukhovich. *Variational Analysis and Generalized Differentiation I: Basic Theory*. Grundlehren der mathematischen Wissenschaften, Vol 330, Springer, Berlin, 2005.
15. J. Nocedal and S.J. Wright. *Numerical optimization*. Springer Series in Operations Research and Financial Engineering. Springer, New York, second edition, 2006.
16. J.-P. Penot. *Calculus without derivatives*, volume 260 of Graduate Texts in Mathematics. Springer, New York, 2013.
17. A. Ruszczyński and R.T. Rockafellar. Amenable functions in optimization. In *Nonsmooth optimization: methods and applications* (Erice, 1991), pages 338–353. Gordon and Breach, Montreux, 1992.
18. A. Ruszczyński and R.T. Rockafellar. Prox-regular functions in variational analysis. *Trans. Amer. Math. Soc.*, 348:1805–1838, 1996.
19. R.T. Rockafellar. Lagrange multipliers and optimality. *SIAM Rev.*, 35(2):183–238, 1993.
20. R.T. Rockafellar and R.J-B. Wets. *Variational Analysis*. Grundlehren der mathematischen Wissenschaften, Vol 317, Springer, Berlin, 1998.
21. A. Shiota. *Semialgebraic sets*. Birkhäuser Boston Inc., Cambridge, MA, USA, 1997.
22. L. van den Dries. *Tame topology and o-minimal structures*, volume 248 of *LMS Lecture Note Series*. Cambridge University Press, Cambridge, 1998.
23. L. van den Dries and C. Miller. Geometric categories and o-minimal structures. *Duke Math. J.*, 84:497–540, 1996.
