Radius Theorems for Monotone Mappings

A. L. Dontchev[*], A. Eberhard[†] and R. T. Rockafellar[‡]

Abstract. For a Hilbert space X and a mapping $F : X \rightrightarrows X$ (potentially set-valued) that is maximal monotone locally around a pair $(\bar{x}, \bar{y})$ in its graph, we obtain a radius theorem of the following kind: the infimum of the norm of a linear and bounded single-valued mapping B such that $F + B$ is not locally monotone around $(\bar{x}, \bar{y} + B\bar{x})$ equals the monotonicity modulus of F. Moreover, the infimum is not changed if taken with respect to B symmetric, negative semidefinite and of rank one, and also not changed if taken with respect to all functions $f : X \to X$ that are Lipschitz continuous around $\bar{x}$ and $\|B\|$ is replaced by the Lipschitz modulus of f at $\bar{x}$. As applications, a radius theorem is obtained for the strong second-order sufficient optimality condition of an optimization problem, which in turn yields a radius theorem for quadratic convergence of the Newton method applied to that problem. A radius theorem is also derived for mappings that are merely hypomonotone.

Key Words. monotone mappings, maximal monotone, locally monotone, radius theorem, optimization problem, Newton method.

AMS Subject Classification. 47H05, 49J53, 90C31.

1 Introduction

Throughout, unless stated otherwise, X is a real Hilbert space with inner product $\langle \cdot, \cdot \rangle$ and associated norm $\| \cdot \|$. A mapping F from X to X, allowed to be set-valued in general as indicated by the notation $F : X \rightrightarrows X$, has graph $\operatorname{gph} F = \{ (x, y) \in X \times X \mid y \in F(x) \}$ and domain $\operatorname{dom} F = \{ x \in X \mid F(x) \neq \emptyset \}$. It is *single-valued*, indicated by $F : X \to X$, if $\operatorname{dom} F = X$ and $F(x)$ consists of only one y for each x; then $y \in F(x)$ can be written as $F(x) = y$. (Single-valued mappings and "functions" are synonomous in this sense.) The inverse

[*]Mathematical Reviews, 416 Fourth Street, Ann Arbor, MI 48107-8604. Supported by the NSF Grant 156229, by Austrian Science Foundation (FWF) Grant P26640-N25, and the Australian Research Council, Discovery grant DP160100854.

[†]RMIT University, Melbourne, Australia 3001. Research supported by the Australian Research Council, Discovery grant DP140100985.

[‡]Department of Mathematics, University of Washington, Seattle, WA 98195–4350.

of a mapping $F : X \rightrightarrows X$ is the mapping $F^{-1} : X \rightrightarrows X$ defined by $F^{-1}(y) = \{\, x \in X \mid y \in F(x) \,\}$, which of course may be set-valued even if F is single-valued.

The *monotonicity* for a mapping $F : X \rightrightarrows X$ means that

$$\langle y - y', x - x' \rangle \geq 0 \ \text{ whenever } \ (x, y), (x', y') \in \operatorname{gph} F,$$

while *maximal* monotonicity refers to the impossibility of enlarging $\operatorname{gph} F$ without violating this inequality. On the local level, F is *monotone at* $\bar{x}$ *for* $\bar{y}$ *if* $\bar{y} \in F(\bar{x})$ and the monotonicity inequality holds when $\operatorname{gph} F$ is replaced by $(U \times V) \cap \operatorname{gph} F$ for some neighborhoods U of $\bar{x}$ and V of $\bar{y}$. It is *maximal* monotone at $\bar{x}$ for $\bar{y}$ if, in addition, nothing more in $U \times V$ can be added to $\operatorname{gph} F$ without upsetting that. The *modulus* of monotonicity at $\bar{x}$ for $\bar{y} \in F(\bar{x})$ is

$$\operatorname{mon}(F; \bar{x} \,|\, \bar{y}) = \liminf_{\substack{x, x' \to \bar{x},\, x \neq x' \\ y, y' \to \bar{y} \\ (x,y),(x',y') \in \operatorname{gph} F}} \frac{\langle y - y', x - x' \rangle}{\|x - x'\|^2},$$

which could be ∞ (and is interpreted to be so in particular in the degenerate case where $\operatorname{dom} F = \{\bar{x}\}$). When $\operatorname{mon}(F; \bar{x} \,|\, \bar{y}) > 0$, F is said to be *strongly* monotone at $\bar{x}$ for $\bar{y}$. This corresponds the existence of $\sigma > 0$ for which there are neighborhoods U of $\bar{x}$ and V of $\bar{y}$ such that

$$\text{(1)} \qquad \langle y - y', x - x' \rangle \geq \sigma \|x - x'\|^2 \ \text{ whenever } \ (x, y), (x', y') \in (U \times V) \cap \operatorname{gph} F.$$

Any positive $\sigma < \operatorname{mon}(F; \bar{x} \,|\, \bar{y})$ will fit this description.

Although monotonicity properties are the focus of this paper, Lipschitz continuity will also have a role. The Lipschitz modulus of a single-valued mapping $f : X \to X$ at a point $\bar{x}$ is defined by

$$\operatorname{lip}(f; \bar{x}) = \limsup_{\substack{x, x' \to \bar{x} \\ x \neq x'}} \frac{\|f(x) - f(x')\|}{\|x - x'\|}.$$

Having $\operatorname{lip}(f; \bar{x}) < l$ corresponds to having a neighborhood U of $\bar{x}$ such that f is Lipschitz continuous on U with Lipschitz constant l. Conversely, if f is Lipschitz continuous around $\bar{x}$ with Lipschitz constant l then we have $\operatorname{lip}(f; \bar{x}) \leq l$. If f is not Lipschitz continuous around $\bar{x}$ then $\operatorname{lip}(f; \bar{x}) = \infty$.

As a main result in this paper we obtain the following, where as usual the space of continuous linear mappings $B : X \to X$ is denoted by $L(X, X)$.

Theorem 1 (radius theorem for monotonicity). *For a mapping $F : X \rightrightarrows X$ that is maximal monotone at $\bar{x}$ for $\bar{y}$,*

$$\text{(2)} \qquad \inf_{B \in L(X,X)} \left\{ \|B\| \;\middle|\; F + B \ \textit{not monotone at } \bar{x} \textit{ for } \bar{y} + B\bar{x} \ \right\} = \operatorname{mon}(F; \bar{x} \,|\, \bar{y}).$$

Moreover, the infimum in the left side of (2) is not changed if taken with respect to $B \in L(X, X)$ symmetric, negative semidefinite and of rank one, and also not changed if taken with respect to all functions $f : X \to X$ that are Lipschitz continuous around $\bar{x}$, with $\|B\|$ replaced by $\operatorname{lip}(f; \bar{x})$.

The proof of Theorem 1 will be given in Section 3 after preparations in Section 2 which include a demonstration that local strong monotonocity is stable under perturbations by single-valued mappings with small Lipschitz constants. Some applications will be taken up in Section 4.

The result in Theorem 1 is in line with the general pattern of so-called radius theorems: given a mapping with a certain desirable property, the "radius" of this property is the "minimal perturbation" such that the perturbed mapping may lose this property. In other words, the radius for a property of a mapping tells us the "distance" from that mapping to mappings that do not have the property. An early example in linear algebra is a radius theorem that goes back to a paper by Eckart and Young [5]: for any nonsingular matrix A,

$$\inf\left\{\|B\| \,\middle|\, A+B \ \text{singular}\right\} = \frac{1}{\|A^{-1}\|}. \tag{3}$$

Far reaching generalizations of that result were obtained in [2] and in [3], see also [4, Section 6A], establishing that typically

$$\mathrm{rad} = \frac{1}{\mathrm{reg}}, \tag{4}$$

where rad is the radius for the property in question, in analogy with the left side of (2) in Theorem 1, and reg is the regularity modulus for the property. Specifically, these papers considered three basic properties of set-valued mappings that are important in variational analysis and optimization, namely metric regularity, strong metric regularity, and strong metric subregularity, which extend to set-valued/nonlinear mappings the classical properties of surjectivity, invertibility and injectivity, respectively, of continuous linear mappings in linear normed spaces. Radius theorems are sometimes called condition number theorems in the literature, see e.g. [11]; indeed, the equality (3) can be also written in terms of the (absolute) condition number of the matrix A. We do not use such terminology here because there are variants of "condition number" that are not tied with any concept of radius.

2 Stability under small Lipschitz perturbations

A mapping F acting from a metric space to possibly another metric space is said to have a *single-valued localization at $\bar{x}$ for $\bar{y}$*, where $(\bar{x},\bar{y}) \in \operatorname{gph} F$, if there exist neighborhoods U of $\bar{x}$ and V of $\bar{y}$ such that the truncated inverse $V \ni y \mapsto F^{-1}(y) \cap U$ is a single-valued on V. If F^{-1} has a single-valued localization at $\bar{y}$ for $\bar{x}$ that is Lipschitz continuous around $\bar{y}$, then F is said to be *strongly metrically regular* at $\bar{x}$ for $\bar{y}$. This term was coined by S. M. Robinson in his seminal paper [9] for a formally different property whose definition was subsequently adapted to the form we use here; for a broad coverage of the major developments around this property as well as other related properties of mappings in variational analysis and optimization; see [4]. The following result was proved in [3, Theorem 4.6]; see also [4, Theorem 6A.8]:

Theorem 2 (radius theorem for strong metric regularity). *Consider a mapping $F : I\!\!R^n \rightrightarrows I\!\!R^m$ that is strongly metrically regular at $\bar{x}$ for $\bar{y}$; specifically, F^{-1} has a single-valued localization s at $\bar{y}$ for $\bar{x}$ with $\mathrm{lip}(s;\bar{y}) < \infty$. Then*

$$\inf_{B\in L(R^n,R^m)}\left\{ \|B\| \,\middle|\, F+B \text{ is not strongly metrically regular at } \bar{x} \text{ for } \bar{y}+B\bar{x} \right\} = \frac{1}{\mathrm{lip}(s;\bar{y})}. \tag{5}$$

Moreover, the infimum remains unchanged when taken either with respect to linear mappings of rank one or with respect to all functions f that are Lipschitz continuous around $\bar{x}$, with $\|B\|$ replaced by the Lipschitz modulus $\operatorname{lip}(f;\bar{x})$ of f at $\bar{x}$.

Clearly Theorem 2 covers the Eckart-Young equality (3) but goes far beyond that in revealing key importance of Lipschitz properties in the study of solution stability.

The proof of Theorem 2 presented in [3] uses a radius theorem for metric regularity given in [2], the proof of which is based on a criterion involving a generalized derivative of a set-valued mapping: the so-called coderivative. Another proof of the Theorem 2, given in [4, Theorem 6A.7], uses a criterion involving another kind of generalized derivative, namely the graphical (or contingent) derivative. However, the derivative and coderivative criteria are essentially valid in finite-dimensional spaces only. In both proofs the inequality $\geq$ in (5) is derived first from an extended version of the Lyusternik-Graves theorem, cf. [4, Theorem 5F.1], for a mapping F acting in Banach spaces and a perturbation B represented by a Lipschitz covers the case $\operatorname{lip}(s;\bar{y}) = 0$ under the convention that $1/0 = +\infty$. Then a specific linear mapping B of rank 1 is constructed such that the sum $F + B$ violates either the derivative or the coderivative criterion; this limits the proof to a finite dimensional setting.

It is still an open problem as to whether it may be possible to extend the radius theorem for (strong) metric regularity to mappings acting in infinite-dimensional spaces; see [6] for a related discussion.

Theorem 2 is quite general but cannot be applied to situations where the perturbed mapping $F + B$ ought to have a specific form, i.e., to *structured* perturbations. For instance, it cannot be applied satisfactorly to the mapping that describes the Karush-Kuhn-Tucker (KKT) conditions in nonlinear programming, because the perturbed mapping there ought to have the form corresponding to a KKT condition. It is an open whether one might find a radius theorem in that sense even for the standard nonlinear programming problem. However, we will give a partial answer to this question in the Section 4 of this paper for the problem of minimizing a $\mathcal{C}^2$ function over a fixed convex polyhedral set.

There is a well known link between strong metric regularity with strong maximal monotonicity, which can be established with the help of the Minty parameterization. The following proposition is a modified version of a part of [7, Lemma 3.3].

Proposition 3. *Consider a mapping $F : X \rightrightarrows X$ that is maximal monotone at $\bar{x}$ for $\bar{y}$ with $\operatorname{mon}(F;\bar{x}\,|\,\bar{y}) > 0$. Then the inverse F^{-1} has a Lipschitz continuous single-valued graphical localization s at $\bar{y}$ for $\bar{x}$. Moreover*

$$\max\left\{ \sigma > 0 \,\middle|\, \operatorname{lip}\left(\frac{1}{2\sigma}I - s;\bar{y}\right) \leq \frac{1}{2\sigma} \right\} = \operatorname{mon}(F;\bar{x}\,|\,\bar{y})$$

and so

$$\operatorname{lip}(s;\bar{y}) = \frac{1}{\operatorname{mon}(F;\bar{x}\,|\,\bar{y})}. \tag{6}$$

Proof. Suppose that F is strongly maximal monotone at $\bar{x}$ for $\bar{y}$. Let $0 < \sigma < \operatorname{mon}(F;\bar{x}\,|\,\bar{y})$. From the strong monotonicity there exist neighborhoods U of $\bar{x}$ and V of $\bar{y}$ such that for every $x, x' \in \operatorname{dom} F \cap U$ and $y \in F(x) \cap V$, $y' \in F(x') \cap V$, the inequality (1) holds. The rest of the proof in this direction will follow the localization argument presented in a part of the proof of [7, Lemma 3.3]. First, note that the mapping $\tilde{F} = (F - \sigma I) \cap W$ is maximal monotone at $\bar{x}$ for $\bar{y}$, where I is the identity mapping and W is a neighborhood of $(\bar{x}, \bar{y})$, and also $\operatorname{gph} F^{-1} \cap (U \times V) = \operatorname{gph}(\sigma I + \tilde{F})^{-1} \cap (U \times V)$. According to Minty's theorem, cf. [1, Theorem 21.1], the mapping $(\sigma I + \tilde{F})^{-1}$ is Lipschitz continuous with a constant σ^{-1}, hence $\operatorname{lip}(s; \bar{y}) \le \sigma^{-1}$, or $\operatorname{lip}(s; \bar{y}) \le \frac{1}{\operatorname{mon}(F; \bar{x}\,|\,\bar{y})}$. In particular we have from the strong monotonicity that

$$
\begin{aligned}
(7)\ \|y' - y'' - 2\sigma(s(y') - s(y''))\|^2 &= \|y' - y''\|^2 - 4\sigma[\langle y' - y'', s(y') - s(y'')\rangle - \sigma\|y' - y''\|^2] \\
&\le \|y' - y''\|^2.
\end{aligned}
$$

Thus the mapping $y \mapsto [\frac{1}{2\sigma}I - s](y)$ is Lipschitz continuous with Lipschitz constant $\frac{1}{2\sigma}$, i.e.

$$
\operatorname{lip}(\frac{1}{2\sigma}I - s; \bar{y}) \le \frac{1}{2\sigma}. \tag{8}
$$

This yields the inequality $\max\left\{ \sigma > 0 \,\middle|\, \operatorname{lip}(\frac{1}{2\sigma}I - s; \bar{y}) \le \frac{1}{2\sigma} \right\} \ge \operatorname{mon}(F; \bar{x}\,|\,\bar{y})$. Now suppose this inequality is strict. Take $\sigma > \operatorname{mon}(F; \bar{x}\,|\,\bar{y}) > 0$ with $\operatorname{lip}(\frac{1}{2\sigma}I - s; \bar{y}) \le \frac{1}{2\sigma}$. Then we have within some neighbourhoods U of $\bar{x}$ and V of $\bar{y}$ and y', $y'' \in V$, that $x' := s(y')$, $x'' := s(y'') \in U$ with

$$
\|\frac{1}{2\sigma}(y' - y'') - (s(y') - s(y''))\|^2 \le \frac{1}{2\sigma}\|y' - y''\|^2. \tag{9}
$$

Expanding this yields $\langle x' - x'', y' - y'' \rangle \ge \sigma\|x' - x''\|^2$, which in turn implies $\operatorname{mon}(F; \bar{x}\,|\,\bar{y}) \ge \sigma$, a contradiction.

Now suppose, contrary to assertion, that $\operatorname{mon}(F; \bar{x}\,|\,\bar{y}) < 1/\operatorname{lip}(s\,|\,\bar{y})$. Take $\operatorname{mon}(F; \bar{x}\,|\,\bar{y}) < \sigma < 1/\operatorname{lip}(s\,|\,\bar{y})$, so $\operatorname{lip}(s\,|\,\bar{y}) < 1/\sigma$ and $\operatorname{mon}(F; \bar{x}\,|\,\bar{y}) < \sigma$. The later inequality implies the existence of y', $y'' \in V$, for which

$$
\langle s(y') - s(y''), y' - y'' \rangle < \sigma\|y' - y''\|^2. \tag{10}
$$

Using the identity (7) we have the inequality (10) implying the following inequality

$$
\|(\frac{1}{2\sigma}y' - s(y')) - (\frac{1}{2\sigma}y'' - s(y''))\|^2 \ge (\frac{1}{2\sigma})^2\|y' - y''\|^2,
$$

which shows that the mapping $y \mapsto \frac{1}{2\sigma}y - s(y)$ is not $\frac{1}{2\sigma}$ locally Lipschitz, or in other words that $\operatorname{lip}(\frac{1}{2\sigma}I - s; \bar{y}) > \frac{1}{2\sigma}$, a contradiction. ◻

Note that (6) enables the relationship in (2) of Theorem 1 to be cast equally in the pattern of (4) usually followed by radius theorems, and similaly in other such results below.

Strong metric regularity, along with its "siblings" metric regularity and strong metric regularity, has an important property sometimes called the "inverse function theorem paradigm." Namely, mappings do not lose these regularity properties when perturbed by a function having a "small" Lipschitz constant. A typical such perturbation appears in the case when the mapping is a strictly differentiable function and then the perturbation is its linearization around the reference point minus the function itself.

For metric regularity, this kind of stability comes from the classical Lyusternik-Graves theorem, already mentioned after Theorem 2. For strong metric regularity, the corresponding result is known as Robinson's inverse function theorem; see [4, Theorem 5F.1]. A simplified version of that is as follows. Consider a mapping $F : X \rightrightarrows X$ with $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ and a function $f : X \to X$ that is strictly differentiable at $\bar{x}$. Then the sum $f + F$ is strongly metrically regular at $\bar{x}$ for $f(\bar{x}) + \bar{y}$ if and only of its partial linearization $x \mapsto f(\bar{x}) + Df(\bar{x})(x - \bar{x}) + F(x)$ has that property. A broad coverage of the inverse function theorem paradigm together with numerous applications is furnished in [4].

It turns out that strong monotonicity has a property which is analogous to the inverse function theorem paradigm. Specifically, we have the following theorem:

Theorem 4 (perturbation stability)**.** *Consider a mapping $F : X \rightrightarrows X$ that is monotone at $\bar{x}$ for $\bar{y}$. Then for any $f : X \to X$ we have*

$$\operatorname{mon}(F; \bar{x}\,|\,\bar{y}) + \operatorname{lip}(f; \bar{x}) \geq \operatorname{mon}(f + F; \bar{x}\,|\,\bar{y} + f(\bar{x})) \geq \operatorname{mon}(F; \bar{x}\,|\,\bar{y}) - \operatorname{lip}(f; \bar{x}). \tag{11}$$

Proof. Suppose $\operatorname{mon}(F; \bar{x}\,|\,\bar{y}) = 0$ and $\operatorname{lip}(f; \bar{x}) < l$. Then for any $x_k, x'_k \to \bar{x}$ we have eventually

$$\|f(x_k) - f(x'_k)\| \leq l\|x_k - x'_k\|.$$

There exist sequences $\tau_k \searrow 0$ and $(x_k, y_k), (x'_k, y'_k) \in \operatorname{gph} F$, $(x_k, y_k), (x'_k, y'_k) \to (\bar{x}, \bar{y})$, such that

$$\langle y_k - y'_k, x_k - x'_k \rangle \leq \tau_k \|x_k - x'_k\|^2.$$

But then

$$\langle y_k - y'_k + f(x_k) - f(x'_k), x_k - x'_k \rangle \leq (\tau_k + l)\|x_k - x'_k\|^2.$$

That is, $\operatorname{mon}(f + F; \bar{x}\,|\,\bar{y} + f(\bar{x})) \leq l$. Also, $\operatorname{mon}(f + F; \bar{x}\,|\,\bar{y} + f(\bar{x})) \geq 0 - l$, and the arbitrariness of l gives (11).

Let $\operatorname{mon}(F; \bar{x}\,|\,\bar{y}) > 0$ and choose $0 < \tau < \operatorname{mon}(F; \bar{x}\,|\,\bar{y})$ and $\kappa > \operatorname{lip}(f; \bar{x})$. Then there exist neighborhoods U of $\bar{x}$ and V of $\bar{y}$ such that

$$\langle y - y', x - x' \rangle \geq \tau \|x - x'\|^2 \quad \text{for all } (x, y), (x', y') \in \operatorname{gph} F \cap (U \times V)$$

and

$$\langle f(x) - f(x'), x - x' \rangle \geq -\kappa \|x - x'\|^2 \quad \text{for all } x, x' \in U.$$

Then

$$\langle y - y' + (f(x) - f(x'), x - x' \rangle \geq (\tau - \kappa)\|x - x'\|^2 \quad \text{for all } (x, y), (x', y') \in \operatorname{gph} F \cap (U \times V),$$

which confirms that the middle quantity of (11) is not greater than the right side. Now apply this proven inequality to the mapping $f + F$ and for the Lipschitz function $-f$, noting that $\operatorname{lip}(-f; \bar{x}) = \operatorname{lip}(f; \bar{x})$, to get

$$\operatorname{mon}(F; \bar{x}\,|\,\bar{y}) = \operatorname{mon}(-f + f + F; \bar{x}\,|\,\bar{y} + f(\bar{x}) - f(\bar{x})) \geq \operatorname{mon}(f + F; \bar{x}\,|\,\bar{y} + f(\bar{x})) - \operatorname{lip}(-f; \bar{x})$$

and so

$$\operatorname{mon}(F; \bar{x}\,|\,\bar{y}) + \operatorname{lip}(f; \bar{x}) \geq \operatorname{mon}(f + F; \bar{x}\,|\,\bar{y} + f(\bar{x})).$$

$\square$

Corollary 5 (stability under linearization). *Consider a mapping of the form $f+F : X \rightrightarrows X$ with $(\bar{x},\bar{y}) \in \operatorname{gph}(f+F)$, where $f : X \to X$ is strictly differentiable at $\bar{x}$ and F is general set-valued mapping. Then*

$$\operatorname{mon}(f+F;\bar{x}\,|\,\bar{y}) = \operatorname{mon}(f(\bar{x}) + Df(\bar{x})(\cdot - \bar{x}) + F(\cdot);\bar{x}\,|\,\bar{y}).$$

Proof. Let

$$g(x) := -[f(x) - f(\bar{x}) - Df(\bar{x})(x-\bar{x})]$$

and note that

$$\operatorname{lip}(g;\bar{x}) = \limsup_{\substack{x,x'\to\bar{x}\\ x\neq x'}} \frac{g(x)-g(x')}{\|x-x'\|} = 0.$$

Now apply Theorem 4 using the mapping $f+F$ and the Lipschitz mapping g to get

$$\begin{aligned}\operatorname{mon}(f+F;\bar{x}\,|\,\bar{y}) &= \operatorname{mon}(g+f+F;\bar{x}\,|\,\bar{y}+g(\bar{x}))\\ &= \operatorname{mon}(f(\bar{x})+Df(\bar{x})(\cdot-\bar{x})+F(\cdot);\bar{x}\,|\,\bar{y}).\end{aligned}$$

□

3 Proof of Theorem 1

A first step toward the proof of Theorem 1 is the following weaker result:

Theorem 6 (radius theorem for strong monotonicity). *For a mapping $F : X \rightrightarrows X$ that is maximal monotone at $\bar{x}$ for $\bar{y} \in F(\bar{x})$,*

$$\inf_{B\in L(X,X)} \left\{ \|B\| \,\Big|\, F+B \text{ not strongly monotone at } \bar{x} \text{ for } \bar{y}+B\bar{x} \right\} = \operatorname{mon}(F;\bar{x}\,|\,\bar{y})\,. \tag{12}$$

Moreover, the infimum in the left side of (12) is not changed if taken with respect to $B \in L(X,X)$ symmetric, negative semidefinite, and of rank one, and also not changed if taken with respect to all functions $f : X \to X$ that are Lipschitz continuous around $\bar{x}$ and then $\|B\|$ is replaced by $\operatorname{lip}(f;\bar{x})$.

Proof. Let F be maximal monotone at $\bar{x}$ for $\bar{y}$ but not strongly monotone there; $\operatorname{mon}(F;\bar{x}\,|\,\bar{y}) = 0$. Then (12) holds with B the zero mapping, and we are done. Next let F be strongly monotone at $\bar{x}$ for $\bar{y}$ and let $\varepsilon > 0$. From the definition of the monotonicity modulus, there exist neighborhoods U of $\bar{x}$ and V of $\bar{y}$ such that

$$\langle y-y', x-x'\rangle \geq (\operatorname{mon}(F;\bar{x}|\bar{y}) - \varepsilon)\|x-x'\|^2 \quad \text{for all } (x,y),(x',y') \in \operatorname{gph} F \cap (U\times V). \tag{13}$$

From Proposition 3 the inverse mapping F^{-1} has a Lipschitz localization s around $\bar{y}$ for $\bar{x}$. Choose any $\gamma > 0$ and let $\alpha > 0$ and $\beta > 0$ be such that $I\!B_\alpha(\bar{x}) \subset U$, $I\!B_\beta(\bar{y}) \subset V$, the mapping $I\!B_\beta(\bar{y}) \ni y \mapsto s(y) := F^{-1}(y) \cap I\!B_\alpha(\bar{x})$ is single-valued and Lipschitz continuous, and

$$\beta + \frac{\alpha}{(\operatorname{mon}(F;|\bar{y}) - \varepsilon)(\operatorname{lip} s(\cdot) - \varepsilon)^2} \leq \gamma. \tag{14}$$

From the definition of the Lipschitz modulus, there exist $(x,y),(x',y') \in (I\!B_\alpha(\bar{x})\times I\!B_\beta(\bar{y}))\cap \operatorname{gph} F$, $x\neq x'$, $y'\neq y$, such that

$$\frac{\|y-y'\|^2}{\|x-x'\|} \geq \operatorname{lip}(s;\bar{y}) - \varepsilon,$$

which is the same as

$$\frac{\|y-y'\|^2}{\|s(y)-s(y')\|} \leq \frac{1}{\operatorname{lip}(s;\bar{y})-\varepsilon}. \tag{15}$$

With the so chosen y, y', define the continuous linear mapping $B : X \to X$ by

$$\xi \mapsto B(\xi) = -\frac{\langle y-y',\xi\rangle(y-y')}{\langle y-y', s(y)-s(y')\rangle}\,. \tag{16}$$

This mapping B is symmetric, negative semidefinite, and of rank one. Furthermore, from (13) and (15) we have

$$\begin{aligned}
\|B\| &= \frac{\|y-y'\|^2}{\langle y-y', s(y)-s(y')\rangle} \\
&\leq \frac{\|y-y'\|^2}{(\operatorname{mon}(F;\bar{x}\,|\,\bar{y})-\varepsilon)\|s(y)-s(y')\|^2} \leq \frac{1}{(\operatorname{mon}(F;\bar{x}\,|\,\bar{y})-\varepsilon)(\operatorname{lip} s(\cdot)-\varepsilon)^2}\,.
\end{aligned} \tag{17}$$

We also have from (14) that

$$\|y+B(x)-\bar{y}-B(\bar{x})\| \leq \|y-\bar{y}\| + \|B\|\|s(y)-\bar{x}\| \leq \gamma,$$

and the same for $y'-Bs(y')$.

Since $B(s(y)-s(y')) = y'-y$ we obtain

$$\begin{aligned}
F(s(y))+B(s(y))-y'-B(s(y')) &= F(s(y))+B(s(y)-s(y'))-y' \\
&= F(s(y))+y'-y-y' = F(s(y))-y \ni 0.
\end{aligned}$$

Hence, $F(s(y))+B(s(y)) \ni y'+B(s(y'))$, that is,

$$s(y) \in (F+B)^{-1}(y'+B(s(y'))).$$

But since $F(s(y')) \ni y'$, we also have $F(s(y'))+B(s(y')) \ni y'+B(s(y'))$, that is,

$$s(y') \in (F+B)^{-1}(y'+B(s(y'))).$$

We see that any localization of the mapping $(F+B)^{-1}$ around $(\bar{x},\bar{y}+B\bar{x})$ cannot be single-valued on $I\!B_\gamma(\bar{y}+B\bar{x})$. Because γ could be arbitrarily small, this means that $F+B$ is not strongly monotone.

Let

$$\rho_1(F) = \inf_{\substack{B\in L(X,X) \text{ symmetric,} \\ \text{negative semidefinite, rank 1}}} \left\{\, \|B\| \;\middle|\; F+B \text{ is not strongly monotone at } \bar{x} \text{ for } \bar{y}+B\bar{x} \right\}.$$

Since the mapping B defined in (16) is symmetric, negative semidefinite, and of rank 1, we get from the bound (17) that

$$\rho_1(F) \le \frac{1}{(\mathrm{mon}(F;\bar{x}|\bar{y})-\varepsilon)(\mathrm{lip}(s;\bar{y})-\varepsilon)^2}\,.$$

Since $\rho_1(F)$ doesn't depend on ε, passing with ε to zero and taking into account (6), we obtain

$$\rho_1(F) \le \mathrm{mon}(F;\bar{x}\,|\,\bar{y})\,. \tag{18}$$

Denote by $\rho_2(F)$ the left side of (12), where the infinum is taken with respect to all functions $g : X \to X$ that are Lipschitz continuous around $\bar{x}$. Take any such g with $\mathrm{lip}(g;\bar{x}) < \mathrm{mon}(F;\bar{x}\,|\,\bar{y})$ and choose μ and ν such that $\mathrm{mon}(F;\bar{x}\,|\,\bar{y}) > \mu > \nu > \mathrm{lip}(g;\bar{x})$. Let $(x',y'),(x'',y'') \in \mathrm{gph}\,F$ be sufficiently close to $(\bar{x},\bar{y})$, so that strong monotonicity of F at $\bar{x}$ for $\bar{y}$ and Lipschitz continuity of g around $\bar{x}$ apply. Note that

$$\langle g(x')-g(x''), x''-x'\rangle \ge -\nu\|x''-x'\|^2\,. \tag{19}$$

Adding and subtracting the expressions in (19) in (1) we get

$$\langle (y'+g(x'))-(y''+g(x'')), x'-x''\rangle \ge (\mu-\nu)\|x'-x''\|^2.$$

Thus $F+g$ is strongly monotone and therefore

$$\rho_2(F) \ge \mu-\nu\,.$$

Since $\rho_2(F)$ does not depend on the choice of μ and ν, by passing to the limits $\alpha \to \mathrm{mon}(F;\bar{x}\,|\,\bar{y})$ and $\beta \to \mathrm{lip}(g;\bar{x})$ we obtain

$$\rho_2(F) \ge \mathrm{mon}(F;\bar{x}\,|\,\bar{y}) - \mathrm{lip}(g;\bar{x})\,. \tag{20}$$

But $\rho_2(F)$ doesn't depend on the particular g either, hence we can take the infimum of $\mathrm{lip}(g;\bar{x})$ in (20) which is zero. Therefore,

$$\rho_2(F) \ge \mathrm{mon}(F;\bar{x}\,|\,\bar{y})\,. \tag{21}$$

Taking into account that $\rho_2(F)$ is the infimum in the left side of (12) taken over a set that is larger than the one over which the infimum is taken to define $\rho_1(F)$, we obtain

$$\rho_1(F) \ge \rho_2(F)\,.$$

This inequality, combined with (18) and (21), completes the proof of (12). ◻

Remark. When $X = I\!R^n$, the first part of the proof of Theorem 6 can be easily obtained from Theorem 2. Indeed, from Proposition 3, the left side of (12) is not greater than the left side of (5). But the letter is equal to the reciprocal of $\mathrm{lip}(s;\bar{x})$, which equals $\mathrm{mon}(F;\bar{x}\,|\,\bar{y})$.

Completion of the proof of Theorem 1. First let $\mathrm{mon}(F;\bar{x}\,|\,\bar{y}) = 0$. Then for any $\varepsilon > 0$ and any neighborhood $U \times V$ of $(\bar{x}, \bar{y})$ there exist $(x_1, y_1), (x_2, y_2) \in (U \times V) \cap \mathrm{gph}\, F$, $x_1 \neq x_2$ such that

$$0 \leq \frac{\langle x_1 - x_2, y_1 - y_2 \rangle}{\|x_1 - x_2\|^2} < \varepsilon. \tag{22}$$

The mapping

$$x \mapsto H(x) = -\varepsilon x$$

is linear, symmetric and negative definite, and since $(x_1, y_1 - \varepsilon x_1), (x_2, y_2 - \varepsilon x_2) \in \mathrm{gph}(F + H)$ can be arbitrarily close to $(\bar{x}, \bar{y})$, taking into account (22), we have

$$\langle x_1 - x_2, (y_1 - \varepsilon x_1) - (y_2 - \varepsilon x_2) \rangle < \varepsilon \|x_1 - x_2\|^2 - \varepsilon \langle x_1 - x_2, x_1 - x_2 \rangle = 0.$$

Hence $F + H$ is not monotone at $\bar{x}$ for $\bar{y} + H\bar{x}$, and since $\|H\|$ can be arbitrarily small, we conclude that the left side of (2) is zero; this proves the theorem for $\mathrm{mon}(F;\bar{x}\,|\,\bar{y}) = 0$.

Now suppose that $\mathrm{mon}(F;\bar{x}\,|\,\bar{y}) > 0$. From the proof of Theorem 6 we know that for every $\delta > 0$ the symmetric and negative semidefinite mapping B defined in (16) is such that $F + B$ is not strongly monotone, and $\|B\| \leq \mathrm{mon}(F;\bar{x}\,|\,\bar{y}) + \delta$. If $F + B$ is not monotone, this implies that the infimum in the left side of (2) is less than $\mathrm{mon}(F;\bar{x}\,|\,\bar{y})$. On the other hand, if $F + B$ is monotone at $\bar{x}$ for $\bar{y} + B\bar{x}$, then $\mathrm{mon}(F + B;\bar{x}\,|\,\bar{y}) = 0$ and, by repeating the argument in the first part of the proof, for any $\varepsilon > 0$ we find a symmetric and negative definite linear mapping H such that $\|H\| < \varepsilon$ and $F + B + H$ is not monotone around $\bar{x}$ for $\bar{y} + (B + H)\bar{x}$. But the sum $B + H$ is a symmetric and negative definite linear mapping with $\|B + H\| \leq \mathrm{mon}(F;\bar{x}\,|\,\bar{y}) + \delta + \varepsilon$. Since δ and ε can be arbitrarily small, we conclude that the infimum in the left side of (2) is less or equal to $\mathrm{mon}(F;\bar{x}\,|\,\bar{y})$.

To end the proof, observe that the set of nonmonotone mappings is contained in the set of the mappings that are not strongly monotone. Hence from Theorem 6 the infimum on the left side of (2) is not less than $\mathrm{mon}(F;\bar{x}\,|\,\bar{y})$, and we are done. □

4 Applications

An application to optimization will be developed first. Consider the problem

$$\text{minimize} \;\; g(x) \;\; \text{over} \;\; x \in C, \tag{23}$$

where C is a nonempty *polyhedral* convex subset of $I\!R^n$ and $g : I\!R^n \to I\!R$ is twice continuously differentiable everywhere (for simplicity). As is well known, the first-order necessary optimality condition for this problem is

$$\nabla g(x) + N_C(x) \ni 0. \tag{24}$$

A solution of the variational inequality (24) is said to be a *critical point* in the problem. The *critical cone* at $\bar{x}$ for $-\nabla g(\bar{x})$ is defined as

$$K := K_C(\bar{x}, -\nabla g(\bar{x})) = T_C(\bar{x}) \cap [-\nabla g(\bar{x})]^{\perp}$$

and the *critical subspace* is

$$L = K - K.$$

Let $A = \nabla^2 g(\bar{x})$. The strong second-order sufficient condition (SOSC) for problem (23) has the form

$$\langle u, Au \rangle > 0 \quad \text{for all nonzero } u \in L. \tag{25}$$

Theorem 7 (strong second-order sufficient condition). *Let $\bar{x}$ be a critical point for (23). Then the following are equivalent:*

(a) *the strong second-order sufficient condition* (25) *holds at $\bar{x}$;*

(b) *the point $\bar{x}$ is a strict local minimizer in (23) and the mapping $\nabla g + N_C$ is locally strongly monotone at $\bar{x}$ for* 0*, in which case*

$$\operatorname{mon}(\nabla g + N_C; \bar{x}\,|\,0) = \min_{\substack{u \in L \\ \|u\|=1}} \langle u, Au \rangle. \tag{26}$$

Proof. From Corollary 5 we obtain that

$$\operatorname{mon}(\nabla g + N_C; \bar{x}\,|\,0) = \operatorname{mon}(\nabla g(\bar{x}) + A(\cdot - \bar{x}) + N_C; \bar{x}\,|\,0).$$

Furthermore, from the Reduction Lemma [4, 2E.4] we have that

$$O \cap [\operatorname{gph} N_C - (\bar{x}, -\nabla g(\bar{x})] = O \cap \operatorname{gph} N_K$$

for some neighborhood O of $(0, 0)$. Thus,

$$\operatorname{mon}(\nabla g + N_C; \bar{x}\,|\,0) = \operatorname{mon}(A + N_K; 0\,|\,0).$$

By definition $\operatorname{mon}(A + N_K; 0|0)$ is the infimum of σ such that

$$\langle y - y', x - x' \rangle \geq \sigma \|x - x'\|^2$$

for $x, x' \in K$ with $x \neq x'$, and $y \in Ax + N_K(x)$, $y' \in Ax' + N_K(x')$, where we drop the neighborhoods because the mapping $A + N_K$ is positively homogeneous. Thus for $\sigma < \operatorname{mon}(A + N_K; 0|0)$ we have

$$\langle A(x - x') + z - z', x - x' \rangle \geq \sigma \|x - x'\|^2$$

for all $x, x' \in K$, and $z \in N_K(x)$, $z' \in N_K(x')$. In particular we may choose $z = 0 \in N_K(x)$ and $z' = 0 \in N_K(x')$ to deduce that:

$$\langle Au, u \rangle \geq \sigma \|u\|^2 \quad \text{for all } u \in L.$$

It follows that

$$\min_{\substack{u \in L \\ \|u\|=1}} \langle u, Au \rangle \geq \operatorname{mon}(A + N_K; 0|0).$$

Now suppose this inequality is strict and take $\sigma > 0$ with

$$\min_{\substack{u \in L \\ \|u\|=1}} \langle u, Au \rangle > \sigma > \operatorname{mon}(A + N_K; 0|0). \tag{27}$$

Then for any neighbourhoods U and V of 0 there exists $x, x' \in K \cap U,\ y, y' \in V$ with $y \in Ax + N_K(x)$, $y' \in Ax' + N_K(x')$, such that

$$\langle y - y', x - x' \rangle < \sigma \|x - x'\|^2. \tag{28}$$

Let $z \in N_K(x)$ and $z' \in N_K(x')$ be such that $y = Ax + z$ and $y' = Ax' + z'$. Then we have from the monotonicity of $x \mapsto N_K(x)$ that $\langle z - z', x - x' \rangle \geq 0$. It is well know that

$$z \in N_K(x) \quad \Leftrightarrow \quad x \in K,\ z \in K^*,\ \langle x, z \rangle = 0,$$

where K^* is the polar cone to K. Hence $x - x' \in K - K = L$. From (27) and (28) we have

$$\sigma \leq \sigma + \langle \frac{z - z'}{\|x - x'\|}, \frac{x - x'}{\|x - x'\|} \rangle < \langle A \frac{x - x'}{\|x - x'\|}, \frac{x - x'}{\|x - x'\|} \rangle + \langle \frac{z - z'}{\|x - x'\|}, \frac{x - x'}{\|x - x'\|} \rangle < \sigma,$$

a clear contradiction. This gives us (26). □

Note that the expression on the right side of (26) is equal to the minimal eigenvalue of the matrix $V^T AV$ where the rows of V form a basis for the critical subspace L(.

We now obtain a radius theorem for problem (23). Along with (23) we consider the perturbed problem

$$\min\, [g(x) + h(x)] \ \text{ over } \ x \in C, \tag{29}$$

where $h : I\!R^n \to I\!R$ is a $\mathcal{C}^2$ function standing for a perturbation.

Theorem 8 (radius theorem for stronq second-order sufficiency)**.** *Let $\bar{x}$ be a critical point for* (23) *at which the strong second-order sufficient condition holds, and consider the perturbed problem* (29) *with h being a $\mathcal{C}^2$ function. Then*

$$\inf_{h \in \mathcal{C}^2} \Big\{ ||\nabla^2 h(\bar{x})|| \, \Big| \ \bar{x} \text{ is a critical point for (29) at which SOSC fails } \Big\} = \min_{\substack{u \in L \\ \|u\|=1}} \langle u, Au \rangle. \tag{30}$$

Proof. From combining Corollary 5 and Theorem 6, the quantity on the left side of (30) is the same as the quantity

$$\inf \Big\{ \|\nabla^2 h(\bar{x})\| \,\Big|\ \begin{array}{l} \bar{x} \text{ is a local minimizer of (29) and the mapping } \nabla g + \nabla h + N_C \\ \text{is not strongly monotone at } \bar{x} \text{ for } 0 \end{array} \Big\}. \tag{31}$$

It is well known that the mapping $\nabla g + N_C$ is maximal monotone. From Theorem 6 the quantity (31) is equal to $\operatorname{mon}(\nabla g + N_C; \bar{x}|0)$. It remains to invoke the equality (26) proven in Theorem 7. □

As another application, consider the Newton method as applied to (24), this being known also as the sequential quadratic programming method:

$$\nabla g(x_k) + \nabla^2 g(x_k)(x_{k+1} - x_k) + N_C(x_{k+1}) \ni 0, \tag{32}$$

where we assume that the function g is $\mathcal{C}^2$ around a solution $\bar{x}$ of (23) and its second derivative is Lipschitz continuous there. As is well known, see e.g. [4, Theorem 6D.2], the strong second-order sufficient condition (25) is a sufficient condition for local quadratic convergence of the method; namely, if SOSC holds then there exists a neighborhood O of $\bar{x}$ such that for any starting point $x_0 \in O$ there is a unique in O sequence $\{x_k\}$ generated by (32) which is quadratically convergent to $\bar{x}$. From Theorem 8 we obtain:

Theorem 9 (radius theorem for Newton's method)**.** *Let $\bar{x}$ be a critical point for* (23) *at which the strong second-order condition holds, and consider the perturbed problem* (29) *with h being a $\mathcal{C}^2$ function. Then*

$$\inf_{h\in\mathcal{C}^2}\Big\{\ \|\nabla^2 h(\bar{x})\|\ \Big|\ \begin{array}{l}\bar{x} \text{ is a critical point for (29) and in every neighborhood of } \bar{x}\\ \text{there exists a point } x_0 \text{ such that the iteration (32)}\\ \text{applied to (29) and starting from } x_0\\ \text{is not quadratically convergent to } \bar{x}\end{array}\Big\} \geq \min_{\substack{u\in L\\ \|u\|=1}}\langle u, Au\rangle.$$

Observe that in Theorem 8 it is assumed that the reference solution $\bar{x}$ is also a solution of the perturbed problem. This requirement could be possibly relaxed by employing results concerning stability/sensitivity of the solution mapping. We leave this question for future research.

Lastly, we apply Theorem 1 to obtain a radius theorem for hypomonotonicity. Recall that a mapping $F : X \rightrightarrows X$ is said to be *hypomonotone* at $\bar{x}$ for $\bar{y}$ if $\bar{y} \in F(\bar{x})$ if (1) holds for some $\sigma < 0$ and neighborhoods U of $\bar{x}$ and V of $\bar{y}$, which corresponds to $F + \rho I$ being monotone at $\bar{x}$ for $\bar{y} + \rho\bar{x}$ with $\rho = -\sigma > 0$. This can also be articulated in terms of negative values of the monotonicity modulus $\mathrm{mon}(F; \bar{x}\,|\,\bar{y})$. (Trivially, monotonicity entails hypomonotonicity.)

Examples and applications of hypomonotone mappings can be found in [10], p. 550 and p. 615. Hypomonotonicity is important as well in extensions of the proximal point algorithm beyond monotonicity; see [8]. Notably, every mapping which has a Lipschitz localization at $\bar{x}$ for $\bar{y}$ is hypomonotone at $\bar{x}$ for $\bar{y}$.

From Theorem 1 and the definition of hypomonoticity we obtain:

Theorem 10 (radius theorem for hypomonotonicity)**.** *For any mapping $F : X \rightrightarrows X$ that is locally maximal monotone at $\bar{x}$ for $\bar{y}$,*

$$\text{(33)}\qquad \inf_{B\in L(X,X)}\Big\{\ \|B\|\ \Big|\ F + B \ \text{ not hypomonotone at } \bar{x} \text{ for } \bar{y} + B\bar{x}\ \Big\} = \inf_{\rho>0} \mathrm{mon}(F + \rho I; \bar{x}\,|\,\bar{y}).$$

Moreover, the infimum in the left side of (33) is not changed if taken with respect to $B \in L(X,X)$ symmetric, negative semidefinite and of rank one, and also not changed if taken with respect to all functions $f : X \to X$ that are Lipschitz continuous around $\bar{x}$, with $\|B\|$ replaced by $\mathrm{lip}(f; \bar{x})$.

Proof. Let $\alpha >$ the left side of (33); then there exists $B \in L(X,X)$ with $\|B\| \leq \alpha$ such that $F + B$ is not hypomonotone at $\bar{x}$ for $\bar{y} + B\bar{x}$. This means that for every $\rho > 0$ the mapping $F + B + \rho I$ is not monotone at $\bar{x}$ for $\bar{y} + B\bar{x}$. But then, from Theorem 1,

$$\alpha \geq \|B\| > \mathrm{mon}(F + \rho I; \bar{x}\,|\,\bar{y}) \geq \inf_{\rho>0} \mathrm{mon}(F + \rho I; \bar{x}\,|\,\bar{y}).$$

Hence, the left side of (33) is not less than the right side.

Conversely, let $\alpha >$ the right side of (33); then for some $\rho > 0$ we obtain that $\alpha > \mathrm{mon}(F + \rho I; \bar{x}\,|\,\bar{y})$. Then from Theorem 1 we have the existence of $B \in L(X,X)$ with $\|B\| < \alpha$ such that $F + B + \rho I$ is not monotone at $\bar{x}$ for $\bar{y} + B\bar{x}$, hence $F + B$ is not hypomonotone at $\bar{x}$ for $\bar{y} + B\bar{x}$. Then the left side of (33) is not greater than α, and therefore not greater than the right side. This proves the equality (33). ◻

References

[1] H. H. Bauschke, P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*, Springer, New York, 2011.

[2] A. L. Dontchev, A. S. Lewis, R. T. Rockafellar, The radius of metric regularity, *Trans. Amer. Math. Soc.* **355** (2003), pp. 493–517.

[3] A. L. Dontchev, R. T. Rockafellar, Regularity and conditioning of solution mappings in variational analysis, *Set-Valued Anal.* **12** (2004), pp. 79–109.

[4] A. L. Dontchev, R. T. Rockafellar, *Implicit Functions and Solution mappings. A View From Variational Analysis*. Springer Monographs in Mathematics, Second Edition, Springer, Dordrecht, 2014.

[5] C. Eckart, G. Young, The approximation of one matrix by another of lower rank, *Psychometrica* **1** (1936), pp. 211–218.

[6] A. D. Ioffe, On stability estimates for the regularity property of maps, in *Topological Methods, Variational Methods and Their Applications*, World Sci. Publ., River Edge, NJ, 2003, pp. 133–142.

[7] B. S. Mordukhovich, T. T. A. Nghia, Local monotonicity and full stability for parametric variational systems, *SIAM J. Optim.* **26** (2016), pp. 1032–1059.

[8] Pennanen, T., Local convergence of the proximal point algorithm and multiplier methods without monotonicity, *Mathematics of Operations Research* **27** (2002), pp. 170–191.

[9] S. M. Robinson, Strongly regular generalized equations, *Mathematics of Operations Research* **5** (1980), pp. 43–62.

[10] R. T. Rockafellar, R. J-B Wets, *Variational analysis*, Springer, Berlin, 1998.

[11] T. Zolezzi, A condition number theorem in convex programming, *Mathematical Programming* **149** (2015), pp. 195–207.