Journal of Convex Analysis
Volume 14 (2007), No. 3, 589–606

Prox-Regularity and Stability of the Proximal Mapping

W. L. Hare*
IRMACS, Simon Fraser University,
Burnaby, BC V5A 1S6, Canada
whare@irmacs.sfu.ca

R. A. Poliquin†
University of Alberta, Department of Mathematical
and Statistical Sciences, Edmonton, AB, Canada
rene.poliquin@ualberta.ca

Received: December 9, 2005
Revised manuscript received: April 17, 2006

Fundamental insights into the properties of a function come from the study of its *Moreau envelopes* and *Proximal point mappings*. In this paper we examine the stability of these two objects under several types of perturbations. In the simplest case, we consider *tilt-perturbations*, i.e. perturbations which correspond to adding a linear term to the objective function. We show that for functions that have single-valued Lipschitz continuous proximal mappings, in particular for prox-regular functions, tilt-perturbations result in stable, i.e. single-valued Lipschitz continuous, proximal point mappings.

In the more complex case, we consider the class of parametrically prox-regular functions. These include most of the functions that arise in the framework of nonlinear programming and its extensions (e.g. convex, lower-$\mathcal{C}^2$, strongly amenable (convexly composite)). New characterizations of prox-regularity are given and more general perturbations along the lines of [12] are studied. We show that under suitable conditions (compatible parameterization, positive coderivative...), the proximal point mappings of the function $f_u(x) = f(x, u)$ depends in a Lipschitz fashion on the parameter u and the prox-parameter r.

Keywords: Prox-regular, proximal mapping, Moreau envelope, stability, convex, convexly composite, strongly amenable, compatible parameterization, perturbations, sensitivity analysis

2000 Mathematics Subject Classification: Primary 49A52, 58C06, 58C20; Secondary 90C30

1. Introduction

Fundamental insights into the properties of a function $f : \mathcal{X} \to \overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$, where $\mathcal{X}$ is a Hilbert space, come from the study of its *Moreau envelopes* $e_r f$, defined for $r > 0$ by

$$e_r f(x) = \inf_w \left\{ f(w) + \frac{r}{2}|w - x|^2 \right\}, \tag{1}$$

and the associated *proximal mappings* $P_r f$, defined by

$$P_r f(x) := \operatorname*{argmin}_w \left\{ f(w) + \frac{r}{2}|w - x|^2 \right\} \tag{2}$$

*Research by this author was supported in part by CNPq Grant No. 150234/2004-0, while visiting at IMPA, Rio de Janeiro, Brazil.
†Research by this author was supported in part by the Natural Sciences and Engineering Research Council of Canada under grant OGP41983.

ISSN 0944-6532 / $ 2.50

The parameter r is referred to as the *prox-parameter*, and the point x as the *prox-center*.

First introduced by Moreau in [17] and [16], the proximal mapping has since become a fundamental tool in convex and nonconvex optimization. Shortly after its introduction, it was shown that a fixed point iteration on the proximal mapping could be used to develop a simple optimization algorithm [13], [14], [23]. Since then, the ideas behind this basic algorithm have been used, studied, and improved repeatedly. For example, past lines of research have included (but are not limited to), the finite convergence of the proximal point algorithm [4], [15], [6], new primal-dual algorithms based on proximal points [8], [5], the replacement norms with Bergman distances in proximal point algorithms [9], [3], and variants of the proximal point algorithm for dealing with nonmonotone or nonconvex functions [18], [10], [7]. On the theoretical side, research involving proximal points has included (but is not limited to), the study of generalized convexity [1], [2], epi- and proto-differentiability [20], generalized Hessians [11], and analysis of the behavior of the proximal mapping itself [19], [20], [22].

The proximal mapping has three components: the prox-center, the prox-parameter, and the function for which we seek the proximal point. In this paper we examine the consequences, on the proximal mapping, of small perturbations in each of these components. Our goal is to show that, under reasonable conditions, perturbations in any of these components will result in single-valued Lipschitz continuous proximal mappings. Using relationships between the proximal mapping and the derivative of the Moreau envelope, we are also able to obtain results for the Moreau envelopes.

It is well known that a convex function has a proximal mapping that is locally single-valued and Lipschitz continuous. The same is true of the much broader class of *prox-regular* functions. Recall that a function $f : \mathcal{X} \to \overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ is prox-regular at $\bar{x}$ for $\bar{v} \in \partial f(\bar{x})$ ($\partial f(x)$ stands for the set of limiting proximal subgradients at the point x) if f is locally lsc at $\bar{x}$ and there exist $\epsilon > 0$ and $r > 0$ such that

$$f(x') \geq f(x) + \langle v, x' - x \rangle - \frac{r}{2}|x' - x|^2 \tag{3}$$

whenever $|x' - \bar{x}| < \epsilon$ and $|x - \bar{x}| < \epsilon$ with $x' \neq x$ and $|f(x) - f(\bar{x})| < \epsilon$ while $|v - \bar{v}| < \epsilon$ with $v \in \partial f(x)$. We say that f is continuously prox-regular at $\bar{x}$ for $\bar{v}$ if, in addition, $f(x)$ is continuous as a function of $(x, v) \in \operatorname{gph} \partial f$ at $(\bar{x}, \bar{v})$.

Prox-regular functions were first introduced in [20] and further studied in [19], [21], [22], [1], [6], [7], [11]. The class of continuously prox-regular functions includes all convex functions, all lower-$\mathcal{C}^2$ functions (a function is lower-$\mathcal{C}^2$ if the function plus a multiple of the norm square is locally convex) and all strongly amenable functions. In finite dimensions, strong amenability of f at $\bar{x}$ refers to the existence of a representation $f = g \circ F$ in a neighborhood of $\bar{x}$ for a mapping $F : \mathbb{R}^n \to \mathbb{R}^m$ that is C^2 and a proper lsc convex function $g : \mathbb{R}^m \to \overline{\mathbb{R}}$ satisfying a constraint qualification at $\bar{x}$; see [1] for a discussion of strong amenability in infinite dimensions.

In [20], and later in [1] for functions defined on a Hilbert space, it was shown that if the function f is prox-bounded and prox-regular at $\bar{x}$ for the subgradient $\bar{v}$, then in a neighborhood of $\bar{x} + (1/r)\bar{v}$ the proximal mapping $P_r f$, for r sufficiently large, is single-valued and Lipschitz continuous near $\bar{x} + (1/r)\bar{v}$. Recall that a function f is *prox-bounded* if there exists $r > 0$ and a point $\bar{x}$ such that $e_r f(\bar{x}) > -\infty$. The infimum of the set of all such r, denoted by r_{pb}, is the *threshold of prox-boundedness*.

In addition, it was shown in [20] and [1] that for prox-regular functions there exists some $\epsilon > 0$ with

$$P_r f = (I + \frac{1}{r}T)^{-1} \tag{4}$$

in a neighborhood of $\bar{x}+(1/r)\bar{v}$ where T is the *f-attentive ϵ-localization*[1] of the subgradient mapping ∂f around $(\bar{x}, \bar{v})$. Recall that the f-attentive ϵ-localization of the subgradient mapping ∂f around $(\bar{x}, \bar{v})$ is the mapping T defined by

$$T(x) = \begin{cases} \{v \in \partial f(x) | |v - \bar{v}| < \epsilon\} & \text{if } |x - \bar{x}| < \epsilon \text{ and } |f(x) - f(\bar{x})| < \epsilon, \\ \emptyset & \text{otherwise.} \end{cases} \tag{5}$$

In this paper we add to the list of equivalent characterizations of prox-regular functions by showing that (4) is both a necessary and sufficient condition for prox-regularity. This result is presented in Theorem 2.3. In that theorem we also provide two new characterizations of prox-regularity. In particular, we show, in finite dimensions, that $(I + (1/r)T)^{-1}$ single-valued (not necessarily equal to $P_r f$ nor Lipschitz continuous) guarantees prox-regularity. In Theorem 2.4 we list some of the effects on the Moreau envelopes of having stable (single-valued Lipschitz continuous) proximal mappings. Many of the equivalent characterizations summarized in Theorem 2.4 first appeared in [1]. In this paper simplified proofs are given of the known results and we add to the list of equivalent characterizations by showing that a function has single-valued Lipschitz continuous proximal mappings if and only if the Moreau envelopes are lower-$\mathcal{C}^2$.

In Section 3 we examine the effect of a *tilt-perturbation* of the objective function on the proximal point mapping. Tilt perturbations correspond to adding a linear term to the objective; in [21] and [12] the behavior of a minimizing point when an objective function is tilted by adding a linear term is studied. We show that for functions that have single-valued Lipschitz continuous proximal mappings, in particular prox-regular functions, tilt-perturbations of the objective function result in stable i.e. Lipschitz continuous behaviour in the proximal point mapping. In our Theorem 3.1 we show that if a function f has a stable proximal mapping of parameter r, then the functions $f_u(x) := f(x) - \langle u, x\rangle$ have a stable proximal mapping of parameter r. With the addition of prox-regularity, in Theorem 3.2, we also obtain that the proximal mappings are stable as a function of r also.

Finally in Section 4, more general perturbations are studied along the lines of [12]. We show, in Theorem 4.6, that under suitable conditions (compatible parameterization, positive coderivative...), the proximal mappings of the functions $f_u(x) = f(x, u)$ depend in a Lipschitz fashion on the parameter u and the prox-parameter r.

2. Prox-regularity and the Proximal Mapping

We begin by recalling some useful (known) facts regarding the proximal mapping of prox-bounded functions. Except when otherwise stated, we will follow the notation of [24]. In particular, ∂ represents the regular subdifferential (i.e. the set of limiting proximal subgradients) and ∂^∞ represents the horizon subdifferential (see [24, Def. 8.3]).

[1]In [20] and [1] it is shown that a function is prox-regular if and only if there is an *f-attentive ϵ-localization* T of the subgradient mapping ∂f that is hypomonotone i.e. for some $r > 0$, $rI + T$ is monotone.

Proposition 2.1 (Proximal Mapping). *A proper lsc function* $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ *is prox-bounded with threshold* r_{pb} *if and only if* $f(\cdot) + (r/2)|\cdot - x|^2$ *is bounded below for any point* x *and any value* $r > r_{pb}$ [24, Ex. 1.24].

If a proper lsc function f *is prox-bounded with threshold* r_{pb}*, then for any* $r > r_{pb}$ *the following are true:*

(a) $e_r f$ *depends continuously on* (r, x)*,* [24, Thm. 1.25]*;*
(b) $P_r f(x)$ *is nonempty and compact (for all* x*),* [24, Thm. 1.25]*;*
(c) $\partial e_r f \subseteq r(I - P_r f)$ *and is locally bounded,* [24, Ex. 10.32]*;*
(d) $\partial(-e_r f) = r(\operatorname{conv} P_r f - I)$*,* [24, Ex. 10.32]*;*
(e) $-e_r f$ *is lower-*$\mathcal{C}^2$*,* [24, Ex. 10.32]*;*
(f) $P_r f$ *is monotone,* [24, Prop. 12.19]*;*
(g) $rx - rP_r f(x) \in \partial f(P_r f(x))$*, (first order optimality condition);*
(h) $P_r f \subseteq (I + (1/r)\partial f)^{-1}$*, (by (*$g$*)); and*
(i) *if* $w_k \in P_r f(x_k)$ *with* $x_k \to \bar{x}$ *then the sequence* $\{w_k\}$ *is bounded and all its accumulation points lie in* $P_r f(\bar{x})$ [24, Thm. 1.25].

Remark. It is an easy exercise to confirm that the preliminary statement (on thresholds of prox-boundedness), as well as parts (a), (f), (g), and (h) of Prop. 2.1, also hold true in Hilbert spaces. (The proofs laid out in [24] make no use of any of the special properties of $\mathbb{R}^n$ over a Hilbert space.)

We next explore the effects of adding a tilt parameter on the Moreau envelopes and the proximal mappings.

Lemma 2.2 (Tilting proximal points). *For a proper lsc function* $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ *and a vector* $v \in \mathbb{R}^n$ *let* f_v *be the tilted function given by* $f_v(x) = f(x) - \langle v, x\rangle$*. Then for all* x*,*

(a) $e_r f_v(x) = e_r f(x + \frac{1}{r}v) - \langle x, v\rangle - \frac{1}{2r}|v|^2$,
(b) $P_r f_v(x) = P_r f(x + \frac{1}{r}v)$*. (holds in a Hilbert space)*

Therefore, if the function f_v *has a well behaved proximal mapping near* $\bar{x}$*, then* f *has a well behaved proximal mapping near* $\bar{x} + (1/r)v$*.*

Proof. Without loss of generality we may assume that f is prox-bounded and that r is greater than the threshold of prox-boundedness. Indeed if f is not prox-bounded, or if r is below the threshold of prox-boundedness, then $e_r f \equiv -\infty$ and $P_r f = \emptyset$. As tilting f by a linear term does not alter the threshold of prox-boundedness, we also have $e_r f_v \equiv -\infty$ and $P_r f_v = \emptyset$, so the formulae hold.

If r is greater than the threshold of prox-boundedness then (b) can be derived as follows,

$$\begin{aligned} P_r f_v(x) &= \operatorname*{argmin}_w \left\{ f(w) - \langle v, w\rangle + \frac{r}{2}|w - x|^2 \right\} \\ &= \operatorname*{argmin}_w \left\{ f(w) - \langle v, w\rangle - \langle rx, w\rangle + \frac{r}{2}|w|^2 \right\} \\ &= \operatorname*{argmin}_w \left\{ f(w) - r\langle x + \frac{1}{r}v, w\rangle + \frac{r}{2}|w|^2 + \frac{r}{2}|x + \frac{1}{r}v|^2 \right\} \\ &= \operatorname*{argmin}_w \left\{ f(w) + \frac{r}{2}|w - (x + \frac{1}{r}v)|^2 \right\} = P_r f(x + \frac{1}{r}v). \end{aligned}$$

For (a) first note that $P_r f_v(x)$ is nonempty (Prop. 2.1(b)). Let $w \in P_r f_v(x) = P_r f(x + (1/r)v)$. We have

$$\begin{aligned}
e_r f_v(x) &= f_v(w) + \frac{r}{2}|x - w|^2 \\
&= f(w) - \langle v, w\rangle + \frac{r}{2}|x + \frac{1}{r}v - w - \frac{1}{r}v|^2 \\
&= f(w) - \langle v, w\rangle + \frac{r}{2}|x + \frac{1}{r}v - w|^2 - \langle x + \frac{1}{r}v - w, v\rangle + \frac{1}{2r}|v|^2 \\
&= e_r f(x + \frac{1}{r}v) - \langle v, w\rangle - \langle x, v\rangle - \frac{1}{r}|v|^2 + \langle v, w\rangle + \frac{1}{2r}|v|^2 \\
&= e_r f(x + \frac{1}{r}v) - \langle x, v\rangle - \frac{1}{2r}|v|^2.
\end{aligned}$$

□

Remark. By using the previous lemma, we can always assume that the point of interest is a critical point of the function. To elaborate, given a function f which is prox-regular at $\bar{x}$ for $\bar{v}$, we can create the function $f_{\bar{v}}(\cdot) = f(\cdot) - \langle \bar{v}, \cdot\rangle$. Any result which holds for the proximal mapping of $f_{\bar{v}}$ at $\bar{x}$ must then hold for the original function at $\bar{x} + (1/r)\bar{v}$.

As mentioned in the introduction, it is well-known that a prox-regular function has single-valued Lipschitz continuous proximal mappings and that the proximal mapping or parameter r is given by $(I + (1/r)T)^{-1}$. In the next theorem we show that this formula is necessary and sufficient for prox-regularity; two additional (new) characterizations of prox-regularity are also given. In particular, we show, in finite dimensions, that $(I + (1/r)T)^{-1}$ single-valued (not necessarily equal to $P_r f$ nor Lipschitz continuous) guarantees prox-regularity.

Theorem 2.3. *Let $\mathcal{X}$ be a Hilbert space and $f : \mathcal{X} \to \overline{\mathbb{R}}$ be a proper lsc prox-bounded function which is finite at the point $\bar{x}$, then the following are equivalent:*

- (a) *f is prox-regular at $\bar{x}$ for $\bar{v} \in \partial f(\bar{x})$.*
- (b) *There exists an f-attentive ϵ-localization T of ∂f around $(\bar{x}, \bar{v})$ such that for all r sufficiently large, $P_r f = (I + (1/r)T)^{-1}$ near $\bar{x} + (1/r)\bar{v}$ with $P_r f(\bar{x} + (1/r)\bar{v}) = \bar{x}$.*
- (c) *There exists an f-attentive ϵ-localization T of ∂f around $(\bar{x}, \bar{v})$ such that for all r sufficiently large, $r(I - P_r f) = ((1/r)I + T^{-1})^{-1}$ near $\bar{x} + (1/r)\bar{v}$ with $P_r f(\bar{x} + (1/r)\bar{v}) = \bar{x}$.*

Any of the previous conditions imply

- (d) *There exists an f-attentive ϵ-localization T of ∂f around $(\bar{x}, \bar{v})$ such that for all r sufficiently large $(I + (1/r)T)^{-1}$ is single-valued near $\bar{x} + (1/r)\bar{v}$ with $P_r f(\bar{x} + (1/r)\bar{v}) = \bar{x}$.*

If the space $\mathcal{X}$ is finite dimensional then (d) is equivalent to the previous conditions.

Remark. Under (a), (b), or (c), the proximal mapping $P_r f$ is actually Lipschitz continuous near $\bar{x} + (1/r)\bar{v}$.

Proof. Clearly we can assume without loss of generality that $\bar{x} = 0$ and that $f(\bar{x}) = 0$. We will show that $(a) \Rightarrow (b) \Leftrightarrow (c) \Rightarrow (a)$, $(a) \Rightarrow (d)$ and that $(d) \Rightarrow (b)$ in a finite dimensional space.

For any set-valued mapping T, $((1/r)I + T^{-1})^{-1} = r[I - (I + (1/r)T)^{-1}]$ (see [20, Lem. 4.5] for details). Therefore $r(I - P_r f) = r[I - (I + (1/r)T)^{-1}]$ if and only if $-P_r f = -(I + (1/r)T)^{-1}$, which shows that (b) and (c) are equivalent.

Condition (a) implies (b) by [20, Thm. 4.4] (or [24, Prop. 13.37]) for $\bar{v} = 0$ in finite dimensions and [1, Prop. 4.4] for arbitrary $\bar{v}$ in a Hilbert space. Moreover, under (a), $P_r f$ is single-valued Lipschitz continuous near $\bar{x} + (1/r)\bar{v}$; this shows that (a) implies (d).

To see that (b) implies (a), recall that the proximal mapping $P_r f$ is always monotone (Prop. 2.1 part (f) which also works in a Hilbert space)). Thus (b) implies that $(I + (1/r)T)^{-1}$ is monotone (once r is sufficiently large). This in turn implies that $rI + T$ is monotone (see for example [24, Ex. 12.4], noting that the proof also works in a Hilbert space). Thus f has an f-attentive ϵ-localization T around $(\bar{x}, \bar{v})$ such that $T + rI$ is monotone. This implies that (a) holds by [20, Thm. 3.2] (or [24, Thm. 13.36]) for $\bar{v} = 0$ and [1, Thm. 3.4] for arbitrary $\bar{v}$ in a Hilbert space.

Finally, if (d) holds then $f_{\bar{v}}(x) := f(x) - \langle \bar{v}, x\rangle > -(r/2)|x|^2$ for all $x \neq \bar{x}$ (from the fact that $P_r f(\bar{x} + (1/r)\bar{v}) = \bar{x}$). When working in a finite dimensional space, we can then use [20, Prop. 4.2 (c)] to conclude that for any $\epsilon' > 0$ there exists a neighborhood $\mathcal{O}$ of $\bar{x}$ (which depends on ϵ') such that if $x' \in P_r f_{\bar{v}}(x)$ with $x \in \mathcal{O}$ then $|x'| < \epsilon'$, $|f_{\bar{v}}(x')| < \epsilon'$ and $|r(x - x')| < \epsilon'$ (recall $f(\bar{x}) = 0$).

Now pick $\epsilon' > 0$ small enough so that $\epsilon' + |\bar{v}|\epsilon' < \epsilon$. We claim that

$$P_r f(x + (1/r)\bar{v}) \subseteq (I + (1/r)T)^{-1}(x + (1/r)\bar{v})$$

for all $x \in \mathcal{O}$. This inclusion with the fact that $P_r f \neq \emptyset$ (by prox-boundedness and finite dimensionality) implies (b). To establish the inclusion, take $x' \in P_r f(x + (1/r)\bar{v}) = P_r f_{\bar{v}}(x)$ (by Lemma 2.2) and $x \in \mathcal{O}$. We must show that $r(x - x') + \bar{v} \in T(x')$. Since $|x'| < \epsilon' < \epsilon$, $|f(x')| < \epsilon' + |\bar{v}|\epsilon' < \epsilon$, $|r(x - x') + \bar{v} - \bar{v}| = |r(x - x')| < \epsilon' < \epsilon$ and $r(x - x') \in \partial f_{\bar{v}}(x')$ (Prop. 2.1 (g)) i.e. $r(x - x') + \bar{v} \in \partial f(x')$; this shows that $r(x - x') + \bar{v} \in T(x')$. □

We conclude this section by listing some of the consequences of having stable proximal mappings on the Moreau envelopes. The fact that single-valued Lipschitz continuous proximal mappings, condition $(A1)$ in the following theorem, is equivalent to having $\mathcal{C}^{1+}$ (Lipschitz continuous derivative) envelopes $(A2)$ along with the formula in $(B3)$, was first established in [1, Prop. 5.3]; here we give a different proof of these results and show that they are equivalent, in finite dimensions, to having lower-$\mathcal{C}^2$ envelopes (which is new to this paper). The equivalence between $(B2)$ and $(B4)$ along with formula $(B3)$ can be found in [1, Prop. 5.1], the fact that this is also equivalent to $(B1)$, in finite dimensions, is new.

Theorem 2.4. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a proper lsc prox-bounded function which is finite at the point $\bar{x}$ then the following are equivalent and are implied by prox-regularity of f at $\bar{x}$ for $\bar{v}$.*

$(A1)$ *For all r sufficiently large $P_r f$ is single-valued and Lipschitz continuous near $\bar{x} + (1/r)\bar{v}$.*

$(A2)$ *For all r sufficiently large $e_r f$ is $\mathcal{C}^{1+}$ near $\bar{x} + (1/r)\bar{v}$.*

$(A3)$ *For all r sufficiently large $e_r f$ is lower-$\mathcal{C}^2$ near $\bar{x} + (1/r)\bar{v}$.*

Moreover, each of the above conditions implies the following equivalent conditions:

($B1$) *For all r sufficiently large $P_r f$ is single-valued near $\bar{x} + (1/r)\bar{v}$.*

($B2$) *For all r sufficiently large $e_r f$ is $\mathcal{C}^1$ near $\bar{x} + (1/r)\bar{v}$.*

($B3$) *For all r sufficiently large $\nabla e_r f = r(I - P_r f)$ near $\bar{x} + (1/r)\bar{v}$.*

($B4$) *For all r sufficiently large $P_r f$ is single-valued continuous near $\bar{x} + (1/r)\bar{v}$.*

Proof. We split the proof into two parts and take a bottom up approach.

Part I: $(A1) \Rightarrow (B1) \Rightarrow (B2) \Rightarrow (B3) \Rightarrow (B4) \Rightarrow (B1)$

That condition ($A1$) implies condition ($B1$) is obvious. To see that ($B1$) implies ($B2$), first note that when $P_r f$ is single-valued near $\bar{x}$ then, according to Prop. 2.1 part (c), $\partial e_r f$ is a singleton, when nonempty, near $\bar{x}$. Furthermore, by Prop. 2.1 parts (a) and (c), $e_r f$ is continuous and $\partial e_r f$ is locally bounded. Therefore by [24, Cor. 9.19] (which states that a function is $\mathcal{C}^1$ if and only if it is lsc and its subdifferential is locally bounded and single-valued when nonempty) we conclude that $e_r f \in \mathcal{C}^1$.

That ($B2$) implies ($B3$) follows from Prop. 2.1 part (d). That either ($B3$) or ($B4$) implies condition ($B1$) is obvious, while ($B2$) with ($B3$) clearly imply ($B4$).

Part II: Prox-regularity of $\bar{x}$ at $\bar{v} \Rightarrow (A1) \Leftrightarrow (A2) \Leftrightarrow (A3)$

Prox-regularity of $\bar{x}$ at $\bar{v}$ implies ($A1$) by [20, Thm. 4.4] (or [24, Prop. 13.37]) for $\bar{v} = 0$ and [1, Prop. 5.3] for arbitrary $\bar{v}$. To see the equivalence of ($A1$) and ($A2$), we must first note that either ($A1$) or ($A2$) imply ($B3$). Indeed, ($A1$) implies ($B1$), and ($A2$) implies ($B2$) trivially, while Part I (above) shows both of these conditions are equivalent to ($B3$). Having ($B3$) we now note that $P_r f$ is single-valued Lipschitz continuous if and only if $r(I - P_r f)$ is single-valued Lipschitz continuous if and only if $e_r f$ is $\mathcal{C}^{1+}$ (by ($B3$)).

To see that ($A2$) are ($A3$) equivalent, recall that a function g is $\mathcal{C}^{1+}$ if and only if g is both upper-$\mathcal{C}^2$ (i.e. $-g$ is lower-$\mathcal{C}^2$) and lower-$\mathcal{C}^2$ [24, Prop. 13.34]. The Moreau envelopes are always upper-$\mathcal{C}^2$ (Prop. 2.1 part (e)). □

Remark. On page 618 of [24] there is an example that shows that condition ($A1$) does not imply prox-regularity. As of yet, it is unknown if the (A) and (B) conditions are in fact equivalent.

3. Tilt-Stability of Proximal Mapping

In this section we examine the stability of the proximal mapping under one of the simplest perturbations of the objective: tilt-perturbations. We first show that a prox-bounded function that satisfies any of the equivalent (A) conditions of Theorem 2.3 has a stable proximal mapping under tilt-perturbations. In particular this is true of a prox-regular function.

Theorem 3.1. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a proper lsc prox-bounded function with threshold r_{pb} which is finite at the point $\bar{x}$. Let $f_u(x) = f(x) - \langle u, x\rangle$. Fix a prox-parameter $r > r_{pb}$, prox-center $\bar{x}$ and error parameter $\bar{u}$. If $P_r f$ is single-valued and Lipschitz continuous near $\bar{x} + (1/r)\bar{u}$, then there exist $K > 0$, neighborhoods $\mathcal{O}$ and U of $\bar{x}$ and $\bar{u}$ such that*

$$|P_r f_u(x) - P_r f_{u'}(x')| \leq K\big(|x - x'| + |u' - u|\big)$$

and

$$|\nabla e_r f_u(x) - \nabla e_r f_{u'}(x')| \leq K\big(|x - x'| + |u' - u|\big),$$

for all $x, x' \in \mathcal{O}$, u, u' *in* U.

In particular, whenever f *is prox-regular, tilt-perturbations result in Lipschitz perturbations of the proximal mapping.*

Proof. There exists $\varepsilon > 0$ and $k > 0$ such that if $z, z' \in B_\varepsilon(\bar{x} + (1/r)\bar{u})$ (the open ball of radius ε around $(\bar{x} + (1/r)\bar{u})$) then

$$|P_r f(z) - P_r f(z')| \leq k|z - z'|.$$

Consider now any $x, x' \in B_{\varepsilon/2}(\bar{x})$ and $u, u' \in B_{r\varepsilon/2}(\bar{u})$. Clearly $x + (1/r)u, x' + (1/r)u' \in B_\varepsilon(\bar{x} + (1/r)\bar{u})$. Hence

$$\begin{aligned}
|P_r f(x + \frac{1}{r}u) - P_r f(x' + \frac{1}{r}u')| &\leq k|x + \frac{1}{r}u - x' - \frac{1}{r}u'| \\
&\leq k|x - x'| + \frac{k}{r}|u - u'| \\
&\leq K_1\big(|x - x'| + |u - u'|\big)
\end{aligned}$$

where $K_1 = \max\{k, k/r\}$. By Lemma 2.2 $P_r f(x + (1/r)u) = P_r f_u(x)$ and $P_r f(x + (1/r)u') = P_r f_{u'}(x)$ and therefore

$$|P_r f_u(x) - P_r f_{u'}(x')| \leq K_1\big(|x - x'| + |u - u'|\big)$$

To see the statement regarding the Moreau envelopes, recall that $\nabla e_r f = r(I - P_r f)$ under condition $(B1)$ of Theorem 2.3. Hence

$$\begin{aligned}
|\nabla e_r f(x + \frac{1}{r}u) - \nabla e_r f(x' + \frac{1}{r}u')| &= \big|r\big(x + \frac{1}{r}u) - P_r f(x' + \frac{1}{r}u)\big) \\
&\quad - r\big(x' + \frac{1}{r}u') - P_r f(x' + \frac{1}{r}u')\big)\big| \\
&\leq r|x - x'| + |u - u'| + K_1\big(|x - x'| + |u - u'|\big) \\
&\leq K_2\big(|x - x'| + |u - u'|\big)
\end{aligned}$$

where $K_2 = K_1 + \max\{r, 1\}$. From Lemma 2.2 we have $\nabla e_r f(x + (1/r)u) = \nabla e_r f_u(x) + u$ and $\nabla e_r f(x' + (1/r)u') = \nabla e_r f_{u'}(x') + u'$. Hence

$$|\nabla e_r f_u(x) + u - \nabla e_r f_{u'}(x') - u'| \leq K_2\big(|x - x'| + |u - u'|\big)$$

which implies that

$$|\nabla e_r f_u(x) - \nabla e_r f_{u'}(x')| \leq (K_2 + 1)\big(|x - x'| + |u - u'|\big).$$

Letting $K = K_2 + 1$ completes the proof. □

When the function is prox-regular then the proximal mapping of parameter r is single-valued and Lipschitz continuous for all $r > r_{pb}$. In this case, we get added stability in that the prox-parameter r can vary (within a given range) without affecting the Lipschitz behaviour of the proximal point mapping.

Theorem 3.2. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a proper lsc prox-bounded function with threshold r_{pb} that is prox-regular at $\bar{x}$ for $0 \in \partial f(\bar{x})$. Let $f_u(x) = f(x) - \langle u, x\rangle$. Fix $\bar{r} > r_{pb}$. There exist $K > 0$, $\rho > 0$, neighborhoods $\mathcal{O}$ and U of $\bar{x}$ and 0 such that*

$$|P_r f_u(x) - P_r f_{u'}(x')| \leq K\big(|x - x'| + |u' - u|\big)$$

and

$$|\nabla e_r f_u(x) - \nabla e_r f_{u'}(x')| \leq K\big(|x - x'| + |u' - u|\big),$$

for all $x, x' \in \mathcal{O}$, u, u' in U, with $|r - \bar{r}| < \rho$.

Proof. According to [20, Thm. 4.4] for every $r > r_{pb}$ the mapping $P_r f$ is single-valued and Lipschitz continuous on a neighborhood of $\bar{x}$ with Lipschitz constants $1/(1 - \big(r_{pb}/r)\big)$. These neighborhoods form a nested family of neighborhoods that decrease as r increases. Note that the Lipschitz constants are bounded in any neighborhood of $\bar{r}$. Therefore there exists $\varepsilon > 0$ and $k > 0$ such that if $z, z' \in B_\varepsilon(\bar{x})$ (the open ball of radius ε around $(\bar{x})$) and $|r - \bar{r}| < \varepsilon$ then

$$|P_r f(z) - P_r f(z')| \leq k|z - z'|.$$

If $|x - \bar{x}| < \varepsilon/2$, $|u| < \varepsilon r_{pb}/2$ with $r > r_{pb}$ then

$$\begin{aligned} |x + \frac{1}{r}u - \bar{x}| &\leq |x - \bar{x}| + \frac{1}{r}|u| \\ &\leq |x - \bar{x}| + \frac{1}{r_{pb}}|u| \\ &< \varepsilon/2 + \varepsilon/2 = \varepsilon \end{aligned}$$

Hence if $x, x' \in B_{\varepsilon/2}(\bar{x})$, $|u| < \varepsilon r_{pb}/2, |u'| < \varepsilon r_{pb}/2$, $r > r_{pb}$ with $|r - \bar{r}| < \varepsilon$ then

$$\begin{aligned} |P_r f(x + \frac{1}{r}u) - P_r f(x' + \frac{1}{r}u')| &\leq k|x + \frac{1}{r}u - x' - \frac{1}{r}u'| \\ &\leq k|x - x'| + \frac{k}{r}|u - u'| \\ &\leq k|x - x'| + \frac{k}{r_{pb}}|u - u'| \\ &\leq K_1\big(|x - x'| + |u' - u|\big) \end{aligned}$$

for $K_1 = \max\{k, \frac{k}{r_{pb}}\}$. By Lemma 2.2, we have $P_r f(x + (1/r)u) = P_r f_u(x)$ and $P_r f(x' + (1/r')u') = P_r f_{u'}(x')$. Hence

$$|P_r f_u(x) - P_r f_{u'}(x')| \leq K_1\big(|x - x'| + |u - u'|\big)$$

The statement regarding the Moreau envelopes can now be derived in the same manner as in Theorem 3.1. □

4. Full Stability of Proximal Mapping

In order to further examine the stability of proximal mappings, we will make use of the stability theory developed in [12]. A simplified version of the main theorem of [12] is reproduced in Theorem 4.2 below. Before stating it we need to recall some definitions and set up some notation.

For the function

$$\begin{array}{rcl} h: \mathbb{R}^n \times \mathbb{R}^s & \to & \overline{\mathbb{R}} \\ (x,p) & \to & h(x,p) \end{array}$$

(where $p \in \mathbb{R}^s$ represents the perturbation parameter) that is finite at the point (x_0, p_0), then for parameters $v \in \mathbb{R}^n$ and $\delta \geq 0$ define the function m_δ^h and mapping M_δ^h by

$$m_\delta^h(p,v) := \inf_{|x-x_0|\leq\delta} \{h(x,p) - \langle v, x\rangle\} \tag{6}$$

$$M_\delta^h(p,v) := \operatorname*{arg\,min}_{|x-x_0|\leq\delta} \{h(x,p) - \langle v, x\rangle\}. \tag{7}$$

A point x_0 is a *fully stable* (as defined in [12]) locally optimal solution for h relative to p_0 and v_0 if x_0 is a locally optimal solution to $\inf_x\{h(x,p_0) - \langle v_0, x\rangle\}$ and for some $\delta > 0$ and a neighborhood of (p_0, v_0), the mappings M_δ^h and m_δ^h are (single-valued) Lipschitz continuous on this neighborhood with $M_\delta^h(p_0, v_0) = x_0$; (recall that $\bar{x}$ is a locally optimal solution to $\inf_x f(x)$ if for some $\delta > 0$, $f(\bar{x}) \leq f(x)$ for all x with $|x - \bar{x}| \leq \delta$).

Our goal is to apply the main result of [12] to the setting

$$h(x,u,r) := f(x,u) + \frac{r}{2}|x-\bar{x}|^2.$$

i.e. $p = (u, r)$, where $\bar{x}$ is some fixed point in $\mathbb{R}^n$.

In the main result of [12] the notion of *compatible parameterization* is used and a "positive" coderivative condition is employed. Recall that the coderivative of a set-valued mapping S from $\mathbb{R}^n$ to $\mathbb{R}^d$ at z for $w \in S(z)$, denoted $D^*S(z|w)$, is the mapping from $\mathbb{R}^d$ to $\mathbb{R}^n$ defined by $D^*S(z|w)(w') = \{z' | (z', -w') \in N_{gphS}(z,w)\}$ where $N_{gphS}(z,w)$ denotes the normal cone to the graph of S at (z,w).

Recalling, $\partial_x h(x,p) = \partial h_p(x)$ where $h_p(x) = h(x,p)$, we define compatible parameterization:

Definition 4.1 (Compatible Parameterization). A lsc function $h(x,p) : \mathbb{R}^n \times \mathbb{R}^s \to \overline{\mathbb{R}}$ is *prox-regular in* x at the point x_0 with *compatible parameterization in* p at p_0 for the subgradient $v_0 \in \partial_x h(x_0, p_0)$ with parameters ρ and ε if

$$h(x',p) \geq h(x,p) + \langle v, x'-x\rangle - \frac{\rho}{2}|x'-x|^2 \text{ for all } x' \text{ with } |x'-x_0| < \varepsilon, \tag{8}$$

whenever $|x-x_0| < \varepsilon$, $|p-p_0| < \varepsilon$, $v \in \partial_x h(x,p)$ with $|v-v_0| < \varepsilon$, and $h(x,p) \leq h(x_0,p_0)+\varepsilon$. It is continuously prox-regular in x at x_0 for v_0 with compatible parameterization by p at p_0 if, in addition, $h(x,p)$ is continuous as a function of $(x,p,v) \in \operatorname{gph} \partial_x h$ at (x_0,p_0,v_0).

We are now in a position to state a version of [12, Thm. 2.3] which has been simplified for application in our setting.

Theorem 4.2 (Parts of) **[12, Thm. 2.3].** *Let $h(x,p) : \mathbb{R}^n \times \mathbb{R}^s \to \overline{\mathbb{R}}$ be finite at the point (x_0, p_0). Suppose the constraint qualification*

$$(0,y) \in \partial^\infty h(x_0,p_0) \Rightarrow y = 0 \tag{9}$$

holds and that h is continuously prox-regular in x at x_0 with compatible parameterization by p at p_0 for $0 \in \partial_x h(x_0,p_0)$. Then

$$x_0 \textit{ is a fully stable locally optimal solution for } h \textit{ relative to } p_0 \textit{ and } 0$$

if and only if

$$(0,p') \in D^*(\partial_x h)(x_0,p_0|0)(0) \Rightarrow p' = 0, \textit{ and}$$
$$(x',p') \in D^*(\partial_x h)(x_0,p_0|0)(v') \textit{ with } v' \neq 0 \Rightarrow \langle x', v'\rangle > 0.$$

In this case, then for $\delta > 0$ sufficiently small and (p,v) near $(p_0,0)$ one has

$$M_\delta^h(p,v) = \operatorname*{argmin}_x \{h(x,p) - \langle v,x\rangle\} \bigcap \{x : |x-x_0| < \delta\}.$$

In order to make use of Theorem 4.2 we must show that $h(x,u,r) = f(x,u) + \frac{r}{2}|x-\bar{x}|^2$ satisfies the framework of this theorem at the point $(x_0,p_0) = (x_0,u_0,r_0)$. To do this we must answer three questions:

i. When is $h(x,u,r)$ continuously prox-regular in x at the point x_0 with compatible parameterization by (u,r) at (u_0,r_0) for the subgradient $0 \in \partial_x h(x_0,u_0,r_0)$?
ii. When does h satisfy the constraint qualification (9) at (x_0,p_0) for $p_0 = (u_0,r_0)$?
iii. When does h satisfy the two conditions placed on the coderivative by Theorem 4.2 at (x_0,u_0,r_0)?

We approach these three questions in order, presenting our answers in Lemmas 4.3, 4.4 and 4.5.

Lemma 4.3. *Suppose the proper lsc function $f(x,u) : \mathbb{R}^n \times \mathbb{R}^d \to \overline{\mathbb{R}}$ is (continuously) prox-regular at x_0 with compatible parameterization by u at u_0 for $\bar{v} \in \partial_x f(x_0,u_0)$. Fix $\bar{x} \in \mathbb{R}^n$, then for any $r_0 > 0$ the function*

$$h(x,u,r) := f(x,u) + \frac{r}{2}|x-\bar{x}|^2$$

is (continuously) prox-regular in x at x_0 for $v_0 = \bar{v} + r_0(x_0 - \bar{x})$ with compatible parameterization by (u,r) at (u_0,r_0).

Proof. First we note that $\partial_x h(x,u,r) = \partial_x f(x,u) + r(x-\bar{x})$, thus $v_0 \in \partial_x h(x_0,u_0,r_0)$.

Since f is prox-regular at x_0 with compatible parameterization by u at u_0 for $\bar{v}$ we know that there exists parameters ρ and ε such that

$$f(x',u) \geq f(x,u) + \langle v, x'-x\rangle - \frac{\rho}{2}|x'-x|^2 \text{ when } \begin{cases} |x'-x_0| < \varepsilon, |x-x_0| < \varepsilon, \\ |u-u_0| < \varepsilon, \\ v \in \partial_x f(x,u), |v-\bar{v}| < \varepsilon, \\ f(x,u) \leq f(x_0,u_0) + \varepsilon. \end{cases} \tag{10}$$

What we seek to show is the existence of a $\varepsilon^* > 0$ such that

$$h(x', u, r) \geq h(x, u, r) + \langle v^*, x' - x\rangle - \frac{\rho}{2}|x' - x|^2 \text{ when } \begin{cases} |x' - x_0| < \varepsilon^*, |x - x_0| < \varepsilon^*, \\ |u - u_0| < \varepsilon^*, |r - r_0| < \varepsilon^*, \\ v^* \in \partial_x h(x, u, r), \\ |v^* - v_0| < \varepsilon^*, \\ h(x, u, r) \leq h(x_0, u_0, r_0) + \varepsilon^*. \end{cases} \tag{11}$$

From (10)

$$\begin{aligned} &f(x', u) + \frac{r}{2}|x' - \bar{x}|^2 - \frac{r}{2}|x' - \bar{x}|^2 \\ &\geq f(x, u) + \frac{r}{2}|x - \bar{x}|^2 + \langle v + r(x - \bar{x}), x' - x\rangle - \frac{\rho}{2}|x' - x|^2 - \frac{r}{2}|x - \bar{x}|^2 - r\langle x - \bar{x}, x' - x\rangle \end{aligned}$$

will hold whenever $|x' - x_0| < \varepsilon$, $|x - x_0| < \varepsilon$, $|u - u_0| < \varepsilon$, $v \in \partial_x f(x, u)$ with $|v - \bar{v}| < \varepsilon$, and $f(x, u) \leq f(x_0, u_0) + \varepsilon$. In terms of h this leads to

$$\begin{aligned} h(x', u, r) &\geq h(x, u, r) + \langle v + r(x - \bar{x}), x' - x\rangle - \frac{\rho}{2}|x' - x|^2 + \frac{r}{2}|x' - \bar{x}|^2 - \frac{r}{2}|x - \bar{x}|^2 \\ &\qquad - r\langle x - \bar{x}, x' - x\rangle \\ &= h(x, u, r) + \langle v + r(x - \bar{x}), x' - x\rangle - \frac{\rho}{2}|x' - x|^2 + \frac{r}{2}|x' - \bar{x}|^2 \\ &\qquad - r\langle x - \bar{x}, x' - \bar{x}\rangle + r\langle x - \bar{x}, x - \bar{x}\rangle - \frac{r}{2}|x - \bar{x}|^2 \\ &= h(x, u, r) + \langle v + r(x - \bar{x}), x' - x\rangle - \frac{\rho}{2}|x' - x|^2 + \frac{r}{2}|x' - \bar{x}|^2 \\ &\qquad - r\langle x - \bar{x}, x' - \bar{x}\rangle + \frac{r}{2}|x - \bar{x}|^2 \\ &= h(x, u, r) + \langle v + r(x - \bar{x}), x' - x\rangle - \frac{\rho}{2}|x' - x|^2 + \frac{r}{2}\big|(x - \bar{x}) - (x' - \bar{x})\big|^2 \\ &\geq h(x, u, r) + \langle v + r(x - \bar{x}), x' - x\rangle - \frac{\rho}{2}|x' - x|^2. \end{aligned}$$

Let $v^* = v + r(x - \bar{x})$. It follows that $v^* \in \partial_x h(x, u, r)$ and we have

$$h(x', u, r) \geq h(x, u, r) + \langle v^*, x' - x\rangle - \frac{\rho}{2}|x' - x|^2 \text{ whenever } \begin{cases} |x' - x_0| < \varepsilon, |x - x_0| < \varepsilon, \\ |u - u_0| < \varepsilon, \\ v^* = v + r(x - \bar{x}), v \in \partial_x f(x, u), \\ |v - \bar{v}| < \varepsilon, \text{ and} \\ f(x, u) \leq f(x_0, u_0) + \varepsilon. \end{cases}$$

Let $\varepsilon^* > 0$ be small enough so that $r_0 - \varepsilon^* > 0$ and

$$\max\left\{\varepsilon^*, \left[\varepsilon^* + (r_0 + \varepsilon^*)\varepsilon^*\right], \frac{\varepsilon^*}{2}|x_0 - \bar{x}|^2, \frac{(r_0 + \varepsilon^*)\varepsilon^*}{2}\left(2|x_0 - \bar{x}| + \varepsilon^*\right)\right\} \leq \frac{\epsilon}{3}$$

We shall show that

$$\left.\begin{array}{l} |x'-x_0|<\varepsilon^*, |x-x_0|<\varepsilon^*, \\ |u-u_0|<\varepsilon^*, |r-r_0|<\varepsilon^* \\ v^*\in\partial_x h(x,u,r), |v^*-v_0|<\varepsilon^*, \\ h(x,u,r)\le h(x_0,u_0,r_0)+\varepsilon^*, \end{array}\right\} \Rightarrow \begin{cases} |x'-x_0|<\varepsilon, |x-x_0|<\varepsilon, \\ |u-u_0|<\varepsilon, \\ v^*=v+r(x-\bar{x}), v\in\partial_x f(x,u), \\ |v-\bar{v}|<\varepsilon, \text{ and} \\ f(x,u)\le f(x_0,u_0)+\varepsilon, \end{cases}$$

which shows that equation (11) holds for our given ε^*.

Clearly $|x'-x_0|<\varepsilon^*, |x-x_0|<\varepsilon^*, |u-u_0|<\varepsilon^*$ implies $|x'-x_0|<\varepsilon, |x-x_0|<\varepsilon, |u-u_0|<\varepsilon$, while $v^*\in\partial_x h(x,u,r)$ implies $v^*=v+r(x-\bar{x})$ for some $v\in\partial f(x,u)$. Next note that $|v^*-v_0|<\varepsilon^*$ implies that $|v+r(x-\bar{x})-\bar{v}-r(x_0-\bar{x})|<\varepsilon^*$. This gives that $|v-\bar{v}|<\varepsilon^*+r|x-x_0|\le\varepsilon^*+r\varepsilon^*\le\varepsilon^*+(r_0+\varepsilon^*)\varepsilon^*<\varepsilon$ by our choice of ε^*.

Finally $h(x,u,r)\le h(x_0,u_0,r_0)+\varepsilon^*$ gives that

$$f(x,u)+\frac{r}{2}|x-\bar{x}|^2\le f(x_0,u_0)+\frac{r_0}{2}|x_0-\bar{x}|^2+\varepsilon^*$$

which implies that

$$\begin{aligned} f(x,u) &\le f(x_0,u_0)+\varepsilon^*+\frac{r_0}{2}|x_0-\bar{x}|^2-\frac{r}{2}|x-\bar{x}|^2 \\ &= f(x_0,u_0)+\varepsilon^*+\frac{r_0-r}{2}|x_0-\bar{x}|^2+\frac{r}{2}\big(|x_0-\bar{x}|^2)-|x-\bar{x}|^2\big) \\ &= f(x_0,u_0)+\varepsilon^*+\frac{r_0-r}{2}|x_0-\bar{x}|^2+\frac{r}{2}\langle x_0-x, x_0+x-2\bar{x}\rangle \quad \text{(easy calculation)} \\ &= f(x_0,u_0)+\varepsilon^*+\frac{r_0-r}{2}|x_0-\bar{x}|^2+\frac{r}{2}\langle x_0-x, 2x_0-2\bar{x}+x-x_0\rangle \\ &\le f(x_0,u_0)+\varepsilon^*+\frac{|r_0-r|}{2}|x_0-\bar{x}|^2+\frac{r}{2}|x_0-x|\;|2x_0-2\bar{x}+x-x_0|. \end{aligned}$$

Applying $|r-r_0|<\varepsilon^*$, $|x-x_0|<\varepsilon^*$, and the definition of ε^* implies,

$$\begin{aligned} f(x,u) &\le f(x_0,u_0)+\varepsilon^*+\frac{\varepsilon^*}{2}|x_0-\bar{x}|^2+\frac{(r_0+\varepsilon^*)\varepsilon^*}{2}\big(2|x_0-\bar{x}|+\varepsilon^*\big) \\ &\le f(x_0,u_0)+\frac{\varepsilon}{3}+\frac{\varepsilon}{3}+\frac{\varepsilon}{3} \\ &= f(x_0,u_0)+\varepsilon. \end{aligned}$$

If f is continuous as a function of $(x,u,v)\in\text{gph}\partial_x f$ then h is continuous as a function of $(x,u,r,v)\in\text{gph}\partial_x h$, so the statement of continuous prox-regularity holds. □

We now turn our attention to examining when h satisfies the constraint qualification (9).

Lemma 4.4. *Suppose the proper lsc function $f(x,u):\mathbb{R}^n\times\mathbb{R}^d\to\overline{\mathbb{R}}$ satisfies*

$$(0,y)\in\partial^\infty f(x_0,u_0)\Rightarrow y=0.$$

Then for any point $\bar{x}$ and $r_0>0$, the function $h(x,u,r):=f(x,u)+(r/2)|x-\bar{x}|^2$ satisfies

$$(0,y_u,y_r)\in\partial^\infty h(x_0,u_0,r_0)\Rightarrow y_u=0, \text{ and } y_r=0.$$

Proof. Note that (see [24, Prop. 10.5] for example)

$$\partial^\infty h(x,u,r) \subseteq [\partial^\infty f(x,u) \times \{0\}] + \partial^\infty(\frac{r}{2}|x-\bar{x}|^2).$$

As $(r/2)|x-\bar{x}|^2$ is $\mathcal{C}^2$, we have $\partial^\infty((r/2)|x-\bar{x}|^2) = 0$, which completes the proof. □

Finally we examine the two conditions regarding the coderivative of h.

Lemma 4.5. *Suppose the proper lsc function $f(x,u) : \mathbb{R}^n \times \mathbb{R}^d \to \overline{\mathbb{R}}$ is finite at the point (x_0,u_0) with $\bar{v} \in \partial_x f(x_0,u_0)$. Let $v_0 = \bar{v} + r_0(x_0 - \bar{x})$, then the function $h(x,u,r) := f(x,u) + (r/2)|x-\bar{x}|^2$ satisfies*

$$\begin{aligned} &(x',u',r') \in D^*(\partial_x h)(x_0,u_0,r_0|v_0)(v') \\ \Leftrightarrow\ &(x',u') \in D^*(\partial_x f)(x_0,u_0|\bar{v})(v') + (r_0 v', 0), \text{ and } r' = \langle (x_0-\bar{x}), v'\rangle \end{aligned} \tag{12}$$

Therefore if f satisfies

$$(x',u') \in D^*(\partial_x f)(x_0,u_0|\bar{v})(v'), \quad v' \neq 0 \quad \Rightarrow \quad \langle x',v'\rangle > -\rho|v'|^2, \tag{13}$$

then for any $r_0 \geq \rho$ one has

$$(x',u',r') \in D^*(\partial_x h)(x_0,u_0,r_0|v_0)(v') \text{ with } v' \neq 0 \quad \Rightarrow \quad \langle x',v'\rangle > 0.$$

Proof. Equation (12) follows from [24, Ex. 10.43], noting that $\nabla_x(r/2)|x-\bar{x}|^2 = r(x-\bar{x})$, thus $D^*(\nabla_x(r/2)|x-\bar{x}|^2)(x_0,u_0,r_0|v_0)(v') = (r_0v', 0, \langle(x_0-\bar{x}),v'\rangle)$.

For the second part of the proof, assume equation (13) holds. Select any $r_0 \geq \rho$ and consider some $(x',u',r') \in D^*(\partial_x h)(x_0,u_0,r_0|v_0)(v')$. By equation (12) we have $x' = \tilde{x} + r_0 v'$ for some $\tilde{x}$ with $(\tilde{x},u') \in D^*(\partial_x f)(x_0,u_0|\bar{v})(v')$. Therefore, if $v' \neq 0$, equation (13) tells us

$$\langle x',v'\rangle = \langle \tilde{x},v'\rangle + r_0|v'|^2 > -\rho|v'|^2 + r_0|v'|^2 \geq 0,$$

as desired. □

We are now ready to state our result on the full stability of the proximal mappings.

Theorem 4.6. *Let the proper lsc function $f(x,u) : \mathbb{R}^n \times \mathbb{R}^d \to \overline{\mathbb{R}}$ be continuously prox-regular at x_0 with compatible parameterization by u at u_0 for $0 \in \partial_x f(x_0,u_0)$. Assume further that f is prox-bounded with threshold r_{pb}, that f satisfies the constraint qualification*

$$(0,y) \in \partial^\infty f(x_0,u_0) \Rightarrow y = 0,$$

that

$$(0,u') \in D^*(\partial_x f)(x_0,u_0|0)(0) \Rightarrow u' = 0,$$

that for some $\rho > 0$

$$(x',u') \in D^*(\partial_x f)(x_0,u_0|0)(v'), \; v' \neq 0 \Rightarrow \langle x',v'\rangle > -\rho|v'|^2,$$

and that the set-valued mapping $\partial_x f(x_0,\cdot)$ has a "continuous selection" g near u_0 with $g(u_0) = 0$ (i.e. $g(u) \in \partial_x f(x_0,u)$ and g is continuous at u_0). Then for any r_0 sufficiently *large there exists $K > 0$ and a neighborhood of (x_0, u_0, r_0) such that for all x, x', u, u', r, r' in this neighborhood we have*

$$\big|P_r f_u(x) - P_{r'} f_{u'}(x')\big| \leq K\big|(r(x - x_0) - r'(x' - x_0), r - r', u - u')\big|. \tag{14}$$

where $f_u(x) = f(x, u)$.

Proof. By increasing ρ (if necessary) we may assume that for some $\varepsilon > 0$, f is continuously prox-regular at x_0 with compatible parameterization by u at u_0 for $0 \in \partial_x f(x_0, u_0)$ with parameters ρ and ε. Without loss of generality we may also assume $\rho > r_{pb}$.

Let $h(x, u, r) := f(x, u) + (r/2)|x - x_0|^2$. We begin by noting that, as $f(x_0, u_0)$ is finite (by definition of prox-regularity), $h(x_0, u_0, r_0)$ is finite. Let r_0 be sufficiently large so that $P_{r_0} f_{u_0}(x_0) = x_0$ (this is possible since $0 \in \partial f_{u_0}(x_0)$).

From Lemma 4.3 we know that h is continuously prox-regular in x at x_0 with compatible parameterization by (u, r) at (u_0, r_0) for $v_0 = 0 + r_0(x_0 - x_0) = 0$. From Lemma 4.5 we know that

$$\begin{aligned}(x', u', r') \in D^*(\partial_x h)(x_0, u_0, r_0|0)(v') &\text{ with } v' \neq 0 \Rightarrow \langle x', v'\rangle > 0, \text{ and}\\ (0, u', r') \in D^*(\partial_x h)(x_0, u_0, r_0|0)(0) &\Rightarrow u' = r' = 0.\end{aligned}$$

The requirements of Theorem 4.2 are fulfilled, hence x_0 is a fully stable locally optimal solution for h relative to $p_0 = (u_0, r_0)$ and $v_0 = 0$. Thus with

$$\begin{aligned} M_\delta^h(u, r, v) &= \underset{|x - x_0| \leq \delta}{\operatorname{argmin}} \{h(x, u, r) - \langle v, x\rangle\}\\ &= \underset{|x - x_0| \leq \delta}{\operatorname{argmin}} \left\{ f(x, u) - \langle v, x\rangle + \frac{r}{2}|x - x_0|^2 \right\}, \end{aligned}$$

there exists $\delta > 0$, a neighborhood of $(u_0, r_0, 0)$ and $K > 0$ such that

$$|M_\delta^h(u, r, v) - M_\delta^h(u', r', v')| \leq K|(v - v', r - r', u - u')| \tag{15}$$

for all u, u', r, r', v, v' in this neighborhood. We may assume that $r - \rho > 0$ for all r in this neighborhood of r_0.

We will now show that by taking a smaller neighborhood of $(u_0, r_0, 0)$ (if necessary), the minimization in $M_\delta^h(u, r, v)$ can be taken over the entire space i.e.

$$M_\delta^h(u, r, v) = M(u, r, v) := \underset{x}{\operatorname{argmin}} \left\{ f(x, u) - \langle v, x\rangle + \frac{r}{2}|x - x_0|^2 \right\}.$$

By taking a smaller neighborhood of u_0, we may assume (using the existence of a continuous selection of $\partial_x f(x_0, \cdot)$) that for all u in this neighborhood, there exists $v^* \in \partial_x f(x_0, u)$ with $|v^*| < \min\{\delta(r - \rho)/4, \varepsilon\}$. Since $f(x, u)$ is continuously prox-regular at x_0 with compatible parameterization by u at u_0 for $0 \in \partial_x f(x_0, u_0)$ we have

$$f(x, u) \geq f(x_0, u) + \langle v^*, x - x_0\rangle - \frac{\rho}{2}|x - x_0|^2$$

for all x close enough to x_0. Applying the prox-boundedness of f we may assume this inequality is global in x. Therefore,

$$f(x,u) - \langle v, x\rangle + \frac{r}{2}|x-x_0|^2 \geq f(x_0,u) - \langle v,x\rangle + \langle v^*, x-x_0\rangle + \frac{(r-\rho)}{2}|x-x_0|^2$$
$$= f(x_0,u) - \langle v,x_0\rangle - \langle v-v^*, x-x_0\rangle + \frac{(r-\rho)}{2}|x-x_0|^2.$$

If $4|v| \leq \delta(r-\rho)$ then $\delta(r-\rho) \geq 2|v-v^*|$. In this case, whenever $|x-x_0| > \delta$ then $(r-\rho)|x-x_0| > 2|v-v^*|$. This implies that

$$\langle v-v^*, x-x_0\rangle \leq |v-v^*|\;|x-x_0| < \big((r-\rho)/2\big)|x-x_0|^2,$$

i.e. $-\langle v-v^*, x-x_0\rangle + \big((r-\rho)/2\big)|x-x_0|^2 > 0$. Hence for all (u,r,v) in a neighborhood of $(u_0,r_0,0)$ and for all x with $|x-x_0| > \delta$ we have

$$f(x,u) - \langle v,x\rangle + \frac{r}{2}|x-x_0|^2 > f(x_0,u) - \langle v,x_0\rangle \geq m_\delta^h(u,r,v).$$

This shows that $M_\delta^h(u,r,v) \subseteq M(u,r,v)$. Since clearly $M(u,r,v) \subseteq M_\delta^h(u,r,v)$ we have $M_\delta^h(u,r,v) = M(u,r,v)$ for these (u,r,v).

Recall that $f_u(x) = f(x,u)$. Hence

$$M(u,r,v) = \operatorname*{argmin}_x \left\{ f_u(x) - \langle v,x\rangle + \frac{r}{2}|x-x_0|^2 \right\}.$$

In terms of proximal mappings and our notation, $M(u,r,v) = P_r(f_u)_v(x)$, where $(f_u)_v(\cdot) = f_u(\cdot) - \langle v,\cdot\rangle$. Hence $M(u,r,v) = P_r f_u(x_0 + (1/r)v)$ by Lemma 2.2 and, by (15),

$$|P_r f_u(x_0 + \frac{1}{r}v) - P_{r'} f_{u'}(x_0 + \frac{1}{r'}v')| \leq K|(v-v', r-r', u-u')|.$$

This gives for $x = x_0 + (1/r)v$ and $x' = x_0 + (1/r')v'$

$$|P_r f_u(x) - P_{r'} f_{u'}(x')| \leq K|(r(x-x_0) - r'(x'-x_0), r-r', u-u')|.$$

□

Note. If we omit the continuous selection assumption from Theorem 4.6 then we can prove a stability result for the proximal mappings of the functions $f_u(\cdot) + I_{\mathbb{B}(x_0,\delta)}(\cdot)$ where $I_{\mathbb{B}(x_0,\delta)}(x) = 0$ when $|x-x_0| \leq \delta$, $+\infty$ otherwise. To see this simply rewrite equation (15) in terms of proximal mappings of these functions. Hence the continuous selection assumption was needed to obtain a stability result for the proximal mappings of the functions $f_u(\cdot)$ instead of $f_u(\cdot) + I_{\mathbb{B}(x_0,\delta)}(\cdot)$; there may be other assumptions that will have the same effect.

By applying the remark after Lemma 2.2 to Theorem 4.6 we can easily consider the case when (x_0,u_0) is not a critical point of the function.

Corollary 4.7. *Let the proper lsc function* $f(x,u) : \mathbb{R}^n \times \mathbb{R}^d \to \overline{\mathbb{R}}$ *be continuously prox-regular at* x_0 *with compatible parameterization by* u *at* u_0 *for* $v_0 \in \partial_x f(x_0,u_0)$. *If the remaining conditions of Theorem* 4.6 *hold, with* v_0 *instead of the subgradient* 0, *then for any* r_0 *sufficiently large, there exists* $K > 0$ *and a neighborhood of* $(x_0 + (1/r)v_0, u_0, r_0)$ *such that for all* x, x', u, u', r, r' *in this neighborhood we have*

$$\big|P_r f_u(x) - P_{r'} f_{u'}(x')\big| \le K\big|(r(x-x_0) - r'(x'-x_0), r-r', u-u')\big|.$$

where $f_u(x) = f(x,u)$.

Proof. Set $f_{v_0}(\cdot) = f(\cdot) - \langle v_0, \cdot\rangle$, and note that f_{v_0} satisfies the remaining conditions of Theorem 4.6. Applying Theorem 4.6 we gain the stability of $P_r f_{v_0}$ near (x_0,u_0,r_0) which implies, as per the comments following Lemma 2.2, the stability of $P_r f$ near $(x_0 + (1/r)v_0, u_0, r_0)$. □

We conclude with a brief examination of the results of fixing any two of the three parameters in equation (14).

Corollary 4.8. *Under the assumptions of Corollary* 4.7, *there exists* $K > 0$ *and a neighborhood of* $(x_0 + (1/r)v_0, u_0, r_0)$ *such that for all* x, x', u, u', r, r' *in this neighborhood we have*

(*a*) $|P_r f_u(x) - P_{r'} f_u(x)| \le K|r-r'|\big(|x-x_0|+1|\big)$,
(*b*) $|P_r f_u(x) - P_r f_{u'}(x)| \le K|u-u'|$, *and*
(*c*) $|P_r f_u(x) - P_r f_u(x')| \le Kr|x-x'|$

Proof.

(a) $$\begin{aligned}|P_r f_u(x) - P_{r'} f_u(x)| &\le K|(r(x-x_0) - r'(x-x_0), r-r', u-u)|\\ &= K|(r-r')(x-x_0), r-r', 0)|\\ &\le K|r-r'|(|x-x_0|+1).\end{aligned}$$
(b) $|P_r f_u(x) - P_r f_{u'}(x)| \le K|(r(x-x_0) - r(x-x_0), r-r, u-u')| = K|u-u'|$.
(c) $|P_r f_u(x) - P_r f_u(x')| \le K|(r(x-x_0) - r(x'-x_0), r-r, u-u)| = Kr|x-x'|$.

□

References

[1] F. Bernard, L. Thibault. Prox-regular functions in Hilbert spaces, J. Math. Anal. Appl. 303(1) (2005) 1–14.

[2] F. Bernard, L. Thibault, D. Zagrodny: Integration of primal lower nice functions in Hilbert spaces, J. Optimization Theory Appl. 124(3) (2005) 561–579.

[3] C. L. Chiriac: Proximal point algorithm with generalized Bregman functions, Fixed Point Theory 6(1) (2005) 25–32.

[4] M. C. Ferris: Finite termination of the proximal point algorithm, Math. Program., Ser. A 50(3) (1991) 359–366.

[5] A. Hamdi: A primal-dual proximal point algorithm for constrained convex programs, Appl. Math. Comput. 162(1) (2005) 293–303.

[6] W. L. Hare, A. S. Lewis: Identifying active constraints via partial smoothness and prox-regularity, J. Convex Analysis 11 (2004) 251–266.

[7] W. L. Hare, C. Sagastizábal: Computing proximal points of nonconvex functions, Math. Program., to appear.

[8] S. Ibaraki, M. Fukushima: Primal-dual proximal point algorithm for multicommodity network flow problems, J. Oper. Res. Soc. Japan 37(4) (1994) 297–309.

[9] A. Iusem, R. D. C. Monteiro: On dual convergence of the generalized proximal point method with Bregman distances, Math. Oper. Res. 25(4) (2000) 606–624.

[10] A. N. Iusem, T. Pennanen, B. F. Svaiter: Inexact variants of the proximal point algorithm without monotonicity, SIAM J. Optim. 13(4) (2003) 1080–1097.

[11] C. Lemaréchal, F. Oustry, C. Sagastizábal. The U-Lagrangian of a convex function, Trans. Amer. Math. Soc. 352(2) (2000) 711–729.

[12] A. B. Levy, R. A. Poliquin, R. T. Rockafellar: Stability of locally optimal solutions, SIAM J. Optim. 10(2) (2000) 580–604.

[13] B. Martinet. Régularisation d'inéquations variationnelles par approximations successives, Rev. Franç. Inform. Rech. Opér. 4(R-3) (1970) 154–158.

[14] B. Martinet: Détermination approchée d'un point fixe d'une application pseudo-contractante. Cas de l'application prox, C. R. Acad. Sci., Paris, Sér. A 274 (1972) 163–165.

[15] R. Mifflin, C. Sagastizábal: Proximal points are on the fast track, J. Convex Analysis 9(2) (2002) 563–579.

[16] J.-J. Moreau. Proximité et dualité dans un espace hilbertien, Bull. Soc. Math. Fr. 93 (1965) 273–299.

[17] J.-J. Moreau: Propriétés des applications "prox", C. R. Acad. Sci., Paris 256 (1963) 1069–1071.

[18] T. Pennanen: Local convergence of the proximal point algorithm and multiplier methods without monotonicity, Math. Oper. Res. 27(1) (2002) 170–191.

[19] R. A. Poliquin, R. T. Rockafellar: Generalized hessian properties of regularized nonsmooth functions, SIAM J. Optim. 6(4) (1996) 1121–1137.

[20] R. A. Poliquin, R. T. Rockafellar: Prox-regular functions in variational analysis, Trans. Amer. Math. Soc. 348(5) (1996) 1805–1838.

[21] R. A. Poliquin, R. T. Rockafellar: Tilt stability of a local minimum, SIAM J. Optim. 8(2) (1998) 287–299.

[22] R. A. Poliquin, R. T. Rockafellar, L. Thibault: Local differentiability of distance functions, Trans. Amer. Math. Soc. 352(11) (2000) 5231–5249.

[23] R. T. Rockafellar: Monotone operators and the proximal point algorithm, SIAM J. Control Optimization 14(5) (1976) 877–898.

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