Semismoothness of Solutions to Generalized Equations and the Moreau-Yosida Regularization

Fanwen Meng[1], Defeng Sun[2], and Gongyun Zhao[3]

June 11, 2004

Abstract

We show that a locally Lipschitz homeomorphism function is semismooth at a given point if and only if its inverse function is semismooth at its image point. We present a sufficient condition for the semismoothness of solutions to generalized equations over cone reducible (nonpolyhedral) convex sets. We prove that the semismoothness of solutions to the Moreau-Yosida regularization of a lower semicontinuous proper convex function is implied by the semismoothness of the metric projector over the epigraph of the convex function.

Keywords: Semismooth, Generalized Equations, Moreau-Yosida Regularization

1 Introduction

In this paper, we are interested in the semismoothness of solutions to parameterized generalized equations over (nonpolyheral) sets and the semismoothness of solutions to the Moreau-Yosida regularization of a lower semicontinuous proper convex function. Let X, Y, and U be finite dimensional vector spaces each equipped with a scalar product denoted by $\langle\cdot,\cdot\rangle$. Let $\|\cdot\|$ be the norm induced by $\langle\cdot,\cdot\rangle$. We use Z to represent an arbitrary vector space from vector spaces X, Y, and U. Suppose that $F : X \times U \to X$ is a single valued continuously differentiable mapping and $G : X \times U \to Y$ is a single valued twice continuously differentiable mapping. We denote by $J_xG(x,u) : X \times U \to Y$ the derivative mapping of G with respect to $x \in X$. Let $J_xG(x,u)^* : Y \times U \to X$ be the adjoint of the derivative mapping $J_xG(x,u)$ and let $K \subseteq Y$ be a closed convex set. The parameterized generalized equation, considered in Shapiro [36], is to find $x \in X$ such that

$$F(x,u) + J_xG(x,u)^*\lambda = 0\ , \quad \lambda \in N_K(G(x,u))\,, \tag{1.1}$$

where for any closed convex set $D \subseteq Z$, $N_D(z)$ denotes the normal cone of D at z:

$$N_D(z) = \begin{cases} \{d \in Z \mid \langle d, c - z\rangle \le 0 \ \ \forall\, c \in D\} & \text{if } z \in D\,, \\ \emptyset & \text{if } z \notin D\,. \end{cases}$$

[1]Department of Finance and Accounting, School of Business, National University of Singapore, Republic of Singapore. E-mail: fnbv4@nus.edu.sg.

[2]Department of Mathematics, National University of Singapore, Singapore 117543, Republic of Singapore. E-mail: matsundf@nus.edu.sg. Fax: 0065 6779 5452.

[3]Department of Mathematics, National University of Singapore, Singapore 117543, Republic of Singapore. E-mail: matzgy@nus.edu.sg.

It is noted that if the mapping $F(x,u)$ is the derivative of a real valued function $f : X \times U \to \Re$ with respect to x, i.e., $F(x,u) = J_x f(x,u)$, then under some standard constraint qualifications (see, e.g., [31]), (1.1) turns to be the first order necessary optimality conditions of the following parameterized optimization problem

$$\begin{array}{ll} \min & f(x,u) \\ \text{s.t.} & G(x,u) \in K\,, \\ & x \in X\,. \end{array} \tag{1.2}$$

For any closed convex set D of the vector space Z and $z \in Z$, let $\Pi_D(z)$ denote the metric projection of z onto D, i.e.,

$$\Pi_D(z) := \operatorname{argmin}\{\frac{1}{2}\|d - z\|^2 \mid d \in D\}\,.$$

It is well known [42] that the metric projector $\Pi_D(\cdot)$ is contractive, i.e., for any two vectors $z^1, z^2 \in Z$,

$$\|\Pi_D(z^1) - \Pi_D(z^2)\| \leq \|z^1 - z^2\|\,.$$

Then, according to Eaves [5], problem (1.1) is equivalent to the following parameterized nonsmooth equation

$$H(x,\lambda,u) := \left[\begin{array}{c} F(x,u) + J_x G(x,u)^* \lambda \\ G(x,u) - \Pi_K[G(x,u) + \lambda] \end{array} \right] = 0\,. \tag{1.3}$$

For $u \in U$, let $(x(u), \lambda(u))$ (if exists) solve (1.3), i.e., $H(x(u), \lambda(u), u) = 0$. Shapiro [36] studied some perturbed properties including the Lipschitz continuity of $(x(\cdot), \lambda(\cdot))$ at a given point $u_0 \in U$. In this paper, we shall further study the semismoothness of $(x(\cdot), \lambda(\cdot))$ at u_0.

A related yet quite different problem is the Moreau-Yosida regularization of a lower semicontinuous proper convex function $f : X \to R \cup \{+\infty\}$. Let $\varepsilon > 0$ be a positive number. The Moreau-Yosida regularization of f [24, 41] is defined by

$$\begin{array}{rll} \hat{f}_\varepsilon(u) := & \min & \{f(x) + \frac{\varepsilon}{2}\langle u - x, u - x\rangle\}\,, \\ & \text{s.t.} & x \in X\,. \end{array} \tag{1.4}$$

It is well known that $\hat{f}_\varepsilon$ is continuously differentiable on X and for any $u \in X$,

$$\nabla \hat{f}_\varepsilon(u) = \varepsilon(u - x(u)),$$

where $x(u)$ denotes the unique optimal solution of (1.4). It is also known that $x(\cdot)$ is globally Lipschitz continuous, which implies that $\nabla \hat{f}_\varepsilon$ is globally Lipschitz continuous [33, p.546]. Here, we are interested in the semismoothness of $x(\cdot)$ at a given point $u_0 \in X$, which is a key condition for the superlinear convergence of an approximate Newton's method designed in Fukushima and Qi [9] for solving nonsmooth convex optimization problems.

The organization of this paper is as follows. In section 2, we discuss the semismoothness of locally Lipschitz homeomorphism functions. In particular, based on Kummer's inverse function theorem for locally Lipschitz functions [15], we show that a locally Lipschitz homeomorphism function is G-semismooth at a given point if and only if its inverse function is G-semismooth at its image point. In Section 3, we study the semismoothness of solutions to parameterized generalized equations. A sufficient condition is presented for the semismoothness of solutions to parameterized equations over cone reducible (nonpolyhedral) convex sets. In section 4, we reduce the semismoothness of solutions to the Moreau-Yosida regularization of a convex function to the semismoothness of the metric projector over the epigraph of the convex function. We make final conclusions in Section 5.

2 Semismoothness of Locally Lipschitz Homeomorphism Functions

Let X and Y be finite dimensional vector spaces. Let $\mathcal{O}$ be an open set in X and $\Phi : \mathcal{O} \subseteq X \to Y$ be a locally Lipschitz continuous function on the open set $\mathcal{O}$. By Rademacher's theorem, Φ is almost everywhere F(réchet)-differentiable in $\mathcal{O}$. We denote by $\mathcal{D}_\Phi$ the set of points in $\mathcal{O}$ where Φ is F-differentiable. Let $J\Phi(x)$, which is a linear mapping from X to Y, denote the derivative of Φ at $x \in \mathcal{O}$ if Φ is F(réchet)-differentiable at x, and $J\Phi(x)^* : Y \to X$ the adjoint of $J\Phi(x)$. Then, the B-subdifferential of Φ at $x \in \mathcal{O}$, denoted by $\partial_B \Phi(x)$, is the set of V such that

$$V = \lim_{k \to \infty} J\Phi(x^k) \,,$$

where $\{x^k\} \in \mathcal{D}_\Phi$ is a sequence converging to x. The Clarke's generalized Jacobian of Φ at x is the convex hull of $\partial_B \Phi(x)$ (see [3]), i.e.,

$$\partial \Phi(x) = \operatorname{conv}\{\partial_B \Phi(x)\} \,.$$

For $x \in \mathcal{O}$ and $u \in X$, the strict derivative $D_*\Phi(x)(u)$ of Φ at x in the direction u consists of all points $y \in Y$ which is the limit of a sequence

$$y^k := (\Phi(x^k + t_k u) - \Phi(x^k))/t_k, \quad x^k \to x, \; t_k \downarrow 0.$$

The set $D_*\Phi(x)(u)$ was first studied by Thibault [39, 40] (with a different notation) in order to extend Clarke's calculus to functions in abstract spaces. Kummer [15] called it "Thibault's directional derivative" and used it to get a complete characterization of a Lipschitz homeomorphism function (see Definition 2.1). Since then the strict derivative has been studied in Rockafellar and Wets [33, Ch.9] and Levy [18, 19] for both single valued functions and multifunctions.

It is known that $D_*\Phi(x)(u)$ is related to Clarke's generalized Jacobian $\partial\Phi$ and the B-subdifferential $\partial_B\Phi$ in the following way [15]:

$$\partial_B \Phi(x) u \subseteq D_*\Phi(x)(u) \subseteq \partial \Phi(x) u \,, \quad \forall\, u \in X \,. \tag{2.1}$$

In [15], Kummer gave a piecewise linear mapping to show that $D_*\Phi(x)(u) \subset \partial\Phi(x)u$ but $D_*\Phi(x)(u) \neq \partial\Phi(x)u$.

Definition 2.1. *A function $\Psi : \mathcal{O} \subseteq X \to X$ is said to be a locally Lipschitz homeomorphism near $x \in \mathcal{O}$ if there exists an open neighborhood $\mathcal{N} \subseteq \mathcal{O}$ of x such that the restricted map $\Psi\mid_{\mathcal{N}}: \mathcal{N} \to \Psi(\mathcal{N})$ is Lipschitz continuous and bijective, and its inverse is also Lipschitz continuous.*

The following inverse function theorem is obtained by Kummer [15].

Theorem 2.1. *Suppose that $\Psi : \mathcal{O} \subseteq X \to X$ is locally Lipschitz near $x \in \mathcal{O}$. Then Ψ is locally Lipschitz homeomorphism near x if and only if the following nonsingularity condition holds:*

$$\forall\, u \in X, \quad 0 \in D_*\Psi(x)(u) \Longrightarrow u = 0\,.$$

The purpose of this section is to show that if $\Psi : \mathcal{O} \subseteq X \to X$ is locally Lipschitz homeomorphism near $x \in \mathcal{O}$, then Ψ is semismooth at x if and only if Ψ^{-1}, the local inverse mapping of Ψ near x, is semismooth at $\Psi(x)$.

Semismoothness was originally introduced by Mifflin [22] for functionals. For studying the superlinear convergence of Newton's method for solving nonsmooth equations, Qi and Sun [30] extended the definition of semismoothness to vector valued functions. There are several equivalent ways for defining the semismoothness. Here we use the following definition.

Definition 2.2. *Let $\Phi : \mathcal{O} \subseteq X \to Y$ be a locally Lipschitz continuous function on the open set $\mathcal{O}$. We say that Φ is semismooth at a point $x \in \mathcal{O}$ if*

(i) Φ is directionally differentiable at x; and

(ii) for any $\Delta x \to 0$ and $V \in \partial\Phi(x + \Delta x)$,

$$\Phi(x + \Delta x) - \Phi(x) - V(\Delta x) = o(||\Delta x||)\,. \tag{2.2}$$

In the above definition on semismoothness, part (i) and part (ii) do not imply each other. To see that (i) does not imply (ii), one may consider the example given in [22]: $\Phi(x) = x^2 \sin(1/x)$ if $0 \neq x \in \Re$ and $\Phi(x) = 0$ if $x = 0$. On the other hand, Shapiro [35] constructed a one dimensional example to show that (ii) holds while (i) fails to hold. Condition (2.2), together with a nonsingularity assumption on $\partial\Phi$ at a solution point, was used by Kummer [14] to prove the superlinear convergence of Newton's method for locally Lipschitz equations. Gowda [10] called a locally Lipschitz continuous function Φ "semismooth" at x if (2.2) holds. To distinguish Gowda's definition on semismoothness of Φ at x, Pang et al. [26] called Φ to be G-semismooth at x if condition (2.2) holds. A stronger notion than semismoothness is strong semismoothness. We say that Φ is strongly G-semismooth (strongly semismooth) at x, if Φ is G-semismooth (semismooth) at x and for any $\Delta x \to 0$ and $V \in \partial\Phi(x + \Delta x)$,

$$\Phi(x + \Delta x) - \Phi(x) - V(\Delta x) = O(||\Delta x||^2). \tag{2.3}$$

We say that Φ is G-semismooth (semismooth, strongly G-semismooth, strongly semismooth) on a set $D \subseteq \mathcal{O}$ if Φ is G-semismooth (semismooth, strongly G-semismooth, strongly semismooth) at every point of D.

In order to show semismoothness, one often finds the following result useful. For a proof, see [38, Theorem 3.7] and [37, Lemma 2.1].

Lemma 2.1. *Let $\Phi : \mathcal{O} \subseteq X \to Y$ be locally Lipschitz near $x \in \mathcal{O}$. Then Φ is G-semismooth (strongly G-semismooth) at x if and only if for any $\Delta x \to 0$ and $x + \Delta x \in \mathcal{D}_\Phi$,*

$$\Phi(x+\Delta x) - \Phi(x) - J\Phi(x+\Delta x)(\Delta x) = o(||\Delta x||) \quad (= O(||\Delta x||^2)). \tag{2.4}$$

Lemma 2.2. *Let $\Psi : \mathcal{O} \subseteq X \to X$ be locally Lipschitz homeomorphism near $x_0 \in \mathcal{O}$. Then there exists an open neighborhood $\mathcal{N} \subseteq \mathcal{O}$ of x_0 such that Ψ is F-differentiable at $x \in \mathcal{N}$ if and only if Ψ^{-1}, the locally inverse mapping of Ψ near x_0, is F-differentiable at $y := \Psi(x)$; and*

$$(J\Psi(x))^{-1} = J\Psi^{-1}(y). \tag{2.5}$$

Moreover, there exists a positive number $\mu > 0$ such that $||V^{-1}|| \leq \mu$, $||W^{-1}|| \leq \mu$ for all $V \in \partial_B \Psi(u)$, $W \in \partial_B \Psi^{-1}(z)$, $u \in \mathcal{N}$, and $z \in \Psi(\mathcal{N})$.

Proof. By Definition 2.1, there exist a neighborhood $\mathcal{N} \subseteq \mathcal{O}$ of x_0 and a locally Lipschitz function Ψ^{-1} defined on the open neighborhood $\Psi(\mathcal{N})$ of $\Psi(x_0)$ such that $\Psi(\Psi^{-1}(z)) = z$ and $\Psi^{-1}(\Psi(u)) = u$ for any $u \in \mathcal{N}$ and $z \in \Psi(\mathcal{N})$. On the other hand, by Theorem 2.1 and (2.1), we know that any $V \in \partial_B \Psi(x_0)$ is nonsingular. Hence, by shrinking $\mathcal{N}$ if necessary, we know from [28] that there exists a positive number $\mu > 0$ such that $\max\{\|V\|, ||V^{-1}||\} \leq \mu$ for all $V \in \partial_B \Psi(u)$ and $u \in \mathcal{N}$.

Suppose that Ψ is F-differentiable at some point $x \in \mathcal{N}$. Then, because $J\Psi(x) \in \partial_B \Psi(x)$, $J\Psi(x)$ is nonsingular. If Ψ^{-1} is not F-differentiable at $y = \Psi(x)$, then there exists a sequence $\{\Delta y^k\}$ converging to 0 such that $\Delta y^k \neq 0$ and

$$\frac{\Delta_k}{||\Delta y^k||} \to \alpha \neq 0, \tag{2.6}$$

where $\Delta_k := \Psi^{-1}(y + \Delta y^k) - \Psi^{-1}(y) - (J\Psi(x))^{-1}(\Delta y^k)$. Then,

$$\Psi^{-1}(y + \Delta y^k) = \Psi^{-1}(y) + (J\Psi(x))^{-1}(\Delta y^k) + \Delta_k.$$

Now

$$\begin{aligned} & \Psi(\Psi^{-1}(y) + (J\Psi(x))^{-1}(\Delta y^k) + \Delta_k) - \Psi(x) \\ = \; & \Psi(\Psi^{-1}(y + \Delta y^k)) - \Psi(\Psi^{-1}(y)) \;=\; y + \Delta y^k - y \;=\; \Delta y^k. \end{aligned}$$

So, we obtain

$$\begin{aligned} & \Psi(\Psi^{-1}(y) + (J\Psi(x))^{-1}(\Delta y^k) + \Delta_k) - \Psi(x) \\ = \; & \Psi(x + (J\Psi(x))^{-1}(\Delta y^k) + \Delta_k) - \Psi(x) = \Delta y^k. \end{aligned}$$

Then, it follows from the F-differentiablity of Ψ at x and the local Lipschitz continuity of Ψ^{-1} that

$$J\Psi(x)[(J\Psi(x))^{-1}(\Delta y^k) + \Delta_k] + o(||\Delta y^k||) = \Delta y^k.$$

Thus

$$\Delta y^k + J\Psi(x)(\Delta_k) + o(||\Delta y^k||) = \Delta y^k,$$

which implies that

$$J\Psi(x)\left(\frac{\Delta_k}{||\Delta y^k||}\right) + o(1) = 0.$$

Hence, $J\Psi(x)(\alpha) = 0$. Therefore, $\alpha = 0$, which contradicts (2.6). This contradiction shows that Ψ^{-1} is F-differentiable at $y = \Psi(x)$ and $J\Psi^{-1}(y) = (J\Psi(x))^{-1}$.

Similarly, we can show the converse part.

Evidently, $||W^{-1}|| \leq \mu$ for all $W \in \partial_B \Psi^{-1}(z)$ and $z \in \Psi(\mathcal{N})$. $\square$

Theorem 2.2. *Let $\Psi : \mathcal{O} \subseteq X \to X$ be locally Lipschitz homeomorphism near $x_0 \in \mathcal{O}$. Then there exists an open neighborhood $\mathcal{N} \subseteq \mathcal{O}$ of x_0 such that*

(i) Ψ is G-semismooth (strongly G-semismooth) at some point $\bar{x} \in \mathcal{N}$ if and only if Ψ^{-1}, the local inverse mapping of Ψ near x_0, is G-semismooth (strongly G-semismooth) at $\bar{y} := \Psi(\bar{x})$; and

(ii) Ψ is semismooth (strongly semismooth) at some point $\bar{x} \in \mathcal{N}$ if and only if Ψ^{-1} is semismooth (strongly semismooth) at $\bar{y} := \Psi(\bar{x})$.

Proof. (i) By Lemma 2.2, there exist an open neighborhood $\mathcal{N} \subseteq \mathcal{O}$ of x_0 and a positive number $\mu > 0$ such that Ψ is F-differentiable at $x \in \mathcal{N}$ if and only if Ψ^{-1} is F-differentiable at $\Psi(x)$ and that $||V^{-1}|| \leq \mu$, $||W^{-1}|| \leq \mu$ for all $V \in \partial_B \Psi(x)$, $W \in \partial_B \Psi^{-1}(y)$, $x \in \mathcal{N}$, and $y \in \Psi(\mathcal{N})$. Then, by Lemma 2.2, for any $y \in \Psi(N) \cap \mathcal{D}_{\Psi^{-1}}$ and $y \to \bar{y}$ we have

$$\begin{aligned}
&\Psi^{-1}(y) - \Psi^{-1}(\bar{y}) - J\Psi^{-1}(y)(y - \bar{y}) \\
&= \Psi^{-1}(\Psi(x)) - \bar{x} - (J\Psi(x))^{-1}(\Psi(x) - \Psi(\bar{x})) \\
&= x - \bar{x} - (J\Psi(x))^{-1}(\Psi(x) - \Psi(\bar{x})) \\
&= (J\Psi(x))^{-1}[\Psi(x) - \Psi(\bar{x}) - J\Psi(x)(x - \bar{x})] \\
&= O\|\Psi(x) - \Psi(\bar{x}) - J\Psi(x)(x - \bar{x})\|,
\end{aligned} \tag{2.7}$$

where $x := \Psi^{-1}(y)$. Hence, Lemma 2.1, together with (2.7), shows that Ψ^{-1} is G-semismooth (strongly G-semismooth) at $\bar{y}$ if Ψ is G-semismooth (strongly G-semismooth) at $\bar{x}$.

By reversing the above arguments, we obtain that Ψ is G-semismooth (strongly G-semismooth) at $\bar{x}$ if Ψ^{-1} is G-semismooth (strongly G-semismooth) at $\bar{y}$.

(ii) By [17, Lemma 2], Ψ is directionally differentiable at $\bar{x}$ if and only Ψ^{-1} is directionally differentiable at $\Psi(\bar{x})$. Therefore, statement (ii) follows by statement (i) and the definitions of semismoothness and strong semismoothness. $\square$

Note that in [10], among many other results for inverse and implicit function theorems, by assuming that a locally Lipschitz continuous function $\Psi : \mathcal{O} \subseteq X \to X$ is semismooth on the open set $\mathcal{O}$, Gowda provided a necessary and sufficient condition for the existence of a semismooth inverse function of Ψ. Various equivalent forms to Gowda's condition were given in [26]. Here, in Theorem 2.2 we showed that a locally Lipschitz homeomorphism $\Psi : \mathcal{O} \subseteq X \to X$ is semismooth at a point if only if its inverse mapping is semismooth at its image point. A Lipschitz continuous function which is semismooth at a certain point may be not semismooth on an neighborhood of this point. For example, for any closed convex cone $D \subseteq X$, $\Pi_D(\cdot)$ is strongly semismooth at the origin [27] while we know nothing about the semismoothness of $\Pi_D(\cdot)$ at other points. Another simple example would be a locally Lipschitz function which is strictly Fréchet differentiable at one point. Then, by [30, Corollary 2.5], the Lipschitz function is semismooth at this point and there is no guarantee that it is semismooth on an open neighbordhood of this point. So, discussions given in Theorem 2.2 are necessary if we only know the semismoothness of Ψ at the point concerned.

Next, we consider the equation $H(x,u) = q$, where $H : X \times Y \to X$ is locally Lipschitz continuous near $(x_0, u_0) \in X \times Y$ with $H(x_0, u_0) = q_0$. Let us make the following assumption:

Assumption 2.1. *There exist an open neighborhood $\mathcal{N}(q_0, u_0) \subseteq X \times Y$ of (q_0, u_0) and a Lipschitz continuous function $x(q,u)$ defined on $\mathcal{N}(q_0, u_0)$ such that $H(x(q,u), u) = q$ for every $(q,u) \in \mathcal{N}(q_0, u_0)$.*

Under the above assumption, we have the following result.

Corollary 2.1. *Suppose that Assumption 2.1 holds. Then, $x(\cdot,\cdot)$ is G-semismooth (strongly G-semismooth, semismooth, strongly semismooth) at some point $(q,u) \in \mathcal{N}(q_0, u_0)$ if and only if H is G-semismooth (strongly G-semismooth, semismooth, strongly semismooth) at $(x(q,u), u)$.*

Proof. Define the mapping $\Psi : X \times Y \to X \times Y$ in the form of

$$\Psi(x,u) := \begin{bmatrix} H(x,u) \\ u \end{bmatrix}.$$

Then $\Psi(x_0, u_0) = \begin{bmatrix} q_0 \\ u_0 \end{bmatrix}$. By Assumption 2.1, Ψ is locally Lipschitz homeomorphism near (x_0, u_0) with its inverse mapping Ψ^{-1} given by

$$\Psi^{-1}(q,u) := \begin{bmatrix} x(q,u) \\ u \end{bmatrix}, \quad (q,u) \in \mathcal{N}(q_0, u_0).$$

Therefore, by Theorem 2.2, Ψ^{-1} is G-semismooth (strongly G-semismooth, semismooth, strongly semismooth) at some point $(q,u) \in \mathcal{N}(q_0, u_0)$ if and only if H is G-semismooth (strongly G-semismooth, semismooth, strongly semismooth) at $(x(q,u), u)$. This completes the proof. $\square$

In [16], Kummer showed that Assumption 2.1 is equivalent to the following nonsingularity condition:

$$\forall\, d \in X, \quad 0 \in D_* H(x_0, u_0)(d, 0) \Longrightarrow d = 0. \tag{2.8}$$

Then, we have the following implicit function theorem, which is a direct consequence of Corollary 2.1 and does not need a proof.

Corollary 2.2. *Suppose that the nonsingularity condition (2.8) is satisfied. Let $q_0 = 0$. Then, there exist an open neighborhood $\mathcal{N} \subseteq Y$ of u_0 and a Lipschitz continuous function $x(\cdot)$ defined on $\mathcal{N}$ such that $H(x(u), u) = 0$ for every $u \in \mathcal{N}$. Moreover, $x(\cdot)$ is G-semismooth (strongly G-semismooth, semismooth, strongly semismooth) at some point $u \in \mathcal{N}$ if H is G-semismooth (strongly G-semismooth, semismooth, strongly semismooth) at $(x(u), u)$.*

The first part of Corollary 2.2 first appeared in [16]. Based on Clarke's implicit function theorem [2, 3], a related result to the second part of the above corollary is obtained in [37, Theorem 2.1] under Clarke's nonsingularity condition:

$$\forall\, d \in X, \quad 0 \in \partial H(x_0, u_0)(d, 0) \Longrightarrow d = 0, \tag{2.9}$$

which is a more restrictive condition than (2.8) (cf. [15]).

3 Semismoothness of Solutions to Generalized Equations

Let X, Y, and U be finite dimensional vector spaces. Suppose that $F : X \times U \to X$ is continuously differentiable, $G : X \times U \to Y$ is twice continuously differentiable, and $K \subseteq Y$ is a closed convex set. Let $H : X \times Y \times U \to X \times Y$ be defined by (1.3). Now, let us consider the following parameterized generalized equation

$$(GEu) \qquad H(x, \lambda, u) = \begin{bmatrix} F(x, u) + J_x G(x, u)^* \lambda \\ G(x, u) - \Pi_K[G(x, u) + \lambda] \end{bmatrix} = 0\,. \tag{3.1}$$

For a given point $u_0 \in U$ of the parameter vector, we view the corresponding generalized equation (GE_{u_0}) as unperturbed, and write (GE_u) for the parameterized generalized equation. Let (x_0, λ_0) be a solution of (GE_{u_0}).

For $u \in U$ near u_0, let $(x(u), \lambda(u))$ (if exists) solve (GE_u), i.e., $H(x(u), \lambda(u), u) = 0$. In [36], Shapiro studied some perturbed properties of $(x(\cdot), \lambda(\cdot))$ near $u_0 \in U$. In this section, we will investigate its semismoothness at u_0. Denote $S := \{x \in X \,|\, G(x, u_0) \in K\}$. Let $T_K(y)$ denote the tangent cone of K at y and $\mathrm{lin}(C)$ the lineality space of the closed convex cone C, i.e., $\mathrm{lin}(C) = C \cap (-C)$. The following definition of nondegeneracy is taken from [1, 36], which is a basic assumption of this section.

Definition 3.1. *We say that a point $x_0 \in S$ is nondegenerate, with respect to the mapping G and the set K, if*

$$J_x G(x_0, u_0) X + \mathrm{lin}(T_K(y_0)) = Y, \tag{3.2}$$

where $y_0 := G(x_0, u_0)$.

Under the above nondegeneracy assumption at $x_0 \in S$, Shapiro [36] showed that the multiplier λ satisfying $H(x_0, \lambda, u_0) = 0$ is unique.

By using the cone reducibility notion, Shapiro [36] reduced the discussion on sensitivity analysis of $(x(\cdot), \lambda(\cdot))$ to a new problem. In this case, the sensitivity analysis becomes simpler, at least notationally. In the following analysis, we will adopt this idea. The concept of cone reducibility below is taken from [1, 36].

Definition 3.2. *A closed (not necessarily convex) set $C \subseteq Y$ is called cone reducible at a point $y_0 \in C$ if there exist a neighborhood $\mathcal{V} \subseteq Y$ of y_0, a pointed closed convex cone Q in a finite dimensional space Z and a twice continuously differentiable mapping $\Xi : \mathcal{V} \to Z$ such that: (i) $\Xi(y_0) = 0 \in Z$, (ii) the derivative mapping $J\Xi(y_0) : Y \to Z$ is onto, and (iii) $C \cap \mathcal{V} = \{y \in \mathcal{V} \,|\, \Xi(y) \in Q\}$. If C is cone reducible at every point $y_0 \in C$ (possibly to a different cone Q), then we say that C is cone reducible.*

Many interesting sets such as the polyhedral convex set, the second-order cone, and the cone S^n_+ of positive semidefinite $n \times n$ symmetric matrices are all cone reducible [1, 36]. In the subsequent analysis, we assume that the convex set K is cone reducible at the point $y_0 := G(x_0, u_0)$ to a pointed closed convex cone $Q \subseteq Z$ by a mapping Ξ. Define the mapping $\mathcal{G}(x, u) := \Xi(G(x, u))$. Then, it is known [1, 36] that for all (x, u) in a neighborhood of (x_0, u_0), the generalized equations (GE_u) can be written in the following equivalent form

$$(\mathcal{G}E_u) \qquad H_{\mathcal{G}}(x, \mu, u) := \begin{bmatrix} F(x, u) + J_x\mathcal{G}(x, u)^*\mu \\ \\ \mathcal{G}(x, u) - \Pi_Q(\mathcal{G}(x, u) + \mu) \end{bmatrix} = 0 \tag{3.3}$$

in the sense that locally, $(x(u), \lambda(u))$ is a solution of (GE_u) if and only if $(x(u), \mu(u))$ is a solution of $(\mathcal{G}E_u)$ and

$$\lambda(u) = [J\Xi(G(x(u), u))]^*\mu(u)\,. \tag{3.4}$$

Moreover, by Definition 3.2, we can derive that for (x, u) sufficiently close to (x_0, u_0), the multiplier $\mu(u)$ is defined uniquely by (3.4). In particular, the unperturbed problem $(\mathcal{G}E_{u_0})$ has solution (x_0, μ_0) with μ_0 being uniquely determined by

$$\lambda_0 = [J\Xi(G(x_0, u_0))]^*\mu_0\,.$$

Hence, in what follows, we only need to study the semismooth sensitivity of the solution of $(\mathcal{G}E_u)$ near u_0. By Definition 3.2, we have $\mathcal{G}(x_0, u_0) = 0$ with the unique multiplier μ_0.

Theorem 3.1. *Let (x_0, μ_0) be a solution of $(\mathcal{G}E_{u_0})$. Suppose that the convex set K is cone reducible at the point $y_0 := G(x_0, u_0)$ to a pointed convex closed cone $Q \subseteq Z$ by a mapping Ξ. Suppose that the following condition holds:*

$$\forall\ (\Delta x, \Delta\mu) \in X \times Z, \quad 0 \in D_*H_{\mathcal{G}}(x_0, \mu_0, u_0)(\Delta x, \Delta\mu, 0) \Longrightarrow \Delta x = 0, \Delta\mu = 0. \tag{3.5}$$

The following statements hold:

(i) there exist an open neighborhood $\mathcal{N}$ of u_0 and a Lipschitz continuous function $(x(\cdot), \mu(\cdot))$ defined on $\mathcal{N}$ such that $H_{\mathcal{G}}(x(u), \mu(u), u) = 0$ for every $u \in \mathcal{N}$;

(ii) if Π_Q is G-semismooth (semismooth) at $\mathcal{G}(x_0, u_0)+\mu_0$, then $(x(\cdot), \mu(\cdot))$ is G-semismooth (semismooth) at u_0; and

(iii) if Π_Q is strongly G-semismooth (strongly semismooth) at $\mathcal{G}(x_0, u_0)+\mu_0$ and the derivative of F and the second derivative of $\mathcal{G}$ are locally Lipschitz continuous near (x_0, u_0), then $(x(\cdot), \mu(\cdot))$ is strongly G-semismooth (strongly semismooth) at u_0.

Proof. Statements (i) and (ii) are direct consequences of Corollary 2.2.

For statement (iii), we observe that $H_{\mathcal{G}}$ is strongly G-semismooth (strongly semismooth) at (x_0, μ_0, u_0) if Π_Q is strongly G-semismooth (strongly semismooth) at $\mathcal{G}(x_0, u_0) + \mu_0$ and the derivative of F and the second derivative of $\mathcal{G}$ are locally Lipschitz continuous near (x_0, u_0) (cf. [8]). Then, by the same corollary, we get (iii). □

Condition (3.5) used in Theorem 3.1 is mild according to discussions in Section 2, but may be difficult to verify in general. Next, we will consider a sufficient condition for guaranteeing (3.5) to hold. For this purpose, we need the characterizations of Clarke's generalized Jacobian of the metric projector $\Pi_K(\cdot)$, which is Lipschitz continuous with Lipschitz constant 1. Define $\phi : X \to R$ by

$$\phi(y) := \frac{1}{2}[\langle y, y\rangle - \langle y - \Pi_K(y), y - \Pi_K(y)\rangle], \quad y \in X\,. \tag{3.6}$$

It is known [42] that $\theta(y) := \frac{1}{2}||y - \Pi_K(y)||^2,\ y \in X$ is continuously differentiable with

$$\nabla\theta(y) = y - \Pi_K(y)\,, \quad y \in X.$$

Thus, ϕ is also continuously differentiable with

$$\nabla\phi(y) = y - (y - \Pi_K(y)) = \Pi_K(y)\,, \quad y \in X\,.$$

Therefore, by mimicing the proof in [25, 3.3.4], we can get the following result. We omit the details here for brevity.

Lemma 3.1. *Let $K \subseteq X$ be a closed convex set. If Π_K is F-differentiable at $x \in X$, then $J\Pi_K(x)$ is self-adjoint, i.e.,*

$$J\Pi_K(x) = J\Pi_K(x)^*. \tag{3.7}$$

Proposition 3.1. *Let $K \subseteq X$ be a closed convex set. Then, for any $x \in X$ and $V \in \partial\Pi_K(x)$, we have*

(i) V is self-adjoint;

(ii) $\langle d, Vd\rangle \geq 0,\ \forall\, d \in X$; and

(iii) $\langle Vd, d - Vd\rangle \geq 0,\ \forall\, d \in X$.

Proof. (i) By Lemma 3.1 and the definition of $\partial_B\Pi_K(x)$, any $V \in \partial_B\Pi_K(x)$ is self-adjoint. This further implies that any $V \in \partial\Pi_K(x) = \text{conv}\partial_B\Pi_K(x)$ is self-adjoint.

(ii) Suppose that $\Pi_K(\cdot)$ is F-differentiable at some point $\bar{x} \in X$. Since $\Pi_K(\cdot)$ is monotone [42], we have

$$\langle \Pi_K(\bar{x} + td) - \Pi_K(\bar{x}), td\rangle \geq 0, \quad \text{for all } t \geq 0,$$

which implies that for all $d \in X$,

$$\langle J\Pi_K(\bar{x})d, d\rangle \geq 0.$$

Hence, by the definition of $\partial\Pi_K(\cdot)$, for any $x \in X$ and $V \in \partial\Pi_K(x)$,

$$\langle d, Vd\rangle \geq 0, \quad \forall d \in X\,.$$

(iii) First, we consider $y \in \mathcal{D}_{\Pi_K}$. By [42], for any $d \in X$ and $t \geq 0$, we have

$$\langle \Pi_K(y + td) - \Pi_K(y), td\rangle \geq ||\Pi_K(y + td) - \Pi_K(y)||^2, \quad \text{for all } t \geq 0.$$

Hence,

$$\langle J\Pi_K(y)d, d\rangle \geq \langle J\Pi_K(y)d, J\Pi_K(y)d\rangle. \tag{3.8}$$

Next, let $V \in \partial\Pi_K(x)$. Then, by Carathéodory's theorem, there exist a positive integer $\kappa > 0$, $V^i \in \partial_B\Pi_K(x)$, $i = 1, 2, \dots, \kappa$ such that

$$V = \sum_{i=1}^{\kappa} \lambda_i V^i \,,$$

where $\lambda_i \geq 0$, $i = 1, 2, \dots, \kappa$, and $\sum_{i=1}^{\kappa} \lambda_i = 1$. Let $d \in X$. For each $i = 1, \dots, \kappa$ and $k = 1, 2, \dots$, there exists $x^{i_k} \in \mathcal{D}_{\Pi_K}$ such that

$$||x - x^{i_k}|| \leq 1/k$$

and

$$||J\Pi_K(x^{i_k}) - V^i|| \leq 1/k$$

By (3.8), we have

$$\langle J\Pi_K(x^{i_k})d, d\rangle \geq \langle J\Pi_K(x^{i_k})d, J\Pi_K(x^{i_k})d\rangle.$$

Hence,

$$\langle V^i d, d\rangle \geq \langle V^i d, V^i d\rangle,$$

and so,

$$\sum_{i=1}^{\kappa} \lambda_i \langle V^i d, d\rangle \geq \sum_{i=1}^{\kappa} \lambda_i \langle V^i d, V^i d\rangle. \tag{3.9}$$

Define $\theta(x) := ||x||^2$, $x \in X$. By the convexity of θ, we have

$$\theta\Big(\sum_{i=1}^{\kappa} \lambda_i V^i d\Big) \leq \sum_{i=1}^{\kappa} \lambda_i \theta(V^i d) = \sum_{i=1}^{\kappa} \lambda_i \langle V^i d, V^i d\rangle = \sum_{i=1}^{\kappa} \lambda_i ||V^i d||^2 \,.$$

Hence,

$$\sum_{i=1}^{\kappa} \lambda_i ||V^i d||^2 \geq \Big\langle \sum_{i=1}^{\kappa} \lambda_i V^i d, \sum_{i=1}^{\kappa} \lambda_i V^i d\Big\rangle. \tag{3.10}$$

By using (3.9) and (3.10), we obtain for all $d \in X$ that

$$\langle Vd, d\rangle \geq \langle Vd, Vd\rangle.$$

The proof is completed. $\square$

Set $\mathcal{L}_0 := \{z \in X \mid J_x\mathcal{G}(x_0, u_0)z \in \mathcal{C}_{\Pi_Q(\mathcal{G}(x_0,u_0)+\mu_0)}\}$, where $\mathcal{C}_{\Pi_Q(\mathcal{G}(x_0,u_0)+\mu_0)} := \{Vh \mid V \in \partial\Pi_Q(\mathcal{G}(x_0, u_0) + \mu_0), h \in Z\}$. For a general closed convex set K, it is not clear what $\mathcal{L}_0$ looks like. But, if K has special structures, then it is likely to know $\mathcal{L}_0$ exactly. For example, $\partial\Pi_Q$, and so $\mathcal{L}_0$, has a complete characterization if Q is either a second order cone or $\mathcal{S}^n_+$ [26].

Proposition 3.2. *Suppose that $x_0 \in S$ is nondegenerate, with respect to the mapping G and the closed convex set K, and K is cone reducible at the point $y_0 := G(x_0, u_0)$ to a pointed closed convex cone $Q \subseteq Z$ by a mapping Ξ. Suppose that*

$$\forall\, 0 \neq \Delta x \in \mathcal{L}_0 \Longrightarrow \langle \Delta x, (J_xF(x_0, u_0) + J^2_{xx}\mathcal{G}(x_0, u_0)^*\mu_0)(\Delta x)\rangle > 0\,. \tag{3.11}$$

Then, the nonsingularity condition (3.5) in Theorem 3.1 holds.

Proof. Let $(\Delta x, \Delta\mu) \in X \times Z$ be such that

$$0 \in \partial H_{\mathcal{G}}(x_0, \mu_0, u_0)(\Delta x, \Delta\mu, 0)\,.$$

Then, according to the definition of $H_{\mathcal{G}}$ and Clarke [3], there exists $V \in \partial\Pi_Q(\mathcal{G}(x_0, u_0)+\mu_0)$ such that

$$\begin{cases} [J_x F(x_0, u_0) + J_{xx}^2 \mathcal{G}(x_0, u_0)^* \mu_0](\Delta x) + J_x\mathcal{G}(x_0, u_0)^*(\Delta\mu) = 0, \\ \\ J_x\mathcal{G}(x_0, u_0)(\Delta x) - V[J_x\mathcal{G}(x_0, u_0)(\Delta x) + \Delta\mu] = 0. \end{cases} \tag{3.12}$$

Let $\Delta H := J_x\mathcal{G}(x_0, u_0)(\Delta x) + \Delta\mu$. Then we have

$$\Delta H - \Delta\mu = J_x\mathcal{G}(x_0, u_0)(\Delta x) = V(\Delta H).$$

So,

$$\langle V(\Delta H), \Delta H - V(\Delta H)\rangle - \langle V(\Delta H), \Delta\mu\rangle = 0. \tag{3.13}$$

Since, by Proposition 3.1, the first term on the left hand side of (3.13) is nonnegative, we have

$$\langle V(\Delta H), \Delta\mu\rangle \geq 0.$$

It follows from (3.12) that

$$\begin{aligned} \langle V(\Delta H), \Delta\mu\rangle &= \langle J_x\mathcal{G}(x_0, u_0)(\Delta x), \Delta\mu\rangle \\ &= \langle \Delta x, J_x\mathcal{G}(x_0, u_0)^*(\Delta\mu)\rangle \\ &= -\langle \Delta x, [J_x F(x_0, u_0) + J_{xx}^2\mathcal{G}(x_0, u_0)^*\mu_0](\Delta x)\rangle\,. \end{aligned}$$

Thus,

$$\langle \Delta x, [J_x F(x_0, u_0) + J_{xx}^2\mathcal{G}(x_0, u_0)^*\mu_0](\Delta x)\rangle \leq 0.$$

From the second equation of (3.12), we know that $\Delta x \in \mathcal{L}_0$. Then, by condition (3.11), we obtain

$$\Delta x = 0.$$

Hence

$$J_x\mathcal{G}(x_0, u_0)^*(\Delta\mu) = 0.$$

Since Q is a pointed closed convex cone, under the assumptions, from [36] we know that the mapping $J_x\mathcal{G}(x_0, u_0)$ is onto. Thus, $\Delta\mu = 0$. This, together with (2.1), shows that the nonsingularity condition (3.5) in Theorem 3.1 holds. □

By using Proposition 3.2 and Theorem 3.1, we get the following result:

Corollary 3.1. *Suppose that $x_0 \in S$ is nondegenerate, with respect to the mapping G and the closed convex set K, and K is cone reducible at the point $y_0 := G(x_0, u_0)$ to a pointed closed convex cone $Q \subseteq Z$ by a mapping Ξ. Then all the conclusions of Theorem 3.1 hold if condition (3.5) in Theorem 3.1 is replaced by (3.11).*

Let $X = \mathcal{S}^n$ and $K := \mathcal{S}^n_+$. Let $(x_0, u_0) \in X \times U$ be such that $G(x_0, u_0) \in K$ of rank r. Then, by [1, Example 3.140], one can construct an infinitely many differentiable (even analytic) reduction mapping from a neighborhood of $G(x_0, u_0)$ into a linear space $\mathcal{S}^{n-r}$ with the corresponding cone $Q := \mathcal{S}^{n-r}_+$. Since, in this case $\Pi_Q(\cdot)$ is strongly semismooth everywhere [38], we can study the semismoothness of $(x(\cdot), \mu(\cdot))$ by Theorem 3.1 and Corollary 3.1.

4 The Moreau-Yosida Regularization

Let X be a finite dimensional vector space and and $f : X \to R \cup \{+\infty\}$ be a lower semicontinuous proper convex function. Let $\varepsilon > 0$ be a positive number. Let $\hat{f}_\varepsilon$ be the Moreau-Yosida regularization of f defined by (1.4). For any $u \in X$, let $x(u)$ denote the unique optimal solution of (1.4). Define $F : X \to X$ by

$$F(u) := \nabla \hat{f}_\varepsilon(u) = \varepsilon(u - x(u)), \quad u \in X. \tag{4.1}$$

The function F is globally Lipschitz continuous because $x(\cdot)$ is so [33, p.546]. However, F may fail to be semismooth. To see this, let us consider the indicator function $f(x) = \delta(x|D)$ defined as $f(x) = 0$ if $x \in D$ and $f(x) = +\infty$ if $x \notin D$, where $D \subset X$ is a nonempty closed convex set. The corresponding Moreau-Yosida regularization of f can be written as

$$\begin{aligned} & \min \quad \{\delta(x|D) + \frac{\varepsilon}{2}\langle u - x, u - x\rangle \,|\, x \in X\} \\ = & \min \quad \{\frac{\varepsilon}{2}\langle u - x, u - x\rangle \,|\, x \in D\}. \end{aligned} \tag{4.2}$$

Hence,

$$F(u) = \varepsilon(u - \Pi_D(u)), \; u \in X.$$

Since $\Pi_D(\cdot)$ is not everywhere directionally differentiable on X [13, 34], it follows that F is not semismooth everywhere. Actually, the same example given in [34] can also be used to show that F fails to be G-semismooth everywhere on X.

Regarding the Moreau-Yosida regularization, see [12, 33] for more general properties and [9, 20, 21, 23, 29] for second order properties. In this section, we will study the semismoothness of F by virtue of the projection onto the epigraph of f. It is evident that (1.4) can be rewritten as

$$\begin{aligned} \hat{f}_\varepsilon(u) = \quad & \min \quad \varepsilon\{\frac{1}{\varepsilon}f(x) + \frac{1}{2}\langle u - x, u - x\rangle\} \\ & \text{s.t.} \quad x \in X. \end{aligned} \tag{4.3}$$

Then, for any $u \in X$, $\hat{f}_\varepsilon(u)$ is the optimal value of the following parameterized optimization problem

$$\begin{aligned} \min \quad & \varepsilon\,\{\hat{t} + \frac{1}{2}\langle u - x, u - x\rangle\} \\ \text{s.t.} \quad & \varepsilon\hat{t} \geq f(x), \\ & (x, \hat{t}) \in Z, \end{aligned} \tag{4.4}$$

where $Z := X \times R$. Let Ω be the epigraph of f, i.e.,

$$\Omega := \text{epi}(f) = \{(x, t) \in X \times R \,|\, t \geq f(x)\}.$$

The set Ω is a closed convex set [32, p.51]. Then, problem (1.4) can be written in the form of

$$\begin{aligned} \min \quad & \{\frac{1}{\varepsilon}t + \frac{1}{2}\langle u - x, u - x\rangle\}, \\ \text{s.t.} \quad & (x, t) \in \Omega, \end{aligned} \tag{4.5}$$

which has a unique optimal solution $(x(u), t(u))$, where $t(u) := f(x(u))$. In the following, we will investigate the semismoothness of $(x(\cdot), t(\cdot))$.

Let $H : X \times R \times X \to X \times R$ be defined by

$$H(x,t,u) := \begin{bmatrix} x \\ t \end{bmatrix} - \Pi_\Omega\left(G(x,t,u)\right), \tag{4.6}$$

where

$$G(x,t,u) := \begin{bmatrix} x \\ t \end{bmatrix} - \begin{bmatrix} x-u \\ 1/\varepsilon \end{bmatrix} = \begin{bmatrix} u \\ t-1/\varepsilon \end{bmatrix}. \tag{4.7}$$

Then, for any $u \in X$, we have

$$H(x(u),t(u),u) = 0\,. \tag{4.8}$$

Lemma 4.1. *For any $u \in X$, $G(x(u),t(u),u) \notin \Omega$.*

Proof. Suppose on the contrary that for some $u \in X$, $G(x(u),t(u),u) \in \Omega$. Then we have

$$\Pi_\Omega(G(x(u),t(u),u)) = G(x(u),t(u),u) = \begin{bmatrix} x(u) \\ t(u) \end{bmatrix} - \begin{bmatrix} x(u)-u \\ 1/\varepsilon \end{bmatrix}.$$

So, it follows from (4.6) that

$$\begin{aligned} H(x(u),t(u),u) = & \begin{bmatrix} x(u) \\ t(u) \end{bmatrix} - \left[\begin{pmatrix} x(u) \\ t(u) \end{pmatrix} - \begin{pmatrix} x(u)-u \\ 1/\varepsilon \end{pmatrix}\right] \\ = & \begin{bmatrix} x(u)-u \\ 1/\varepsilon \end{bmatrix} \neq 0, \end{aligned}$$

which leads to a contradiction with (4.8). Hence, we have $G(x(u),t(u),u) \notin \Omega$. $\square$

Lemma 4.2. *Let C be a closed convex set of a finite dimensional space Y and w be a point in Y with $w \notin C$. Let $x := \Pi_C(w)$. Let $P := \{s \in Y \mid \langle s, w-x\rangle = \alpha\}$ be the supporting hyperplane of C at x such that $C \subseteq P_- := \{s \in Y \mid \langle s, w-x\rangle \leq \alpha\}$, where $\alpha := \langle x, w-x\rangle$. Then for any $z \notin P_-$, one has $||\Pi_C(z)-x|| \leq ||\Pi_P(z)-x||$.*

Proof. Let $y := \Pi_C(z)$. Then, we have

$$\langle z-y, y-c\rangle \geq 0, \quad \forall\, c \in C\,,$$

which, implies

$$\begin{aligned} ||z-x||^2 &= ||z-y||^2 + ||y-x||^2 + 2\langle z-y, y-x\rangle \\ &\geq ||z-y||^2 + ||y-x||^2. \end{aligned} \tag{4.9}$$

Now, since $y \in C \subseteq P_-$ and $z \notin P_-$, we have

$$\langle y, w-x\rangle \leq \alpha < \langle z, w-x\rangle\,.$$

Thus, there exists $\theta \in [0,1]$ such that $\langle \theta z + (1-\theta)y, w-x\rangle = \alpha$, i.e., $\theta z + (1-\theta)y \in P$. Hence,

$$||z - \Pi_P(z)|| \leq \|z - [\theta z + (1-\theta)y]\| \leq \|z-y\|\,,$$

which, together with (4.9), implies

$$\begin{aligned} ||z-y||^2 + ||\Pi_P(z) - x||^2 &\geq ||z - \Pi_P(z)||^2 + ||\Pi_P(z) - x||^2 \\ &= ||z-x||^2 \geq ||z-y||^2 + ||y-x||^2. \end{aligned}$$

Hence, the desired result is valid. $\square$

According to Proposition 3.1 and Lemma 4.2, we derive the following result regarding the B-subdifferential of Π_Ω. For any $z = (x,t) \in X \times R$, let $\Pi_\Omega(z)_x \in X$ and $\Pi_\Omega(z)_t \in \Re$ be such that $\Pi_\Omega(z) = (\Pi_\Omega(z)_x, \Pi_\Omega(z)_t)$. Let $\mathrm{int}(\mathrm{dom}(f))$ denote the interior part of the domain of f.

Proposition 4.1. *Let $z' = (x', t') \in X \times R$ be such that $z' \notin \Omega$ and $\Pi_\Omega(z')_x \in \mathrm{int}(\mathrm{dom}(f))$. Then, there exists $0 < \eta < 1$ such that for any $W \in \partial_B \Pi_\Omega(z')$,*

$$0 \leq \langle e, We \rangle \leq \eta < 1, \tag{4.10}$$

where $e := (0,1) \in X \times R$.

Proof. By the assumption that $z' \notin \Omega$, there exists $\bar{x}' \in X$ such that $(\bar{x}', f(\bar{x}')) = \Pi_\Omega(z')$. Let $\bar{z}' := (\bar{x}', f(\bar{x}'))$. Because f is locally Lipschitz continuous near $\bar{x}'$, there exist a neighborhood $\mathcal{N}_1(\bar{x}')$ of $\bar{x}'$ and a constant $\sigma > 0$ such that

$$\|\xi\| \leq \sigma < \infty, \quad \forall\, \xi \in \partial f(\bar{x}),\, \bar{x} \in \mathcal{N}_1(\bar{x}').$$

Since Π_Ω is globally Lipschitz continuous, there exists a neighborhood $\mathcal{N}_2(z')$ of z' such that for any $z \in \mathcal{N}_2(z')$,

$$z \notin \Omega \quad \text{and} \quad \Pi_\Omega(z)_x \in \mathcal{N}_1(\bar{x}').$$

For any $W \in \partial_B \Pi_\Omega(z')$, there exists a sequence, say $\{z^k\} \subset \mathcal{D}_{\Pi_\Omega}$, satisfying $z^k \to z'$ and $W = \lim_{k\to\infty} J\Pi_\Omega(z^k)$. Thus, by (ii) of Proposition 3.1, to show (4.10), it only needs to show that there exists $0 < \eta < 1$ such that

$$\langle e, J\Pi_\Omega(z)e \rangle \leq \eta, \qquad \forall\, z \in \mathcal{N}_2(z') \cap \mathcal{D}_{\Pi_\Omega}.$$

Let $z \in \mathcal{N}_2(z')$. Then, there exists $\bar{x} \in X$ such that

$$\bar{z} := \begin{bmatrix} \bar{x} \\ f(\bar{x}) \end{bmatrix} = \Pi_\Omega(z).$$

For any given $\Delta t > 0$, let $z + \Delta z := (x, t + \Delta t)$. Then, for all $\Delta t > 0$ sufficiently small, $z + \Delta z \notin \Omega$ and there exists $\hat{x}_t \in X$ such that

$$\hat{z}_t := \begin{bmatrix} \hat{x}_t \\ f(\hat{x}_t) \end{bmatrix} = \Pi_\Omega(z + \Delta z).$$

Hence, for any $z \in \mathcal{N}_2(z') \cap \mathcal{D}_{\Pi_\Omega}$,

$$\langle e, J\Pi_\Omega(z)e \rangle = \lim_{\Delta t \downarrow 0} \frac{\Pi_\Omega(z + \Delta z)_t - \Pi_\Omega(z)_t}{\Delta t} = \lim_{\Delta t \downarrow 0} \frac{f(\hat{x}_t) - f(\bar{x})}{\Delta t}.$$

This means that, to show (4.10), it suffices to show that there exists $0 < \eta < 1$ such that for all $\Delta t > 0$ sufficiently small,

$$|f(\hat{x}_t) - f(\bar{x})| \leq \eta|\Delta t|, \qquad \forall\, z \in \mathcal{N}_2(z') \cap \mathcal{D}_{\Pi_\Omega}.$$

Let $P := \{(s,\tau) \in X \times R \,|\, \langle (s,\tau), (x-\bar{x}, t - f(\bar{x}))\rangle = \alpha\}$ be the supporting plane P of Ω at $(\bar{x}, f(\bar{x}))$ such that

$$\Omega \subseteq P_- := \{(s,\tau) \in X \times R \,|\, \langle (s,\tau), (x-\bar{x}, t - f(\bar{x}))\rangle \leq \alpha\},$$

where $\alpha := \langle (\bar{x}, f(\bar{x})), (x-\bar{x}, t-f(\bar{x}))\rangle$. By shrinking $\mathcal{N}(z')$ if necessary, we may assume $z \notin P_-$ because $z' \notin P_-$. Then, for all $\Delta t > 0$ sufficiently small, $z + \Delta z = (x, t+\Delta t) \notin P_-$ and from Lemma 4.2,

$$|f(\hat{x}_t) - f(\bar{x})| \leq ||(\hat{x}_t, f(\hat{x}_t)) - (\bar{x}, f(\bar{x}))|| = \|\hat{z}_t - \bar{z}\| \leq ||\Pi_P(z+\Delta z) - \bar{z}||.$$

Let θ_t be the angle between $\bar{z} - z$ and $(z+\Delta z) - z$. Since both $(z+\Delta z) - \Pi_P(z+\Delta z)$ and $z - \bar{z}$ are perpendicular to the hyperplane P, we have for all $\Delta t > 0$ sufficiently small that

$$|f(\hat{x}_t) - f(\bar{x})| \leq \|\Pi_P(z+\Delta z) - \bar{z}\| = \|(z+\Delta z) - z\| \sin\theta_t = |\Delta t| \sin\theta_t.$$

It remains to estimate the value of $\sin\theta_t$. Since Ω is the epigraph of f and $\Omega \subseteq P_-$, we have for all $s \in X$ and $\tau \geq f(s)$ that

$$\langle s - \bar{x}, x - \bar{x}\rangle + (\tau - f(\bar{x}))(t - f(\bar{x})) \leq 0,$$

which, implies that $t - f(\bar{x}) < 0$ and for all $s \in X$,

$$f(s) - f(\bar{x}) \geq \langle s - \bar{x}, (x-\bar{x})/(f(\bar{x}) - t)\rangle.$$

Hence, $\xi := (x - \bar{x})/(f(\bar{x}) - t) \in \partial f(\bar{x})$ and for all $\Delta t > 0$ sufficiently small,

$$|\cos\theta_t| = \left|\frac{\langle z - \bar{z}, \Delta z)\rangle}{\|z-\bar{z}\|\|\Delta z\|}\right| = \left|\frac{\langle (\xi, -1), (0, \Delta t)\rangle}{\|(\xi,-1)\|}\right| = \frac{1}{\sqrt{\|\xi\|^2 + 1}}.$$

It follows that for all $\Delta t > 0$ sufficiently small,

$$\sin\theta_t = \sqrt{1 - \cos^2\theta_t} = \frac{\|\xi\|}{\sqrt{\|\xi\|^2+1}}.$$

Hence, for all $\Delta t > 0$ sufficiently small,

$$\sin\theta_t \leq \frac{\sigma}{\sqrt{\sigma^2+1}} =: \eta < 1.$$

This completes the proof. $\square$

Proposition 4.2. *For $u_0 \in X$, let $x_0 := x(u_0)$ and $t_0 := f(x(u_0))$. Suppose that $\Pi_\Omega(G(x_0,t_0,u_0))_x \in \mathrm{int}(\mathrm{dom}(f))$. Then, it holds that*

$$\forall\; h_x \in X,\; h_t \in \Re, \quad 0 \in \partial H(x_0,t_0,u_0)(h_x,h_t,0) \Longrightarrow h_x = 0,\; h_t = 0. \tag{4.11}$$

Proof. Let $h_x \in X$ and $h_t \in \Re$ be such that

$$0 \in \partial H(x_0, t_0, u_0)(h_x, h_t, 0)\,.$$

Let

$$h := \begin{bmatrix} h_x \\ h_t \end{bmatrix}.$$

Then, according to the definition of H and [3], there exists $W \in \partial \Pi_\Omega(G(x_0, t_0, u_0))$ such that

$$0 = \begin{bmatrix} h_x \\ h_t \end{bmatrix} - W(0, h_t) = 0. \tag{4.12}$$

By Lemma 4.1, $G(x_0, t_0, u_0) \notin \Omega$. So, according to Proposition 4.1 and noticing that W is a convex combination of some finitely many elements in $\partial_B \Pi_\Omega(G(x_0, t_0, u_0)$, we have

$$0 \leq \langle e, We \rangle \leq \eta < 1, \tag{4.13}$$

where e and η are defined as in Proposition 4.1. By (4.12), we obtain

$$|h_t|^2 = \langle (0, h_t), (h_x, h_t) \rangle = \langle (0, h_t), W(0, h_t) \rangle = |h_t|^2 \langle e, We \rangle\,,$$

which, together with (4.13), implies $h_t = 0$. By virtue of (4.12), $h_x = 0$. Therefore, condition (4.11) holds. $\square$

By (2.1), Corollary 2.2, and Proposition 4.2, we get the following result immediately.

Theorem 4.1. *For $u_0 \in X$, let $x_0 := x(u_0)$ and $t_0 := f(x(u_0))$. Then $(x(\cdot), t(\cdot))$ and $F(\cdot)$ are G-semismooth (strongly G-semismooth, semismooth, strongly semismooth) at u_0 if $\Pi_\Omega(G(x_0, t_0, u_0))_x \in \mathrm{int}(\mathrm{dom}(f))$ and $\Pi_\Omega(\cdot)$ is G-semismooth (strongly G-semismooth, semismooth, strongly semismooth) at $G(x_0, t_0, u_0)$.*

The significance of Theorem 4.1 is that it allows us to exploit the rich properties of the metric projector over closed convex sets (cf. [6, Ch.4]) to study the semismoothness of solutions to the Moreau-Yosida regularization. The condition $\Pi_\Omega(G(x_0, t_0, u_0))_x \in \mathrm{int}(\mathrm{dom}(f))$ in Theorem 4.1 holds automatically if f is finite valued everywhere.

5 Conclusions

In this paper, we showed that a locally Lipschitz homeomorphism function is G-semismooth (strongly G-semismooth, semismooth, strongly semismooth) at a given point if and only if its inverse function is G-semismooth (strongly G-semismooth, semismooth, strongly semismooth) at its image point. We then used this result and its corollaries to study the semismoothness of solutions to parameterized generalized equations and the Moreau-Yosida regularization of a convex function.

By analyzing the properties of Clarke's generalized Jacobian of the metric projector over closed convex sets, we presented a sufficient condition for the semismoothness of solutions to parameterized generalized equations over cone reducible (nonpolyhedral) convex sets. It would be interesting to see results relaxing this sufficient condition. By a careful study about the structure of Clarke's generalized Jacobian of the metric projector over the epigraph of a convex function, we showed that the semismoothness of solutions to the Moreau-Yosida regularization of a convex function can be obtained via the semismoothness of the metric projector over the epigraph of the convex function. We leave the study on the semismoothness of the metric projector over various interesting closed convex sets such as closed homogeneous cones [4, 7] and the set of correlation matrices [11] as a future research topic.

References

[1] J. F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer-Verlag, New York, 2000.

[2] F.H. Clarke, "On the inverse function theorem," *Pacific Journal of Mathematics* 64 (1976) 97–102.

[3] F. H. Clarke, *Optimization and Nonsmooth Analysis*, John Wiley and Sons, New York, 1983.

[4] C. B. Chua, "Relating homogeneous cones and positive definite cones via T-algebras," *SIAM Journal on Optimization* 14 (2004) 500–506.

[5] B. C. Eaves, "On the basic theorem of complementarity," *Mathematical Programming* 1 (1971) 68–75.

[6] F. Facchinei and J. S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems, Vols I and II*, Springer-Verlag, New York, 2003.

[7] L. Faybusovich, "On Nesterov's approach to semi-infinite programming," *Acta Appl. Math.* 74 (2002) 195–215.

[8] A. Fischer, "Solution of monotone complementarity problems with locally Lipschitzian functions," *Mathematical Programming* 76 (1997) 513–532.

[9] M. Fukushima and L. Qi, "A global and superlinear convergent algorithm for nonsmooth convex minimization," *SIAM Journal on Optimization* 6 (1996) 1106–1120.

[10] M.S. Gowda, "Inverse and implicit function theorems for H-differentiable and semismooth functions," Manuscript, Department of Mathematics and Statistics, University of Maryland, Baltimore (November 2000). To appear in *Optimization Methods and Software.*

[11] N. J. Higham, "Computing the nearest correlation matrix – a problem from finance," *IMA Journal of Numerical Analysis* 22 (2002) 329–343.

[12] J. B. Hiriart-Urruty, C. Lemaréchal, *Convex Analysis and Minimization Algorithms*, Springer Berlin, 1993.

[13] J. Kruskal, "Two convex counterexamples: A discontinuous envelope function and a nondifferentiable nearest-point mapping," *Proceedings of the American Mathematical Society* 23 (1969) 697–703.

[14] B. Kummer, "Newton's method for non-differentiable functions," in *Advances in Mathematical Optimization* (J. Guddat et al., Eds.), pp. 114–125, Akademie-Verlag, Berlin, 1988.

[15] B. Kummer, "Lipschitzian inverse functions, directional derivatives, and applications in $C^{1,1}$-optimization," *Journal of Optimization Theory and Applications* 70 (1991) 559–580.

[16] B. Kummer, "An implicit-function theorem for $C^{0,1}$-equations and parametric $C^{1,1}$-optimization," *Journal of Mathematical Analysis and Applications* 158 (1991) 35–46.

[17] B. Kummer, "Newton's method based on generalized derivatives for nonsmooth functions-convergence analysis," *Lectures Notes in Economics and Mathematics* 182 (1992) 171-194.

[18] A. B. Levy, "Nonsingularity Conditions for multifunctions," *Set-Valued Analysis* 7 (1999) 89–99.

[19] A. B. Levy, "Lipschitzian multifunctions and a Lipschitzian inverse mapping theorem," *Mathematics of Operations Research* 26 (2001) 105–118.

[20] C. Lemaréchal and C. Sagastizábal, "An approach to variable metric methods," in: J. Henry and J. P. Yvon, eds., *Systems Modelling and Pptimization, Lecture Notes in Control and Information Sciences* 197 (Springer-Verlag, Berlin, 1994) pp.144-162.

[21] C. Lemaréchal and C. Sagastizábal, "Practical aspects of the Moreau-Yosida regularization I: theoretical preliminaries," *SIAM Journal on Optimization* 7 (1997) 367–385.

[22] R. Mifflin, "Semismooth and semiconvex functions in constrained optimization," *SIAM Journal on Control and Optimization* 15 (1977) 957–972.

[23] R. Mifflin, L. Qi, and D. Sun, "Properties of the Moreau-Yosida regularization of a piecewise C^2 convex function," *Mathematical Programming* 84 (1999) 269–281.

[24] J. J. Moreau, "Proximite et dualite dans un espace hilbertien," *Bulletin de la Societe Mathematique de France* 93 (1965) 273–299.

[25] J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.

[26] J. S. Pang, D. Sun, and J. Sun, "Semismooth homeomorphisms and strong stability of semidefinite and Lorentz complementarity problems," *Mathematics of Operations Research* 28 (2003) 39–63.

[27] J. P. Penot, Private Communication, 2001.

[28] L. Qi, "Convergence analysis of some algorithms for solving nonsmooth equations," *Mathematics of Operations Research* 18 (1993) 227-244.

[29] L. Qi, "Second-order analysis of the Moreau-Yosida regularization," in *Nonlinear Analysis and Convex Analysis* (Niigata, 1998), 16–25, World Scientific Publishing, River Edge, NJ, 1999.

[30] L. Qi and J. Sun, "A nonsmooth version of Newton's method," *Mathematical Programming* 58 (1993) 353–367.

[31] S. M. Robinson, "First order conditions for general nonlinear optimization," *SIAM Journal on Applied Mathematics* 30 (1976) 597–607.

[32] R. T. Rockafellar, *Convex analysis*, Princeton, New Jersey, 1970.

[33] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Springer-Verlag, New York, 1998.

[34] A. Shapiro, "Directionally nondifferentiable metric projection," *Journal of Optimization Theory and Applications* 81 (1994) 203–204.

[35] A. Shapiro, Private Communication, March 2003.

[36] A. Shapiro, "Sensitivity analysis of generalized equations," *Journal of Mathematical Sciences* 115 (2003) 2554–2565.

[37] D. Sun, "A further result on an implicit function theorem for locally Lipschitz functions," *Operations Research Letters* 28 (2001) 193–198.

[38] D. Sun and J. Sun, "Semismooth matrix valued functions," *Mathematics of Operations Research* 27 (2002) 150–169.

[39] L. Thibault, "Subdifferentials of compactly Lipschitzian vector-valued functions," *Annali di Matematica Pura ed Applicata* 125 (1980) 157-192.

[40] L. Thibault, "On generalized differentials and subdifferentials of Lipschitz vector-valued functions," *Nonlinear Analysis: Theory, Methods, and Applications* 6 (1982) 1037-1053.

[41] K. Yosida, *Functional analysis*, Springer Verlag, Berlin, 1964.

[42] E. H. Zarantonello, "Projections on convex sets in Hilbert space and spectral theory I and II.," In E.H. Zarantonello, editor, Contribution to Nonlinear Functional Analysis, Academic Press (New York 1971) pp. 237–424.