Journal of Convex Analysis
Volume 21 (2014), No. 4, 901–923

Parametrically Prox-Regular Functions

W. Hare*
University of British Columbia, Okanagan Campus (UBCO),
3333 University Way, Kelowna, BC, Canada
warren.hare@ubc.ca

C. Planiden†
University of British Columbia, Okanagan Campus (UBCO),
3333 University Way, Kelowna, BC, Canada
chayneplaniden@hotmail.com

Received: October 23, 2012
Revised manuscript received: June 13, 2013
Accepted: July 12, 2013

Prox-regularity is a generalization of convexity that includes all $\mathcal{C}^2$, lower-$\mathcal{C}^2$, strongly amenable, and primal-lower-nice functions. The study of prox-regular functions provides insight on a broad spectrum of important functions. Parametrically prox-regular (para-prox-regular) functions are a further extension of this family, produced by adding a parameter. Such functions have been shown to play a key role in understanding stability of minimizers in optimization problems. This document discusses para-prox-regular functions in $\mathbb{R}^n$.

We begin with some basic examples of para-prox-regular functions, and move on to the more complex examples of the convex and nonconvex proximal average. We develop an alternate representation of a para-prox-regular function, related to the monotonicity of an f-attentive ϵ-localization as was done for prox-regular functions [18]. This extends a result of Levy [13], who used an alternate approach to show one implication of the relationship (we provide a characterization). We analyze two common forms of parametrized functions that appear in optimization: finite parametrized sum of functions, and finite parametrized max of functions. The example of strongly amenable functions by Poliquin and Rockafellar [20] is given, and a relaxation of its necessary conditions is presented.

1. Introduction

Convex analysis has long been a prominent area of study in mathematics, and in optimization in particular. One method of unifying several classes of functions together with convex functions, that allows us to take advantage of their different properties, is through the notion of prox-regular functions.

Prox-regularity is a local property defined as follows. A function f is said to be prox-regular at $\bar{x}$ for a subgradient $\bar{v}$ if there exist $\epsilon > 0$ and $r \geq 0$ such that

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

*Research by this author was supported by NSERC of Canada.
†Research by this author was supported by UBC UGF and by NSERC of Canada.

ISSN 0944-6532 / $ 2.50 © Heldermann Verlag

under ϵ-neighborhoods of $\bar{x}$, $f(\bar{x})$ and $\bar{v}$ (a formal definition appears in Section 2). Prox-regularity was first introduced in the mid-nineties [7, 17, 18], when it was shown that all proper convex functions are prox-regular, as are lower-$\mathcal{C}^2$, strongly amenable, and primal-lower-nice functions [18]. The class of lower-$\mathcal{C}^2$ functions includes the class of $\mathcal{C}^2$ functions, thus the study of prox-regular functions gives us a way to connect the analyses of smooth functions and convex functions. It is quite common to see results that begin with a function of one of the classes mentioned above, and provide a related function that is of another of those classes. For example, Poliquin and Rockafellar [18] showed that if a function is prox-regular, then its Moreau envelope is a $\mathcal{C}^{1+}$ function.

Much of the theory in the field of prox-regularity has been extended from finite-dimensional space to Hilbert and Banach spaces [4, 3]. Prox-regularity is key in matters of function stability [19, 5, 12, 8], and it has been shown to extend results previously known for convex functions, such as the convergence rates of variational inequalities [16], alternating and averaged projection algorithms [15], active manifold identification [11], proximal averages [10], and proximal bundle methods [9].

Parametric prox-regularity, first introduced in 2000 by Levy, Poliquin and Rockafellar [14], extends prox-regularity to parametrized functions. The formal definition of parametric prox-regularity is presented in Definition 2.6. Little research has been done on parametrically prox-regular (henceforth para-prox-regular) functions to date; the purpose of this article is to take a step in that direction. We present an alternate definition similar to that of prox-regular functions given by Poliquin and Rockafellar [18]. We examine several common styles for constructing parametrized functions, and explore where the results are para-prox-regular.

The organization of this article is as follows. Section 2 starts with the definitions of subgradients, prox-regular functions, and para-prox-regular functions. It also provides some preliminary results for use in subsequent sections. Section 3 presents some examples of prox-regular and para-prox-regular functions. These include the proximal average [2, 1] and the NC-proximal average [10], which are parametrized functions designed to transform one function into another in a continuous manner. Section 4 contains an alternate definition of para-prox-regular functions. This result is similar to Poliquin's and Rockafellar's alternate definition of prox-regular functions, which is based on an f-attentive ϵ-localization T of ∂f around $(\bar{x}, \bar{v})$ (the prox-regularity of f is equivalent to the monotonicity of $T + rI$). Section 5 presents three methods of constructing parametrized functions, and discusses when each method results in a para-prox-regular function. As a corollary, we answer an open question posed in [10]. The methods presented are scalar multiplication of a function by a positive parameter, the parametrized sum of a finite collection of functions, and the finite parametrized max of continuous functions. An expansion of [20, Theorem 3.2] is also seen in Section 5, extending the required two initial prox-regular functions to any finite number of prox-regular functions. Section 6 poses some suggestions for future research.

2. Preliminaries

This section includes the definitions of prox-regular and para-prox-regular functions. Henceforth, functions are assumed to be proper and lower semicontinuous (lsc) as defined in [21], unless otherwise stated.

Definition 2.1 (Little-o). For two real functions f and g, we say $f(x) = o(g(x))$ at x_0 if and only if $\lim_{x\to x_0} \frac{f(x)}{g(x)} = 0$.

The first important definition we will need is that of *subgradients*. There are several types of subgradients; their definitions are restated here.

Definition 2.2 ([21, Def. 8.3]). Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$, $\bar{x} \in \operatorname{dom} f$. A vector $\bar{v} \in \mathbb{R}^n$ is

a) a **regular subgradient** of f at $\bar{x}$, written $\bar{v} \in \hat{\partial} f(\bar{x})$, if
$$f(x) \geq f(\bar{x}) + \langle \bar{v}, x - \bar{x}\rangle + o(|x - \bar{x}|). \tag{1}$$

b) a **(general) subgradient** of f at $\bar{x}$, written $\bar{v} \in \partial f(\bar{x})$, if there exist sequences $\{x^\nu\}_{\nu=1}^\infty \subseteq \operatorname{dom} f$ and $\{v^\nu\}_{\nu=1}^\infty \subseteq \hat{\partial} f(x^\nu)$ such that $(x^\nu, f(x^\nu)) \to (\bar{x}, f(\bar{x}))$ and $v^\nu \to \bar{v}$.

c) a **horizon subgradient** of f at $\bar{x}$, written $v \in \partial^\infty f(\bar{x})$, if the same holds as in b) except that instead of $v^\nu \to v$ one has $\lambda^\nu v^\nu \to v$ for some sequence $\lambda^\nu \searrow 0$, or in other words, $v^\nu \to \operatorname{dir}(v)$ (or $v = 0$).

d) a **proximal subgradient** of f at $\bar{x}$ if it is a regular subgradient whose error term $o(|x - \bar{x}|)$ is $\frac{r}{2}|x - \bar{x}|^2$ for some $r > 0$.

The set of all regular subgradients of f at $\bar{x}$ is denoted $\hat{\partial} f(\bar{x})$ and is called the *regular subdifferential* of f at $\bar{x}$. Similarly, the set of (general) subgradients is denoted $\partial f(\bar{x})$ and is called the *subdifferential* and the set of horizon subgradients is denoted $\partial^\infty f(\bar{x})$ and called the *horizon subdifferential*. For a function $f : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$, $f(x, \lambda)$, we shall also use $\partial_x f(\cdot, \cdot)$ to denote the subdifferential with respect to x.

Now we are ready for the definition of prox-regular functions.

Definition 2.3 ([21, Def. 13.27]). A proper function $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ that is finite at $\bar{x}$ is **prox-regular** at $\bar{x}$ for $\bar{v}$ with parameters $\epsilon > 0, r \geq 0$, where $\bar{v} \in \partial f(\bar{x})$, if and only if f is locally lsc at $\bar{x}$ and
$$f(x') \geq f(x) + \langle v, x' - x\rangle - \frac{r}{2}|x' - x|^2 \tag{2}$$
whenever $|x' - \bar{x}| < \epsilon$, $|x - \bar{x}| < \epsilon$, $|f(x) - f(\bar{x})| < \epsilon$, $v \in \partial f(x)$ and $|v - \bar{v}| < \epsilon$. If this holds true for all $\bar{v} \in \partial f(\bar{x})$, then f is said to be prox-regular at $\bar{x}$ (without reference to any subgradient). If f is prox-regular for all $\bar{x} \in \operatorname{dom} f$, then it is said to be a prox-regular function (without reference to any point).

By choosing ϵ slightly smaller, we may change to a strict inequality and rewrite as follows.

Corollary 2.4 ([20, Def. 1.1]). *A proper function $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is prox-regular at $\bar{x}$ for $\bar{v}$ with parameters $\epsilon > 0$, $r > 0$, if and only if f is locally lsc at $\bar{x}$, $\bar{v} \in \partial f(\bar{x})$, and*

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

whenever $x' \neq x$, $|x' - \bar{x}| < \epsilon$, $|x - \bar{x}| < \epsilon$, $|f(x) - f(\bar{x}| < \epsilon$, $|v - \bar{v}| < \epsilon$ and $v \in \partial f(x)$.

Notice that since we now have a strict inequality for r, we must include the condition $x' \neq x$. Graphically speaking, due to the last term on the right-hand side of (2), we can think of prox-regular functions as those functions which are locally bounded below by a tangent concave quadratic function. An alternate definition of prox-regularity, which will be used in Section 5, is described in the following theorem.

Theorem 2.5 ([18, Theorem 3.2]). *A proper function $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is prox-regular at $\bar{x}$ for $\bar{v}$ with parameters $\epsilon > 0, r \geq 0$ if and only if f is locally lsc at $\bar{x}$, $\bar{v} \in \partial f(\bar{x})$ and*

$$\langle v' - v, x' - x\rangle \geq -r|x' - x|^2$$

whenever $|x'-\bar{x}| < \epsilon$, $|x-\bar{x}| < \epsilon$, $|f(x')-f(\bar{x})| < \epsilon$, $|f(x)-f(\bar{x})| < \epsilon$, $v' \in \partial f(x')$, $v \in \partial f(x)$, $|v' - \bar{v}| < \epsilon$, and $|v - \bar{v}| < \epsilon$.

From Definition 2.3 we see that a prox-regular function is one that is locally bounded below by a quadratic function. Theorem 2.5 gives us an alternate expression that defines prox-regularity in terms of local pre-monotonicity of the subdifferential operator. This form is the inspiration for our main result in Section 4.

Para-prox-regularity is a natural extension of prox-regularity that allows for a parametrized function.

Definition 2.6 ([14, Definition 2.1]). The proper function $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R} \cup \{\infty\}$ is para-prox-regular in x at $\bar{x}$ for $\bar{v}$ with compatible parametrization by λ at $\bar{\lambda}$ (also said to be para-prox-regular in x at $(\bar{x}, \bar{\lambda})$ for $\bar{v}$), with parameters $\epsilon > 0, r \geq 0$, if and only if f is locally lsc at $(\bar{x}, \bar{\lambda})$, $\bar{v} \in \partial_x f(\bar{x}, \bar{\lambda})$ and

$$f(x', \lambda) \geq f(x, \lambda) + \langle v, x' - x\rangle - \frac{r}{2}|x' - x|^2 \tag{4}$$

whenever $|x'-\bar{x}| < \epsilon$, $|x-\bar{x}| < \epsilon$, $|f(x, \lambda)-f(\bar{x}, \bar{\lambda})| < \epsilon$, $v \in \partial_x f(x, \lambda)$, $|v-\bar{v}| < \epsilon$, and $|\lambda - \bar{\lambda}| < \epsilon$. The function f is continuously para-prox-regular in x at $(\bar{x}, \bar{\lambda})$ for $\bar{v}$ if, in addition, $f(x, \lambda)$ is continuous as a function of $(x, v, \lambda) \in \operatorname{gph} \partial_x f$ at $(\bar{x}, \bar{v}, \bar{\lambda})$. It is para-prox-regular in x at $(\bar{x}, \bar{\lambda})$ (with no mention of $\bar{v}$) if it is para-prox-regular in x at $(\bar{x}, \bar{\lambda})$ for all $\bar{v} \in \partial_x f(\bar{x}, \bar{\lambda})$, and it is a para-prox-regular function in x (with no mention of $\bar{x}$ or $\bar{\lambda}$) if it is para-prox-regular in x at $(\bar{x}, \bar{\lambda})$ for all $(\bar{x}, \bar{\lambda}) \in \operatorname{dom} f$.

A para-prox-regular function f of (x, λ) is a function that is prox-regular as a function of x at $(\bar{x}, \bar{\lambda})$, and continues to be so at (x, λ) for all x and for all λ in ϵ-neighborhoods of $\bar{x}$ and $\bar{\lambda}$, respectively. That is, if we do not wander too far away from $(\bar{x}, \bar{\lambda})$ either in x or in λ, the prox-regularity property of f and its corresponding prox-regularity parameters are preserved.

3. Examples

3.1. Basic Examples

We begin with a few basic examples of para-prox-regular functions. The first example is the result of multiplying the function $|x|$ by the parameter λ.

Example 3.1. For $x \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$, define

$$f(x, \lambda) := \lambda |x|.$$

Then f is para-prox-regular at $\bar{x} = 0$ with compatible parametrization by λ at any $\bar{\lambda} > 0$. It is not para-prox-regular at $\bar{x} = 0$ at any $\bar{\lambda} \leq 0$.

Proof. Let $\bar{x} = 0$ and $\bar{\lambda} > 0$. Then

$$f(x', \lambda) \geq f(x, \lambda) + \langle v, x' - x \rangle$$

for any x', x and for any $v \in \partial_x f(x, \lambda)$, as f is convex in x and $\lambda > 0$. Thus, for $\epsilon = \frac{\bar{\lambda}}{2}$ and $r = 0$ we have

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

for $|x' - \bar{x}| < \epsilon$, $|x - \bar{x}| < \epsilon$, $|f(x, \lambda) - f(\bar{x}, \bar{\lambda})| < \epsilon$, $v \in \partial_x f(x, \lambda)$, $|v - \bar{v}| < \epsilon$, and $|\lambda - \bar{\lambda}| < \epsilon$. That is, f is para-prox-regular at $(\bar{x}, \bar{\lambda})$.

For $\bar{\lambda} < 0$, f ceases to be a prox-regular function at $\bar{x} = 0$. This can be verified by noting that $\partial_x f(\bar{x}, \bar{\lambda}) = \{|\bar{\lambda}|, -|\bar{\lambda}|\}$ and that there does not exist (ϵ, r) that satisfies Definition 2.3 in this case. Hence, it is not para-prox-regular either. For $\bar{\lambda} = 0$ para-prox-regularity fails as well, since we require f to be prox-regular in a neighborhood of $\bar{\lambda}$, but any neighborhood of $\bar{\lambda} = 0$ necessarily contains some $\lambda < 0$ and yields a function that is not prox-regular. □

The previous example hints at the next result: if a function is convex in terms of x no matter the value of λ, then it is para-prox-regular. This result will be used in Subsection 3.2 to prove that the proximal average is para-prox-regular.

Example 3.2. Let $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R} \cup \{\infty\}$ be convex as a function of x for all λ such that $|\lambda - \bar{\lambda}| < \epsilon$, where $\epsilon > 0$. Then f is para-prox-regular with compatible parametrization by λ at $\bar{\lambda}$.

Proof. Since $f(x,\bar{\lambda})$ is convex in terms of x, for any $x \in \operatorname{dom} f(\cdot,\bar{\lambda})$ and any $v \in \partial_x f(x,\bar{\lambda})$ we have that

$$f(x',\bar{\lambda}) \geq f(x,\bar{\lambda}) + \langle v, x'-x\rangle \tag{5}$$

for all x'. For every λ such that $|\lambda-\bar{\lambda}|<\epsilon$, since f remains convex, for any $x \in \operatorname{dom} f(\cdot,\lambda)$ and any $v_\lambda \in \partial f_x(x,\lambda)$ we have that

$$f(x',\lambda) \geq f(x,\lambda) + \langle v_\lambda, x'-x\rangle \tag{6}$$

for all x'. In particular, inequality (6) is true for all x' and x such that $|x'-\bar{x}|<\epsilon$, $|x-\bar{x}|<\epsilon$ and $|f(x,\lambda)-f(\bar{x},\bar{\lambda})|<\epsilon$ for any $\bar{x}$ fixed, and for all $v_\lambda \in \partial_x f(x,\lambda)$ such that $|v_\lambda - v|<\epsilon$. This demonstrates para-prox-regularity with parameters $r=0$ and ϵ. □

The next two examples demonstrate that we can take a prox-regular function f and use the parameter λ in the argument of f to create para-prox-regular functions at particular $\bar{\lambda}$ values. Example 3.3 examines a linear shift of the argument by λ and Example 3.4 examines a scalar multiple of the argument by λ.

Example 3.3. Let $\tilde{f}:\mathbb{R}^n \to \mathbb{R}\cup\{\infty\}$ be prox-regular at $\bar{x}$ for $\bar{v}\in\partial\tilde{f}(\bar{x})$, and let $\lambda\in\mathbb{R}^n$. Define

$$f(x,\lambda) := \tilde{f}(x-\lambda).$$

Then f is para-prox-regular at $\bar{x}$ with compatible parametrization by λ at $\bar{\lambda}=0$.

Proof. By assumption, there exist $\tilde{\epsilon}>0$ and $\tilde{r}\geq 0$ such that

$$\tilde{f}(\tilde{x}') \geq \tilde{f}(\tilde{x}) + \langle v, \tilde{x}'-\tilde{x}\rangle - \frac{\tilde{r}}{2}|\tilde{x}'-\tilde{x}|^2 \tag{7}$$

whenever $\tilde{x}'\neq\tilde{x}$, $|\tilde{x}'-\bar{x}|<\tilde{\epsilon}$, $|\tilde{x}-\bar{x}|<\tilde{\epsilon}$, $|\tilde{f}(\tilde{x})-\tilde{f}(\bar{x})|<\tilde{\epsilon}$, $v\in\partial\tilde{f}(\tilde{x})$, $|v-\bar{v}|<\tilde{\epsilon}$. We need to show that when $\bar{\lambda}=0$, for all $\bar{v}\in\partial\tilde{f}(\bar{x})=\partial_x f(\bar{x},\bar{\lambda})$ there exist $\epsilon>0$ and $r\geq 0$ such that

$$f(x',\lambda) \geq f(x,\lambda) + \langle v, x'-x\rangle - \frac{r}{2}|x'-x|^2, \tag{8}$$

when $x'\neq x$, $|x'-\bar{x}|<\epsilon$, $|x-\bar{x}|<\epsilon$, $|\lambda-\bar{\lambda}|<\epsilon$, $|f(x,\lambda)-f(\bar{x},\bar{\lambda})|<\epsilon$, $v\in\partial_x f(x,\lambda)$ and $|v-\bar{v}|<\epsilon$. Let $\epsilon=\frac{\tilde{\epsilon}}{2}$, $r=\tilde{r}$. Since $\epsilon<\tilde{\epsilon}$, inequality (7) holds when we replace $\tilde{r}$ with r and $\tilde{\epsilon}$ with ϵ everywhere. Select any λ such that $|\lambda-\bar{\lambda}|<\epsilon$, and set $\tilde{x}'=x'-\lambda$ and $\tilde{x}=x-\lambda$. Notice that if $|x'-\bar{x}|<\epsilon$ and $|x-\bar{x}|<\epsilon$, then by the triangle inequality $|(x'-\lambda)-\bar{x}|<\tilde{\epsilon}$ and $|(x-\lambda)-\bar{x}|<\tilde{\epsilon}$, which is to say $|\tilde{x}'-\bar{x}|<\tilde{\epsilon}$ and $|\tilde{x}-\bar{x}|<\tilde{\epsilon}$. If in addition we restrict our function values using ϵ instead of $\tilde{\epsilon}$, then we have

$$|\tilde{f}(\tilde{x})-\tilde{f}(\bar{x})|<\epsilon \Leftrightarrow |\tilde{f}(x-\lambda)-\tilde{f}(\bar{x}-\bar{\lambda})|<\epsilon \Leftrightarrow |f(x,\lambda)-f(\bar{x},\bar{\lambda})|<\epsilon.$$

Finally we observe that $v \in \partial \tilde{f}(\tilde{x})$ if and only if $v \in \partial_x f(x, \lambda)$, so we may rewrite inequality (7) as

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

when $x' \neq x$, $|x' - \bar{x}| < \epsilon$, $|x - \bar{x}| < \epsilon$, $|\lambda - \bar{\lambda}| < \epsilon$, $|f(x, \lambda) - f(\bar{x}, \bar{\lambda})| < \epsilon$, $v \in \partial_x f(x, \lambda)$, $|v - \bar{v}| < \epsilon$. That is, inequality (8) holds under the appropriate conditions. □

Example 3.4. Let $\tilde{f} : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be prox-regular at $\bar{x}$ for $\bar{v}$, and let $\lambda \in \mathbb{R}$. Define

$$f(x, \lambda) := \tilde{f}(\lambda x).$$

Then f is para-prox-regular at $\bar{x}$ for $\bar{v}$ with compatible parametrization by λ at $\bar{\lambda} = 1$.

Proof. By assumption, there exist prox-regularity parameters $\tilde{\epsilon} > 0$ and $\tilde{r} \geq 0$ such that

$$\tilde{f}(\tilde{x}') \geq \tilde{f}(\tilde{x}) + \langle v, \tilde{x}' - \tilde{x} \rangle + \frac{\tilde{r}}{2}|\tilde{x}' - \tilde{x}|^2 \tag{9}$$

for $|\tilde{x}' - \bar{x}| < \tilde{\epsilon}$, $|\tilde{x} - \bar{x}| < \tilde{\epsilon}$, $|\tilde{f}(\tilde{x}) - \tilde{f}(\bar{x})| < \tilde{\epsilon}$, $v \in \partial \tilde{f}(\tilde{x})$, and $|v - \bar{v}| < \tilde{\epsilon}$. Substituting $\tilde{x}' = \lambda x'$ and $\tilde{x} = \lambda x$, inequality (9) becomes

$$\tilde{f}(\lambda x') \geq \tilde{f}(\lambda x) + \langle v, \lambda x' - \lambda x \rangle - \frac{\tilde{r}}{2}|\lambda x' - \lambda x|^2 \tag{10}$$

for $|\lambda x' - \bar{x}| < \tilde{\epsilon}$, $|\lambda x - \bar{x}| < \tilde{\epsilon}$, $|\tilde{f}(\lambda x) - \tilde{f}(\bar{x})| < \tilde{\epsilon}$, $v \in \partial \tilde{f}(\lambda x)$, and $|v - \bar{v}| < \tilde{\epsilon}$. The subgradient chain rule [21, Ex. 10.7] gives us $\lambda v \in \partial_x f(x, \lambda)$, so this may be rewritten as

$$f(x', \lambda) \geq f(x, \lambda) + \langle \lambda v, x' - x \rangle - \frac{\tilde{r}|\lambda|^2}{2}|x' - x|^2 \tag{11}$$

for $|\lambda x' - \bar{x}| < \tilde{\epsilon}$, $|\lambda x - \bar{x}| < \tilde{\epsilon}$, $|f(x, \lambda) - f(\bar{x}, \bar{\lambda})| < \tilde{\epsilon}$, $\lambda v \in \partial_x f(x, \lambda)$, and $|v - \bar{v}| < \tilde{\epsilon}$. If we restrict our neighborhood of $\bar{x}$ further, say $|x' - \bar{x}| < \frac{\tilde{\epsilon}}{2}$ and $|x - \bar{x}| < \frac{\tilde{\epsilon}}{2}$, then there exists $\hat{\epsilon} > 0$ such that for all λ with $|\lambda - \bar{\lambda}| < \hat{\epsilon}$ and if $|x' - \bar{x}| < \frac{\tilde{\epsilon}}{2}$ and $|x - \bar{x}| < \frac{\tilde{\epsilon}}{2}$, then we necessarily have that $|\lambda x' - \bar{x}| < \tilde{\epsilon}$ and $|\lambda x - \bar{x}| < \tilde{\epsilon}$. Hence by taking $\epsilon = \min\{\frac{\tilde{\epsilon}}{2}, \hat{\epsilon}\}$, inequality (11) remains true for $|x' - \bar{x}| < \epsilon$, $|x - \bar{x}| < \epsilon$, $|f(x, \lambda) - f(\bar{x}, \bar{\lambda})| < \epsilon$, $\lambda v \in \partial_x f(x, \lambda)$, $|v - \bar{v}| < \epsilon$, and $|\lambda - \bar{\lambda}| < \epsilon$. Since $|\lambda - \bar{\lambda}| < \epsilon$, we may take $r = \max\{\tilde{r}, \tilde{r}|\bar{\lambda} + \epsilon|^2\}$. Renaming λv as $\tilde{v}$ we have that

$$f(x', \lambda) \geq f(x, \lambda) + \langle \tilde{v}, x' - x \rangle - \frac{r}{2}|x' - x|^2 \tag{12}$$

for $|x' - \bar{x}| < \epsilon$, $|x - \bar{x}| < \epsilon$, $|\lambda - \bar{\lambda}| < \epsilon$, $|f(x, \lambda) - f(\bar{x}, \bar{\lambda})| < \epsilon$, $\tilde{v} \in \partial_x f(x, \lambda)$, and $|\tilde{v} - \bar{v}| < \epsilon$. □

3.2. Proximal Average

The next examples are those of the *proximal average* and *NC-proximal average*, for which we will need some background information. Recall that for a proper function f, the *Fenchel conjugate* f^* is defined via

$$f^*(y) := \sup_y\{\langle x, y\rangle - f(y)\}.$$

For two convex functions f_0 and f_1, using the Fenchel conjugate, we define the proximal average.

Definition 3.5 ([1, Definition 4.1]). The *proximal average* PA_{f_0,f_1} of two proper convex functions f_0 and f_1 is defined via

$$PA_{f_0,f_1}(x,\lambda) := \left((1-\lambda)\left(f_0 + \frac{1}{2}|\cdot|^2\right)^* + \lambda\left(f_1 + \frac{1}{2}|\cdot|^2\right)^*\right)^*(x) - \frac{1}{2}|x|^2$$

for $\lambda \in [0,1]$.

The proximal average has been shown to be an effective method of transforming one convex function into another in a continuous manner [1]. It has also been shown to be a convex function [2], and as such it is prox-regular. In fact, it is para-prox-regular, as the following proposition demonstrates.

Proposition 3.6. *The proximal average PA_{f_0,f_1} of two proper convex functions f_0, f_1 is para-prox-regular with compatible parametrization by λ at $\bar{\lambda}$, for all $\bar{\lambda} \in (0,1)$.*

Proof. Bauschke, Matoušková and Reich [2, Theorem 6.1] proved that this function is convex in x for all $\lambda \in (0,1)$, therefore it satisfies the conditions of Example 3.2. □

In [10] an alternate representation of the proximal average function was given using Moreau envelopes. Recall that for a function f the *Moreau envelope* with parameter $r > 0$ is defined as

$$e_r f(x) := \inf_y \left\{f(y) + \frac{r}{2}|y-x|^2\right\},$$

and its associated *proximal point mapping* is defined as

$$P_r f(x) := \operatorname{argmin}_y \left\{f(y) + \frac{r}{2}|y-x|^2\right\}.$$

If f is proper and convex, then $e_r f$ is proper for all $r > 0$. Properties of the Moreau envelope give rise to the equivalent definition of proximal average [10, Equation (4)],

$$PA_{f_0,f_1}(x,\lambda) = -e_1(-(1-\lambda)e_1 f_0 - \lambda e_1 f_1)(x).$$

However, this formulation is only useful if f_0 and f_1 are convex functions. For two not necessarily convex functions f_0 and f_1, the *NC-proximal average function* PA_r is similarly defined. We need one more definition first, that of prox-boundedness.

Definition 3.7. A function f is *prox-bounded* if there exist $r > 0$ and a point $\bar{x}$ such that $e_r f(\bar{x}) > -\infty$. The infimum of the set of all such r is called the *threshold of prox-boundedness*.

Definition 3.8 ([10, Equation (5)]). Let f_0 and f_1 be proper, prox-bounded functions, and let $r > 0$ be greater than the threshold of prox-boundedness for both f_0 and f_1. The *NC-proximal average* of f_0 and f_1 is defined as

$$PA_r(x, \lambda) := -e_{r+\lambda(1-\lambda)}(-(1-\lambda)e_r f_0 - \lambda e_r f_1)(x).$$

Showing that the NC-proximal average is para-prox-regular requires a technical condition from [10, Theorem 4.6], specifically that the Lipschitz constant of $\lambda P_r f_0 + (1-\lambda)P_r f_1 - I$ is bounded by 1. This condition ensures that the NC-proximal average is well-behaved.

Example 3.9. Consider two proper, prox-bounded, prox-regular functions f_0 and f_1. If the Lipschitz constant of $\lambda P_r f_0 + (1-\lambda)P_r f_1 - I$ is bounded by 1, then for r sufficiently large, the NC-proximal average is para-prox-regular for all $\bar{\lambda} \in (0, 1)$.

Proof. Fix $\bar{\lambda} \in (0, 1)$. Let $\bar{v} \in \partial_x PA_r(\bar{x}, \bar{\lambda})$. Enlarging r if necessary, we know that PA_r is locally Lipschitz continuous in λ [10, Theorem 4.6], and that it is a lower-$\mathcal{C}^2$ function in x [10, Proposition 2.5]. Since all lower-$\mathcal{C}^2$ functions are prox-regular, we have that PA_r is prox-regular as a function of x at $\bar{x}$ for $\bar{v}$, say with parameters $\epsilon_1 > 0$ and $\rho \geq 0$. Therefore

$$PA_r(x', \bar{\lambda}) > PA_r(x, \bar{\lambda}) + \langle v, x' - x\rangle - \frac{\rho}{2}|x' - x|^2 \tag{13}$$

when $x' \neq x$, $|x' - \bar{x}| < \epsilon_1$, $|x - \bar{x}| < \epsilon_1$, $|PA_r(x, \bar{\lambda}) - PA_r(\bar{x}, \bar{\lambda})| < \epsilon_1$, $v \in \partial_x PA_r(x, \bar{\lambda})$, and $|v - \bar{v}| < \epsilon_1$. We have established that PA_r is continuous in λ, and therefore there exists $\epsilon_2 > 0$ such that inequality (13) remains true when $\bar{\lambda}$ is replaced by any λ such that $|\lambda - \bar{\lambda}| < \epsilon_2$. We can reduce ϵ_2 if necessary, so that $|\lambda - \bar{\lambda}| < \epsilon_2$ implies $\lambda \in (0, 1)$. Taking $\epsilon_3 = \min\{\epsilon_1, \epsilon_2\}$, we have

$$PA_r(x', \lambda) > PA_r(x, \lambda) + \langle v, x' - x\rangle - \frac{\rho}{2}|x' - x|^2 \tag{14}$$

when $x' \neq x$, $|x' - \bar{x}| < \epsilon_3$, $|x - \bar{x}| < \epsilon_3$, $|PA_r(x, \lambda) - PA_r(\bar{x}, \bar{\lambda})| < \epsilon_3$, $v \in \partial_x PA_r(x, \bar{\lambda})$, $|v - \bar{v}| < \epsilon_3$, and $|\lambda - \bar{\lambda}| < \epsilon_3$. If we use $\frac{\epsilon_3}{2}$-neighborhoods instead of ϵ_3-neighborhoods, then inequality (14) certainly remains true. Since PA_r is $\mathcal{C}^1$ with respect to x, we know that $\partial_x PA_r(x, \bar{\lambda}) = \{\nabla_x PA_r(x, \bar{\lambda})\}$, and by [10, Theorem 4.6] we know that $\nabla_x PA_r(x, \bar{\lambda})$ changes in a continuous manner with respect to λ. Hence, there exists $\epsilon_4 > 0$ such that inequality (14) remains valid when $v \in \partial_x PA_r(x, \bar{\lambda})$ is replaced by $\hat{v} \in \partial_x PA_r(x, \lambda)$ for $|\hat{v} - v| < \epsilon_4$ and

$|\lambda - \bar{\lambda}| < \epsilon_4$. Shrinking ϵ_4 if necessary, we can ensure that $\epsilon_4 < \frac{\epsilon_3}{2}$. Thus we have $|\hat{v} - v| < \frac{\epsilon_3}{2}$ and $|v - \bar{v}| < \frac{\epsilon_3}{2}$, which implies that $|\hat{v} - \bar{v}| < \frac{\epsilon_3}{2} + \frac{\epsilon_3}{2} = \epsilon_3$. Therefore, choosing $\epsilon = \min\{\frac{\epsilon_3}{2}, \epsilon_4\}$ we have that

$$PA_r(x', \lambda) > PA_r(x, \lambda) + \langle \hat{v}, x' - x \rangle - \frac{\rho}{2}|x' - x|^2 \tag{15}$$

when $x' \neq x$, $|x' - \bar{x}| < \epsilon$, $|x - \bar{x}| < \epsilon$, $|PA_r(x, \lambda) - PA_r(\bar{x}, \bar{\lambda})| < \epsilon$, $\hat{v} \in \partial_x PA_r(x, \lambda)$, $|\hat{v} - \bar{v}| < \epsilon$, and $|\lambda - \bar{\lambda}| < \epsilon$. □

4. Monotonicity

This section contains an alternate definition of para-prox-regular functions. We begin with the definition of an *f-attentive localization* of the subdifferential. An f-attentive localization of $\partial_x f$ around $(\bar{x}, \bar{\lambda}, \bar{v})$ for a fixed λ is a set-valued mapping $T_\lambda : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ whose graph in $\mathbb{R}^n \times \mathbb{R}^n$ is the intersection of $\operatorname{gph} \partial_x f$ with an f-attentive neighborhood of $\bar{x}$ and an ordinary neighborhood of $\bar{v}$. (This contrasts with an ordinary localization, in which the f-attentive neighborhood of $\bar{x}$ is relaxed to an ordinary neighborhood.) The following definition is adapted from [18, Def. 3.1].

Definition 4.1. For an $\epsilon > 0$ and a fixed λ, the *f-attentive ϵ-localization of ∂f_x around* $(\bar{x}, \bar{\lambda}, \bar{v})$ is

$$T_\lambda(x) = \begin{cases} \{v \in \partial_x f(x, \lambda) : |v - \bar{v}| < \epsilon\}, & |x - \bar{x}| < \epsilon, |f(x, \lambda) - f(\bar{x}, \bar{\lambda})| < \epsilon \\ \emptyset, & \text{otherwise.} \end{cases} \tag{16}$$

With this definition in mind, we shall present Theorem 4.3, which is a para-prox-regular analog of [18, Theorem 3.2], and Corollary 4.4, which is the alternate definition of para-prox-regular functions that we seek. In order to prove Theorem 4.3, we use the following lemma. The lemma takes a para-prox-regular function at $\bar{x}$ for $\bar{v}$ (with compatible parametrization by λ at $\bar{\lambda}$) and forms a new function that is para-prox-regular at 0 for 0 (with compatible parametrization by λ at $\bar{\lambda}$), in effect shifting the original function. This will allow us, in Theorem 4.3, to make some assumptions without loss of generality in order to simplify the proof.

Lemma 4.2. *Let f be para-prox-regular at $\bar{x}$ for $\bar{v}$ with compatible parametrization by λ at $\bar{\lambda}$. Define*

$$\tilde{f}(x, \lambda) := f(x - \bar{x}, \lambda) - \langle \bar{v}, x - \bar{x} \rangle.$$

Then $\tilde{f}$ is para-prox-regular at 0 for 0 with compatible parametrization by λ at $\bar{\lambda}$.

Proof. By assumption, there exist $\epsilon > 0$ and $r \geq 0$ such that

$$f(x', \lambda) \geq f(x, \lambda) + \langle v, x' - x \rangle - \frac{r}{2}|x - \bar{x}|^2 \tag{17}$$

whenever $x' \neq x$, $|x' - \bar{x}| < \epsilon$, $|x - \bar{x}| < \epsilon$, $|f(x,\lambda) - f(\bar{x},\bar{\lambda})| < \epsilon$, $v \in \partial_x f(x,\lambda)$, $|v - \bar{v}| < \epsilon$, and$|\lambda - \bar{\lambda}| < \epsilon$. For ease of discussion let $f_1(x,\lambda) = f(x - \bar{x}, \lambda)$ and $f_2(x,\lambda) = -\langle \bar{v}, x - \bar{x}\rangle$. Notice that $v \in \partial_x f(x,\lambda)$ implies $v - \bar{v} \in \partial_x f(x,\lambda) - \bar{v}$. Since $f_2 \in \mathcal{C}^2$, applying [21, Exercise 8.8(c)] we have that

$$\partial_x f(x,\lambda) - \bar{v} = \partial_x[f(x - \bar{x}, \lambda) - \langle \bar{v}, x - \bar{x}\rangle] = \partial_x \tilde{f}(x,\lambda).$$

Therefore $v - \bar{v} \in \partial_x \tilde{f}(x,\lambda)$. Translating inequality (17) by $\bar{x}$, we may rewrite it as

$$f(x' - \bar{x}, \lambda) \geq f(x - \bar{x}, \lambda) + \langle v, (x' - \bar{x}) - (x - \bar{x})\rangle - \frac{r}{2}|(x' - \bar{x}) - (x - \bar{x})|^2 \quad (18)$$

now with the conditions $x' \neq x$, $|x'| < \epsilon$, $|x| < \epsilon$, $|\tilde{f}(x,\lambda) - \tilde{f}(0,\bar{\lambda})| < \epsilon$, $v - \bar{v} \in \partial_x \tilde{f}(x,\lambda)$, $|v - \bar{v}| < \epsilon$, and $|\lambda - \bar{\lambda}| < \epsilon$. Subtracting $\langle \bar{v}, x - \bar{x}\rangle$ from both sides of inequality (18) and simplifying, we get

$$\tilde{f}(x',\lambda) \geq \tilde{f}(x,\lambda) + \langle v, x' - x\rangle - \frac{r}{2}|x' - x|^2. \quad (19)$$

Therefore $\tilde{f}$ is para-prox-regular at 0 for 0 with compatible parametrization by λ at $\bar{\lambda}$. □

Theorem 4.3. *Let $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R} \cup \{\infty\}$ be a proper function that is locally lsc near $(\bar{x}, \bar{\lambda})$. Then the following are equivalent.*

a) *The function f is para-prox-regular at $\bar{x}$ for $\bar{v}$ with compatible parametrization by λ at $\bar{\lambda}$.*

b) *The vector $\bar{v}$ is a proximal subgradient of f with respect to x at $(\bar{x}, \bar{\lambda})$, and there exist $\epsilon > 0$, $r > 0$ such that the f-attentive ϵ-localization T_λ of $\partial_x f$ at $(\bar{x}, \bar{\lambda}, \bar{v})$ has $T_\lambda + rI_x$ monotone, for all λ such that $|\lambda - \bar{\lambda}| < \epsilon$. That is, there is a constant r such that*

$$\langle v_1 - v_0, x_1 - x_0\rangle \geq -r|x_1 - x_0|^2 \quad (20)$$

when $v_i \in T_\lambda(x_i)$, $i \in \{0,1\}$, and $|\lambda - \bar{\lambda}| < \epsilon$. Here I_x is the identity mapping with respect to x.

c) *There exist $\epsilon > 0, r > 0$ such that*

$$f(x,\lambda) > f(\bar{x},\bar{\lambda}) + \langle \bar{v}, x - \bar{x}\rangle - \frac{r}{2}|x - \bar{x}|^2 \quad (21)$$

when $|x - \bar{x}| < \epsilon$ and $|\lambda - \bar{\lambda}| < \epsilon$, and the mapping $(\partial_x f + rI_x)^{-1}$ has the following single-valuedness property near $\bar{z} = \bar{v} + r\bar{x}$: if $|z - \bar{z}| < \epsilon$, then for $i \in \{0,1\}$ one has

$$(x_i,\lambda) \in (\partial_x f + rI_x)^{-1}(z) \Rightarrow x_0 = x_1 \quad (22)$$

when $|x_i - \bar{x}| < \epsilon$, $|f(x_i,\lambda) - f(\bar{x},\bar{\lambda})| < \epsilon$, and $|\lambda - \bar{\lambda}| < \epsilon$.

Proof. *a)* $\Rightarrow$ *b)*: We have that f is para-prox-regular at $(\bar{x}, \bar{\lambda})$ for $\bar{v}$, so there exist $\epsilon > 0$ and $r > 0$ such that

$$\begin{aligned} f(x_1, \lambda) &\geq f(x_0, \lambda) + \langle v_0, x_1 - x_0 \rangle - \tfrac{r}{2}|x_1 - x_0|^2 \\ f(x_0, \lambda) &\geq f(x_1, \lambda) + \langle v_1, x_0 - x_1 \rangle - \tfrac{r}{2}|x_0 - x_1|^2 \end{aligned} \tag{23}$$

whenever $i \in \{0, 1\}$, $|x_i - \bar{x}| < \epsilon$, $|f(x_i, \lambda) - f(\bar{x}, \bar{\lambda})| < \epsilon$, $v_i \in \partial_x f(x_i, \lambda)$, $|v_i - \bar{v}| < \epsilon$, and $|\lambda - \bar{\lambda}| < \epsilon$. Adding the two inequalities in (23) together, we get

$$f(x_1, \lambda) + f(x_0, \lambda) \geq f(x_0, \lambda) + f(x_1, \lambda) + \langle v_0 - v_1, x_1 - x_0 \rangle - r|x_1 - x_0|^2,$$

which simplifies to inequality (20). Finally, since f is para-prox-regular at $\bar{x}$ for $\bar{v}$, it is prox-regular there as well, giving us that $\bar{v}$ is a proximal subgradient.

b) $\Rightarrow$ *c)*: For this section of the proof and the next, without loss of generality we assume that $\bar{x} = 0$, $f(0, \bar{\lambda}) = 0$, and $\bar{v} = 0$ (by Lemma 4.2). Let $\bar{\epsilon} > 0$ and $\bar{r} > 0$ be the parameters assumed by condition *b)*. Since $\bar{v}$ is a proximal subgradient of f with respect to x at $(\bar{x}, \bar{\lambda})$, we have

$$f(x, \lambda) > f(\bar{x}, \bar{\lambda}) + \langle \bar{v}, x - \bar{x} \rangle - \frac{r}{2}|x - \bar{x}|^2 \tag{24}$$

when $|x - \bar{x}| < \epsilon$ and $|\lambda - \bar{\lambda}| < \epsilon$, for some $\epsilon \in (0, \bar{\epsilon})$ and some $r \in (\bar{r}, \infty)$. Let $\lambda \in [\bar{\lambda} - \epsilon, \bar{\lambda} + \epsilon]$. By decreasing ϵ and increasing r if necessary, we can arrange that $r \geq 1$ and $\epsilon < \frac{\bar{\epsilon}}{2}$. Suppose that for z such that $|z - (\bar{v} + r\bar{x})| < \epsilon$ we have $(x_i, \lambda) \in (\partial_x f + rI_x)^{-1}(z)$, $i \in \{0, 1\}$. By our assumptions at the beginning of the proof, this simplifies to $|z| < \epsilon$. For $v_i = z - rx_i$ we have $v_i \in \partial_x f(x_i, \lambda)$, and by the triangle inequality $|v_i| \leq |z| + r|x_i| < \frac{\bar{\epsilon}}{2} + \frac{\bar{\epsilon}}{2} = \bar{\epsilon}$. Also $|x_i| < \bar{\epsilon}$ and $|f(x_i, \lambda)| < \bar{\epsilon}$, thus, $v_i \in T_\lambda(x_i)$. So we have that

$$\begin{aligned} -\bar{r}|x_1 - x_0|^2 &\leq \langle v_1 - v_0, x_1 - x_0 \rangle \\ &= \langle (z - rx_1) - (z - rx_0), x_1 - x_0 \rangle = -r|x_1 - x_0|^2. \end{aligned}$$

Since $r > \bar{r}$, we conclude that $x_1 = x_0$.

c) $\Rightarrow$ *a)*: Let $\epsilon > 0$ and $r > 0$ be the parameters assumed by condition *c)*. It will suffice to show that there exists $\bar{\epsilon} \in (0, \epsilon)$ such that para-prox-regularity of f is satisfied by $\bar{\epsilon}$ and r at $\bar{x} = 0$, $f(0, \bar{\lambda}) = 0$ for $\bar{v} = 0$. That is, whenever

$$v_0 \in \partial_x f(x_0, \lambda), \quad |v_0| < \bar{\epsilon}, \quad |f(x_0, \lambda)| < \bar{\epsilon}, \tag{25}$$

we have

$$f(x, \lambda) \geq f(x_0, \lambda) + \langle v_0, x - x_0 \rangle - \frac{r}{2}|x - x_0|^2$$

when $|x| < \bar{\epsilon}$ and $|\lambda - \bar{\lambda}| < \bar{\epsilon}$. In place of the latter we can aim at guaranteeing the stronger condition

$$f(x, \lambda) > f(x_0, \lambda) + \langle v_0, x - x_0 \rangle - \frac{r}{2}|x - x_0|^2 \tag{26}$$

for $x \neq x_0$, $|x| \leq \epsilon$, and $|\lambda - \bar{\lambda}| \leq \epsilon$. Notice that

$$-\frac{r}{2}|x - x_0|^2 = \langle rx_0, x - x_0\rangle - \frac{r}{2}|x|^2 + \frac{r}{2}|x_0|^2,$$

so we can rewrite inequality (26) as

$$\begin{aligned} f(x,\lambda) &> f(x_0,\lambda) + \langle v_0, x - x_0\rangle + \langle rx_0, x - x_0\rangle - \frac{r}{2}|x|^2 + \frac{r}{2}|x_0|^2 \\ &= f(x_0,\lambda) + \langle v_0 + rx_0, x\rangle - \langle v_0 + rx_0, x_0\rangle - \frac{r}{2}|x|^2 + \frac{r}{2}|x_0|^2. \end{aligned}$$

Rearranging, we get

$$f(x,\lambda) - \langle v_0 + rx_0, x\rangle + \frac{r}{2}|x|^2 > f(x_0,\lambda) - \langle v_0 + rx_0, x_0\rangle + \frac{r}{2}|x_0|^2.$$

Thus, we seek to show that for every fixed λ such that $|\lambda - \bar{\lambda}| \leq \bar{\epsilon}$,

$$\operatorname{argmin}_{|x|\leq\bar{\epsilon}} \left\{ f(x,\lambda) + \frac{r}{2}|x|^2 - \langle z_0, x\rangle \right\} = \{x_0\}, \tag{27}$$

where $z_0 = v_0 + rx_0$. In summary, we need only demonstrate that conditions (25) imply equation (27) when $\bar{\epsilon}$ is small enough. Our assumption that f is locally lsc at $(0,0)$ with $f(0,0) = 0$ ensures that f is lsc with respect to a compact set of the form

$$C = \{(x,\lambda) : |x| \leq \hat{\epsilon}, |\lambda| \leq \hat{\epsilon}, f(x,\lambda) \leq 2\hat{\epsilon}\}$$

for some $\hat{\epsilon} > 0$. In particular, f must be lsc at $(0,0)$ itself: $\liminf_{x,\lambda\to 0} f(x,\lambda) = 0$. Shrinking ϵ if necessary, we can arrange that

$$|x| \leq \epsilon, \quad |\lambda| \leq \epsilon \Rightarrow |x| \leq \hat{\epsilon}, \ \|\lambda| \leq \hat{\epsilon}, \quad \text{and} \quad f(x,\lambda) > -2\hat{\epsilon}. \tag{28}$$

Enlarging r if necessary, we can arrange that

$$|x|^2 < \frac{6\hat{\epsilon}}{r} \Rightarrow |x| < \epsilon, \quad |x| < \frac{\epsilon^2}{2\hat{\epsilon}}, \quad \text{and} \quad f(x,\lambda) > -\frac{\epsilon}{2}. \tag{29}$$

Define

$$g(z) := \inf_{|x|,|\lambda-\bar{\lambda}|\leq\epsilon} \left\{ f(x,\lambda) + \frac{r}{2}|x|^2 - \langle z, x\rangle \right\},$$

$$G(z) := \operatorname{argmin}_{|x|,|\lambda-\bar{\lambda}|\leq\epsilon} \left\{ f(x,\lambda) + \frac{r}{2}|x|^2 - \langle z, x\rangle \right\}.$$

Since g is the pointwise infimum of a collection of affine functions of z, it is concave. (Note: $\bar{x} = 0$, $\bar{v} = 0$, and $f(\bar{x},\bar{\lambda}) = 0$.) Our assumptions tell us that $g(0) = 0$ and $G(0) = \{0\}$, whereas $g(z) \leq 0$ in general. We claim that under these circumstances,

$$|z| < \frac{\hat{\epsilon}}{\epsilon} \Rightarrow \begin{cases} G(z) \neq \emptyset \text{ and} \\ |f(x,\lambda)| < 2\hat{\epsilon} \quad \text{for all } (x,\lambda) \in G(z). \end{cases} \tag{30}$$

To see this, observe that because $g \leq 0$, the minimization in the definition of $g(z)$ is unaffected if the attention is restricted to the points (x, λ) satisfying not only $|x| < \epsilon$ and $|\lambda| < \epsilon$, but also $f(x, \lambda) + \frac{r}{2}|x|^2 - \langle z, x\rangle \leq \hat{\epsilon}$, in which case $f(x, \lambda) \leq \hat{\epsilon} + \epsilon|z|$. Therefore, as long as $|z| < \frac{\hat{\epsilon}}{\epsilon}$, attention can be restricted to points (x, λ) satisfying $f(x, \lambda) < 2\hat{\epsilon}$. Recalling conditions (28) and the choice of the set C, we deduce that when $|z| \leq \frac{\hat{\epsilon}}{\epsilon}$,

$$\begin{aligned} g(z) &= \inf_{(x,\lambda)\in C}\{f(x, \lambda) + \tfrac{r}{2}|x|^2 - \langle z, x\rangle\} \text{ and} \\ G(z) &= \operatorname{argmin}_{(x,\lambda)\in C}\{f(x, \lambda) + \tfrac{r}{2}|x|^2 - \langle z, x\rangle\}. \end{aligned} \tag{31}$$

Since f is lsc relative to C, which is compact, the infimum is sure to be attained. Thus, implication (30) is correct. Also we observe that g is finite on a neighborhood of $z = 0$, and hence continuous (by concavity) [21, Theorem 2.35]. Choose $\tilde{\epsilon} \in (0, \epsilon)$ small enough that $\tilde{\epsilon}(1+r) < \frac{\hat{\epsilon}}{\epsilon}$ and $g(z) > -\frac{\epsilon}{2}$ when $|z| < \tilde{\epsilon}(1+r)$. Under this choice we are ready to consider elements satisfying conditions (25) and show that we get equation (27). The vector $z_0 = v_0 + rx_0$ has $|z_0| \leq |v_0| + r|x_0| < \tilde{\epsilon}(1+r)$, hence $|z_0| < \frac{\hat{\epsilon}}{\epsilon}$ and $g(z_0) > -\frac{\epsilon}{2}$. In particular, $G(z_0)$ must be nonempty. Consider any $(x_1, \lambda) \in G(z_0)$. We have to show on the basis of the single-valuedness property that $x_1 = x_0$. We have that $(x_0, \lambda) \in (\partial_x f + rI)^{-1}(z_0)$, due to the fact that $v_0 \in \partial_x f(x_0, \lambda)$ and hence $v_0 + rx_0 \in (\partial_x f + rI_x)(x_0, \lambda)$. If we can establish that

$$|x_1| < \epsilon, \quad |\lambda| < \epsilon, \quad |f(x_1, \lambda)| < \epsilon \quad \text{and} \quad (x_1, \lambda) \in (\partial_f + rI)^{-1}(z_0), \tag{32}$$

then the single-valuedness property can be invoked and we will get $x_1 = x_0$ as required.

Because $|z_0| < \frac{\hat{\epsilon}}{\epsilon}$, we know from implication (30) that $|f(x_1, \lambda)| < 2\hat{\epsilon}$. Then from $\frac{r}{2}|x_1|^2 = \langle z_0, x_1\rangle - f(x_1, \lambda) + g(z_0)$, where $g(z_0) \leq 0$, we know that $\frac{r}{2}|x_1|^2 \leq |z_0||x_1| + |f(x_1, \lambda)| < \frac{\hat{\epsilon}}{\epsilon}\epsilon + 2\hat{\epsilon} = 3\hat{\epsilon}$, so $|x_1|^2 < \frac{6\hat{\epsilon}}{r}$. Through (29) this implies that $|x_1| < \epsilon$ and also that $|x_1| < \frac{\epsilon^2}{2\hat{\epsilon}}$ and $f(x_1, \lambda) > -\epsilon$. Then from $f(x_1, \lambda) = \langle z_0, x_1\rangle - \frac{r}{2}|x_1|^2 + g(z_0) \leq |z_0||x_1| + |g(z_0)|$ and the fact that $g(z_0) < \frac{\epsilon}{2}$ we obtain $f(x_1) < \frac{\hat{\epsilon}}{\epsilon}\frac{\epsilon^2}{2\hat{\epsilon}} + \frac{\epsilon}{2} = \epsilon$. Hence $|x_1| < \epsilon$, $|\lambda| < \epsilon$, and $|f(x_1, \lambda)| < \epsilon$. The fact that the minimum for $g(z_0)$ is attained at x_1 gives us

$$f(x, \lambda) + \frac{r}{2}|x|^2 - \langle z_0, x\rangle \geq f(x_1, \lambda) + \frac{r}{2}|x_1|^2 - \langle z_0, x_1\rangle \tag{33}$$

when $|x| < \epsilon$ and $|\lambda| < \epsilon$. Hence, for all $|x| < \epsilon$ and $|\lambda| < \epsilon$,

$$\begin{aligned} & f(x, \lambda) \\ &\geq f(x_1, \lambda) + \langle z_0, x\rangle - \langle z_0, x_1\rangle - \frac{r}{2}|x|^2 + \frac{r}{2}|x_1|^2 \\ &= f(x_1, \lambda) + \langle z_0, x - x_1\rangle - r\langle x, x_1\rangle + r\langle x, x_1\rangle + r\langle x_1, x_1\rangle - \frac{r}{2}\langle x, x\rangle - \frac{r}{2}\langle x_1, x_1\rangle \\ &= f(x_1, \lambda) + \langle z_0, x - x_1\rangle - \langle rx_1, x - x_1\rangle - \frac{r}{2}(\langle x, x\rangle - 2\langle x, x_1\rangle + \langle x_1, x_1\rangle) \\ &= f(x_1, \lambda) + \langle z_0 - rx_1, x - x_1\rangle - \frac{r}{2}|x - x_1|^2. \end{aligned}$$

Since $|x_1| < \epsilon$, this implies that $z_0 - rx_1 \in \partial_x f(x_1, \lambda)$, so $z_0 \in (\partial_x f + rI_x)(x_1, \lambda)$. The subgradient relation in (32) is therefore correct as well, and our single-valuedness assumption implies $x_0 = x_1$. Thus equation (27) holds and we are done. □

This result allows us to define the following alternate definition of para-prox-regularity.

Corollary 4.4. *Let f be a proper, lsc function. Then f is para-prox-regular in x at $\bar{x}$ for $\bar{v}$ with compatible parametrization by λ at $\bar{\lambda}$, with parameters $\epsilon > 0, r > 0$, if and only if $\bar{v}$ is a proximal subgradient of f with respect to x at $(\bar{x}, \bar{\lambda})$ and*

$$\langle v_1 - v_0, x_1 - x_0 \rangle \geq r|x_1 - x_0|^2$$

when for $i \in \{0, 1\}$ we have $|x_i - \bar{x}| < \epsilon$, $|f(x_i, \lambda) - f(\bar{x}, \bar{\lambda})| < \epsilon$, $v_i \in \partial_x f(x_i, \lambda)$, $|v_i - \bar{v}| < \epsilon$, and $|\lambda - \bar{\lambda}| < \epsilon$.

5. Advanced Examples

In this section we examine some common situations where parameters arise in optimization problems. In all cases we focus on identifying the prox-regularity parameters. We already know of many classes of functions that are prox-regular, some were mentioned in the introduction, and here in general we begin with prox-regular functions and build upon them. We show that the scalar multiplication of a prox-regular function by a positive parameter results in a function that remains prox-regular. This will lead to the construction of para-prox-regular functions via the weighted sums and weighted max of prox-regular functions.

The following proposition and theorem is a result first provided in [20], showing that strongly amenable functions are prox-regular. In that work, Poliquin and Rockafellar simply conclude the function is prox-regular and they do not single out the prox-regularity parameters ϵ and r in the statement of the theorem. In the present work, the theorem has been restated with a proposition beforehand, in order to pull out the prox-regularity parameters from the proof into the statement of the theorem. In later results we require the ability to identify such parameters instead of merely stating that they exist. Other than clarifying the prox-regularity parameters, Proposition 5.1 and Theorem 5.2 and their proofs are essentially the same as [20, Theorem 3.1].

Proposition 5.1. *Let $f(x) = g(F(x))$, where $F : \mathbb{R}^n \to \mathbb{R}^m$ is of class $\mathcal{C}^2$, $g : \mathbb{R}^m \to \bar{\mathbb{R}}$ is lsc and proper, and suppose there is no vector $y \neq 0$ in $\partial^\infty g(F(\bar{x}))$ with $\nabla F(\bar{x})^* y = 0$, where $\bar{x} \in \operatorname{dom} f$. Then*

a) there exists $\epsilon_0 > 0$ such that

$$\partial f(x) \subset \nabla F(x)^* \partial g(F(x))$$

when $|x - \bar{x}| < \epsilon_0$ and $|f(x) - f(\bar{x})| < \epsilon_0$.

b) $S : (x, v) \to \{y \in \partial g(F(x)) : \nabla F(x)^* y = v\}$ *has closed graph and is locally bounded at* $(\bar{x}, \bar{v})$ *for* $\bar{v} \in \partial f(\bar{x})$*, within the same neighborhoods* $B_{\epsilon_0}(\bar{x})$ *and* $B_{\epsilon_0}(f(\bar{x}))$*. In particular,* $S(\bar{x}, \bar{v})$ *is a compact set.*

c) *If* $\bar{v} \in \partial f(\bar{x})$ *is such that for every* $y \in \partial g(F(\bar{x}))$ *with* $\nabla F(\bar{x})^* y = \bar{v}$ *the function* g *is prox-regular at* $F(\bar{x})$ *with parameters* $\epsilon_y > 0, r_y > 0$*, then for the parameters* $\epsilon = \min\{\epsilon_0, \min_y\{\epsilon_y\}\}$ *and* $\bar{r} = \max_y\{r_y\}$*, which are guaranteed to exist since* $S(\bar{x}, \bar{v})$ *is a compact set, we have*

$$\langle y_1 - y_0, u_1 - u_0\rangle \geq -\bar{r}|u_1 - u_0|^2 \tag{34}$$

whenever $y_i \in \partial g(u_i)$, $|u_i - F(\bar{x})| < \bar{\epsilon}$, $|g(u_i) - g(F(\bar{x}))| < \bar{\epsilon}$, $\operatorname{dist}(y_i, S(\bar{x}, \bar{v})) < \bar{\epsilon}$.

Proof. See [20, Theorem 3.1] and the proof thereof.

In Theorem 5.2 we must provide greater detail, in order to fully describe the prox-regularity parameters. In particular, the full details of parameter r only become apparent with the proof itself.

Theorem 5.2. *Let the conditions of Proposition* 5.1 *hold. Assume further that* $\bar{v} \in \partial f(\bar{x})$ *is a vector such that for every* $y \in \partial g(F(\bar{x}))$ *such that* $\nabla F(\bar{x})^* y = \bar{v}$*, the function* g *is prox-regular at* $F(\bar{x})$ *for* y *with parameters* $\epsilon_y > 0, r_y > 0$*. Then* f *is prox-regular at* $\bar{x}$ *for* $\bar{v}$ *with parameters* $\epsilon > 0$ *as in Proposition* 5.1 *and* $r > 0$ *as described by equations* (35)*,* (37)*,* (39)*, and* (41) *below.*

Proof. By our assumptions, g is prox-regular at $F(\bar{x})$ for each $y \in S(\bar{x}, \bar{v})$, with constants ϵ_y and r_y. Since $S(\bar{x}, \bar{v})$ is compact we can make this uniform. We know that there is a finite open cover of $S(\bar{x}, \bar{v})$ using a finite number of the ϵ_y, say $\{\epsilon_1, \ldots, \epsilon_k\}$. Then using the corresponding $\{r_1, \ldots, r_k\}$, we define

$$\bar{r} := \max_{i \in \{1,\ldots,k\}} \{r_i\} > 0. \tag{35}$$

We now have $\epsilon > 0$, $\bar{r} > 0$, $Y = S(\bar{x}, \bar{v})$ a compact set, and $y_i \in \partial g(F(x_i)) \cap Y$ with $\nabla F(x_i)^* y_i = v_i$ such that

$$\langle y_1 - y_0, F(x_1) - F(x_0)\rangle \geq -\bar{r}|F(x_1) - F(x_0)|^2$$

when $|x_i - \bar{x}| < \epsilon$, $|f(x_i) - f(\bar{x})| < \epsilon$, $v_i \in \partial f(x_i)$, and $|v_i - \bar{v}| < \epsilon$. Note that when $v_i = \nabla F(x_i)^* y_i$ then

$$\begin{aligned} &\langle v_1 - v_0, x_1 - x_0\rangle \\ &= \langle \nabla F(x_1)^* y_i - \nabla F(x_0)^* y_1 + \nabla F(x_0)^* y_1 - \nabla F(x_0)^* y_0, x_1 - x_0\rangle, \end{aligned}$$

which gives

$$\begin{aligned} &\langle v_1 - v_0, x_1 - x_0\rangle \\ &= \langle [\nabla F(x_1)^* - \nabla F(x_0)^*] y_1, x_1 - x_0\rangle + \langle \nabla F(x_0)^*(y_1 - y_0), x_1 - x_0\rangle. \end{aligned} \tag{36}$$

Because Y is compact and $F \in \mathcal{C}^2$, we can find

$$r_1 > 0 \text{ such that for all } y \in Y, \quad |[\nabla F(x_1)^* - \nabla F(x_0)^*]y| \leq r_1|x_1 - x_0|. \tag{37}$$

Thus for all $y \in Y$,

$$\langle [\nabla F(x_1)^* - \nabla F(x_0)^*]y, x_1 - x_0 \rangle \geq -r_1|x_1 - x_0|^2. \tag{38}$$

The final term in equation (36) can be written as $\langle \nabla\phi(x_0), x_1 - x_0\rangle$ for the $\mathcal{C}^2$ function $\phi(x) = \langle y_1 - y_0, F(x)\rangle$. Let

$$\begin{array}{l} r_2 \text{ be an upper bound for the eigenvalues} \\ \text{of the Hessian of the mapping } x \to \langle \eta, F(x)\rangle, \end{array} \tag{39}$$

where x ranges over $B_\epsilon(\bar{x})$ and η ranges over the compact set $Y - Y$. Then in particular $\phi(x) \leq \phi(x_0) + \langle \nabla\phi(x_0), x - x_0\rangle + r_2|x - x_0|^2$ when $|x - \bar{x}| \leq \epsilon$. It follows that

$$\begin{aligned} \langle \nabla\phi(x_0), x_1 - x_0\rangle &\geq -r_2|x_1 - x_0|^2 + \phi(x_1) - \phi(x_0) \\ &= -r_2|x_1 - x_0|^2 + \langle y_1 - y_0, F(x_1) - F(x_0)\rangle \\ &\geq -r_2|x_1 - x_0|^2 - \bar{r}|F(x_1) - F(x_0)|^2. \end{aligned} \tag{40}$$

But there is also a constant k such that $|F(x_1) - F(x_0)| \leq k|x_1 - x_0|$ when $|x_i - \bar{x}| < \epsilon$. Using this fact with the inequalities (38) and (40) for the two terms at the end of equation (36) we obtain

$$\langle v_1 - v_0, x_1 - x_0\rangle \geq -r_1|x_1 - x_0|^2 - r_2|x_1 - x_0|^2 - \bar{r}k^2|x_1 - x_0|^2.$$

Thus for

$$r := r_1 + r_2 + \bar{r}k^2 \tag{41}$$

we have that

$$\langle v_1 - v_0, x_1 - x_0\rangle \geq -r|x_1 - x_0|^2$$

whenever $|x_i - \bar{x}| < \epsilon$, $|f(x_i) - f(\bar{x})| < \epsilon$, $v_i \in \partial f(x_i)$, and $|v_i - \bar{v}| < \epsilon$. □

Now we are ready to begin with parametrized functions. Lemma 5.3 demonstrates that multiplying a prox-regular function by a positive parameter does not affect its prox-regularity, a property that we shall use in the theorems that follow.

Lemma 5.3. *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be prox-regular at $\bar{x}$ for $\bar{v}$ with parameters $\tilde{\epsilon} > 0$ and $\tilde{r} > 0$, and let $\lambda \in \mathbb{R}, \lambda > 0$. Then $\lambda f(x)$ is prox-regular at $\bar{x}$ for $\lambda\bar{v}$ with parameters $\epsilon = \min\{\tilde{\epsilon}, \lambda\tilde{\epsilon}\}$ and $r = \lambda\tilde{r}$.*

Proof. By prox-regularity of f, $\tilde{\epsilon}$ and $\tilde{r}$ are such that

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

whenever $x' \neq x$, $|x' - \bar{x}| < \tilde{\epsilon}$, $|x - \bar{x}| < \tilde{\epsilon}$, $|f(x) - f(\bar{x})| < \tilde{\epsilon}$, $v \in \partial f(x)$, and $|v - \bar{v}| < \tilde{\epsilon}$. Multiplying by $\lambda > 0$ yields

$$\lambda f(x') > \lambda f(x) + \langle \lambda v, x' - x \rangle - \frac{\lambda \tilde{r}}{2}|x' - x|^2$$

whenever $x' \neq x$, $|x' - \bar{x}| < \tilde{\epsilon}$, $|x - \bar{x}| < \tilde{\epsilon}$, $|f(x) - f(\bar{x})| < \tilde{\epsilon}$, $v \in \partial f(x)$, and $|v - \bar{v}| < \tilde{\epsilon}$. Since $v \in \partial f(x)$, we know that $\lambda v \in \partial \lambda f(x)$. When $|v - \bar{v}| < \tilde{\epsilon}$ and $|f(x) - f(\bar{x})| < \tilde{\epsilon}$ we have $|\lambda v - \lambda \bar{v}| < \lambda \tilde{\epsilon}$ and $|\lambda f(x) - \lambda f(\bar{x})| < \lambda \tilde{\epsilon}$. Let $\epsilon = \min\{\tilde{\epsilon}, \lambda \tilde{\epsilon}\}$, and let $r = \lambda \tilde{r}$. We conclude that

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

whenever $x' \neq x$, $|x' - \bar{x}| < \epsilon$, $|x - \bar{x}| < \epsilon$, $|\lambda f(x) - \lambda f(\bar{x})| < \epsilon$, $\lambda v \in \partial f(x)$, and $|\lambda v - \lambda \bar{v}| < \epsilon$. □

A direct result of Lemma 5.3 is that the scalar multiplication of a prox-regular function with a positive parameter results in a para-prox-regular function.

Corollary 5.4. *Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be prox-regular at $\bar{x}$ for $\bar{v}$ with parameters $\tilde{\epsilon} > 0$ and $\tilde{r} > 0$, and let $\bar{\lambda} \in \mathbb{R}$, $\bar{\lambda} > 0$. Then $g(x, \lambda) = \lambda f(x)$ is para-prox-regular at $\bar{x}$ for $\bar{\lambda}\bar{v}$ with parameters $\epsilon = \min\{\tilde{\epsilon}, \frac{\bar{\lambda}\tilde{\epsilon}}{2}\}$ and $r = \frac{3\bar{\lambda}\tilde{r}}{2}$ with compatible parametrization by λ at $\bar{\lambda}$.*

Proof. Let $\delta = \frac{\bar{\lambda}}{2}$ and consider $\lambda \in (\bar{\lambda} - \delta, \bar{\lambda} + \delta)$. Then for each such λ, using $\epsilon_\lambda = \min\{\tilde{\epsilon}, \delta\tilde{\epsilon}\}$ we have that

$$g(x', \lambda) > g(x, \lambda) + \langle \lambda v, x' - x \rangle - \frac{\lambda \tilde{r}}{2}|x' - x|^2$$

whenever $x' \neq x$, $|x' - \bar{x}| < \epsilon_\lambda$, $|x - \bar{x}| < \epsilon_\lambda$, $|g(x, \bar{\lambda}) - g(\bar{x}, \bar{\lambda})| < \epsilon_\lambda$, $\lambda v \in \partial_x g(x, \lambda)$, and $|\lambda v - \bar{\lambda}\bar{v}| < \epsilon_\lambda$. Let $\epsilon = \min\{\delta, \tilde{\epsilon}, \delta\tilde{\epsilon}\}$ and $r = \max\{\lambda \tilde{r}\}$. Then we have

$$g(x', \lambda) > g(x, \lambda) + \langle \lambda v, x' - x \rangle - \frac{r}{2}|x' - x|^2$$

when $x' \neq x$, $|x' - \bar{x}| < \epsilon$, $|x - \bar{x}| < \epsilon$, $|g(x, \lambda) - g(\bar{x}, \bar{\lambda})| < \epsilon$, $\lambda v \in \partial_x g(x, \lambda)$, $|\lambda v - \bar{\lambda} v| > \epsilon$, and $|\lambda - \bar{\lambda}| < \epsilon$. □

Next we have a minor extension of [20, Theorem 3.2], where Poliquin and Rockafellar showed that, under certain conditions, the sum of two prox-regular functions is again prox-regular. We extend this to the sum of any finite number of prox-regular functions. The result also extracts the prox-regularity parameters from the proof, thus allowing their use in later results.

Theorem 5.5. *For $i \in \{1, 2, \ldots, m\}$, let $f_i : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$. Let $\bar{x} \in \bigcap_{i=1}^m \operatorname{dom} f_i$ and assume that the only choice of $v_i \in \partial f_i^\infty(\bar{x})$ with $\sum_{i=1}^m v_i = 0$ is $v_i = 0$ for all $i = 1, 2, \ldots, m$. Define $f(x) := \sum_{i=1}^m f_i(x)$ and let $\bar{v} \in \partial f(\bar{x})$. If f_i is prox-regular at $\bar{x}$ for v_i with parameters $\epsilon_i > 0$ and $r_i > 0$, for each $v_i \in \partial f_i(\bar{x})$ such that $\sum_{i=1}^m v_i = \bar{v}$, then f is prox-regular at $\bar{x}$ for $\bar{v}$ with parameters $\epsilon = \min_i\{\epsilon_i\}$ and $r = m \max_i\{r_i\}$.*

Proof. Define $g(u_1, u_2, \ldots, u_m) := \sum_{i=1}^m f_i(u_i)$ and $F(x) := (x, x, \ldots, x)$. Then g is lsc and proper, $F \in \mathcal{C}^2$, and $f(x) = g(F(x))$. Notice that $\nabla F(x) = (I, I, \ldots, I)^T$, so the only possible $y \in \partial^\infty g(F(\bar{x}))$ with $\nabla F(\bar{x})^* y = 0$ is $y = 0$. Therefore the constraint qualification on y in Theorem 5.2 holds. We have that g is prox-regular at $F(\bar{x})$ for every $y \in \partial g(F(\bar{x}))$ with $\nabla F(\bar{x})^* y = \bar{v}$, so all the conditions of Theorem 5.2 are satisfied. Let ϵ and r be as described in the proof of that theorem. In particular, $r_1 = 0$ in this case since $\nabla F(x_1) - \nabla F(x_0) = 0$. Also $r_2 = 0$, since $\langle \eta, F(x)\rangle = \langle \eta, (x, x, \ldots, x)\rangle = x\eta$, which implies $\nabla\langle \eta, F(x)\rangle = \eta$ and $\nabla^2\langle \eta, F(x)\rangle = 0$. Finally,

$$\begin{aligned} |F(x_1) - F(x_0)| &= |(x_1 - x_0, x_1 - x_0, \ldots, x_1 - x_0)| \\ &= \sqrt{m(x_1 - x_0)^2} = \sqrt{m}|x_1 - x_0|, \end{aligned}$$

which gives $\lambda = \sqrt{n}$. Hence $r = m\bar{r} = m \max_i\{r_i\}$. We conclude that f is prox-regular at $\bar{x}$ for $\bar{v}$ with these parameters. □

Corollary 5.6. *For $i \in \{1, 2, \ldots, m\}$, let $f_i : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be prox-regular at $\bar{x}$ for $v_i \in \partial f_i(\bar{x})$ with parameters $\epsilon_i > 0$ and $r_i > 0$. Let $\bar{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_m) > 0$ (i.e. $\lambda_i > 0$ for all i), and define $f(x) := \sum_{i=1}^m \lambda_i f_i(x)$. Then f is prox-regular at $\bar{x}$ for $\bar{\lambda}\bar{v}$ with parameters $\epsilon = \min_i\{\epsilon_i, \lambda_i\epsilon_i\}$ and $r = m \max_i\{\lambda_i r_i\}$, where $\bar{v} = \sum_{i=1}^m \lambda_i v_i \in \partial_x f(\bar{x}, \bar{\lambda})$.*

Proof. Apply Lemma 5.3 to each f_i in Theorem 5.5. □

We now examine parametrized functions by multiplying each of the prox-regular functions f_i in Theorem 5.5 by a parameter λ_i, and we show that the weighted sum of prox-regular functions is a para-prox-regular function.

Theorem 5.7. *For $i \in \{1, 2, \ldots, m\}$, let $f_i : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$. Let $\bar{\lambda} = (\bar{\lambda}_1, \bar{\lambda}_2, \ldots, \bar{\lambda}_m) > 0$. Let $\bar{x} \in \bigcap_{i=1}^m \operatorname{dom} f_i$ and assume that*

$$\text{the only choice of } v_i \in \partial f_i^\infty(\bar{x}) \text{ with } \sum_{i=1}^m v_i = 0 \text{ is } v_i = 0 \text{ for all } i. \tag{42}$$

Define

$$f(x, \lambda) := \sum_{i=1}^m \lambda_i f_i(x)$$

where $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_m)$, and let $\bar{v} \in \partial_x f(\bar{x}, \bar{\lambda})$. If f_i is prox-regular at $\bar{x}$ for all $\bar{v}_i \in \partial f_i(\bar{x})$ such that $\sum_{i=1}^m \bar{\lambda}_i \bar{v}_i = \bar{v}$, with parameters ϵ_i and r_i, then f is para-prox-regular at $\bar{x}$ for $\bar{v}$ with compatible parametrization by λ at $\bar{\lambda}$, with parameters $\epsilon = \min_i\{\frac{\bar{\lambda}_i}{2}, \epsilon_i, \frac{\bar{\lambda}_i}{2}\epsilon_i\}$ and $r = m \max_i\{\frac{3\bar{\lambda}_i}{2} r_i\}$.

Proof. Since $\bar{\lambda}_i > 0$ and f_i is prox-regular at $\bar{x}$ for $\bar{v}_i$ for all i, Lemma 5.3 applies and we have that each $\lambda_i f_i$ is prox-regular at $\bar{x}$ for $\bar{\lambda}_i\bar{v}_i$ with parameters $\tilde{\epsilon}_i = \min\{\epsilon_i, \bar{\lambda}_i\epsilon_i\}$ and $\tilde{r}_i = \bar{\lambda}_i r_i$. Since $\bar{\lambda}_i > 0$, we have $\partial^\infty \bar{\lambda}_i f_i = \partial^\infty f_i$ [21,

equation 10(6)], so condition (42) holds for $v_i \in \partial^\infty \bar{\lambda}_i f_i$. Thus by Corollary 5.6, we have that $f(x,\bar{\lambda})$ is prox-regular in x at $\bar{x}$ for $\bar{v}$ with parameters $\bar{\epsilon} = \min_i\{\tilde{\epsilon}_i\}$ and $\bar{r} = m\max_i\{\tilde{r}_i\}$. (Note that here we are viewing f as a function of x, with $\bar{\lambda}$ fixed.) Define $\delta := \frac{1}{2}\min_i\{\bar{\lambda}_i\}$, and $\mathcal{B} := \{\lambda : \lambda_i \in [\bar{\lambda}_i - \delta, \bar{\lambda}_i + \delta] \text{ for all } i\}$. Since $\lambda > 0$ for all $\lambda \in \mathcal{B}$, by similar arguments as above we have that $f(x,\lambda)$ is prox-regular in x at $\bar{x}$ for $\bar{v}$ for every fixed $\lambda \in \mathcal{B}$, with parameters $\epsilon_\lambda = \min_i\{\epsilon_i, \lambda_i\epsilon_i\}$ and $r_\lambda = m\max_i\{\lambda_i r_i\}$. But $\min_{\lambda\in\mathcal{B}}\{\lambda_i\} = \delta$ and $\max_{\lambda\in\mathcal{B}}\{\lambda_i\} = \max_i\{\bar{\lambda}_i\} + \delta$, so for all $\lambda \in \mathcal{B}$ we have that $f(x,\lambda)$ is prox-regular in x at $\bar{x}$ for $\bar{v}$ with parameters $\hat{\epsilon} = \min\{\epsilon_i, \delta\epsilon_i\}$ and $\hat{r} = m\max_i\{(\bar{\lambda}_i + \delta)r_i\}$. Since $\bar{\lambda}_i + \delta \leq \frac{3\bar{\lambda}_i}{2}$, we may overestimate and set $r = m\max_i\{\frac{3\bar{\lambda}_i}{2}r_i\}$. Therefore we may conclude that f is para-prox-regular at $(\bar{x},\bar{\lambda})$ for $\bar{v}$ with parameters $\epsilon = \min\{\delta,\hat{\epsilon}\} = \min_i\{\frac{\bar{\lambda}_i}{2}, \epsilon_i, \frac{\bar{\lambda}_i}{2}\epsilon_i\}$ and $r = m\max_i\{\frac{3\bar{\lambda}_i}{2}r_i\}$. □

Example 5.8. Let $f_1 : \mathbb{R}^n \to \mathbb{R}\cup\{\infty\}$ and $f_2 : \mathbb{R}^n \to \mathbb{R}\cup\{\infty\}$ be prox-regular at $\bar{x} \in \bigcap_{i=1}^2 \operatorname{dom} f_i$. Assume the only choice of $\hat{v}_1 \in \partial f_1^\infty(\bar{x})$ and $\hat{v}_1 \in \partial f_2^\infty(\bar{x})$ with $\hat{v}_1 + \hat{v}_2 = 0$ is $\hat{v}_1 = \hat{v}_2 = 0$. Define

$$f(x,\lambda) := \lambda f_1(x) + (1-\lambda)f_2(x).$$

Let $\bar{v} \in \partial_x f(\bar{x},\bar{\lambda})$ such that $\bar{v} = \lambda v_1 + (1-\lambda)v_2$ with $v_1 \in \partial f_1(\bar{x})$ and $v_2 \in \partial f_2(\bar{x})$. Then f is para-prox-regular at $\bar{x}$ for $\bar{v} = \lambda v_1 + (1-\lambda)v_2$ with compatible parametrization by λ at any $\bar{\lambda} \in (0,1)$.

Details: For f_1 and f_2 lower-$\mathcal{C}^2$ functions, this result has already been proved [10]. The lower-$\mathcal{C}^2$ requirement can be generalized to prox-regular by using Theorem 5.7. □

The following proposition and lemma are a restatement of a result from [14]. We shall use them in the proof of Theorem 5.11, para-prox-regularity of a weighted finite max of $\mathcal{C}^0$ functions.

Proposition 5.9 ([14, Prop. 2.2]). *Suppose that $f : \mathbb{R}^n \times \mathbb{R}^m$ is strongly amenable in x at $\bar{x}$ with compatible parametrization by λ at $\bar{\lambda}$, in the sense that on some neighborhood of $(\bar{x},\bar{\lambda})$ there is a composite representation $f(x,\lambda) = g(F(x,\lambda))$ in which $F : \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}^m$ is a $\mathcal{C}^2$ mapping and $g : \mathbb{R}^m \to \mathbb{R}$ is a convex, proper, lsc function for which $F(\bar{x},\bar{\lambda}) \in D := \operatorname{dom} g$ and*

$$z \in N_D(F(\bar{x},\bar{\lambda})), \nabla_x F(\bar{x},\bar{\lambda})^* z = 0 \Rightarrow z = 0.$$

Then as long as $\bar{v} \in \partial_x f(\bar{x},\bar{\lambda})$, one has f continuously para-prox-regular in x at $\bar{x}$ for $\bar{v}$ with compatible parametrization by λ at $\bar{\lambda}$.

Corollary 5.10. *For $i \in \{1,2,\ldots,m\}$, let $f_i : \mathbb{R}^n \to \mathbb{R}$ be $\mathcal{C}^2$. Let $\bar{\lambda} = (\bar{\lambda}_1,\bar{\lambda}_2,\ldots,\bar{\lambda}_m) > 0$. Let $\bar{x} \in \bigcap_{i=1}^m \operatorname{dom} f_i$ and $\bar{v} \in \partial_x f(\bar{x},\bar{\lambda})$. Define*

$$f(x,\lambda) := \max_i\{\lambda_i f_i(x)\}$$

where $\lambda = (\lambda_1,\lambda_2,\ldots,\lambda_m)$. Then f is continuously para-prox-regular in x at $\bar{x}$ for $\bar{v}$ with compatible parametrization by λ at $\bar{\lambda}$.

Proof. Let

$$F(x,\lambda) = (\lambda_1 f_1(x), \lambda_2 f_2(x), \ldots, \lambda_m f_m(x))$$

and

$$g(u_1, u_2, \ldots, u_m) = \max\{u_1, u_2, \ldots, u_m\}.$$

Then F is $\mathcal{C}^2$, since each component is a product of two $\mathcal{C}^2$ functions, g is proper lsc convex, and $f(x,\lambda) = g(F(x))$.
Therefore f is strongly amenable. Since $D = \operatorname{dom} g$ is $\mathbb{R}^n$, $N_D(F(\bar{x},\bar{\lambda})) = \{0\}$, so the constraint qualification of Proposition 5.9 holds. We have all the requirements of Proposition 5.9, and its conclusion is our desired result. □

Our next result is a relaxation of the $\mathcal{C}^2$ condition of Corollary 5.10. It is sufficient that the f_i be $\mathcal{C}^1$, prox-regular functions. In relaxing the result, we provide a direct proof and also extract the prox-regularity parameters from the proof.

Theorem 5.11. *For $i \in \{1,2,\ldots,m\}$, let $f_i : \mathbb{R}^n \to \mathbb{R}$ be $\mathcal{C}^1$. Let $\bar{\lambda} = (\bar{\lambda}_1, \bar{\lambda}_2, \ldots, \bar{\lambda}_m) > 0$. Let $\bar{x} \in \bigcap_{i=1}^m \operatorname{dom} f_i$. Define*

$$f(x,\lambda) := \max_i\{\lambda_i f_i(x)\}$$

where $\lambda = (\lambda_1, \lambda_2 \ldots, \lambda_m)$, and let $\bar{v} \in \partial_x f(\bar{x},\bar{\lambda})$. If the f_i are prox-regular at $\bar{x}$ for each $v_i \in \partial f_i(\bar{x})$ with parameters $\epsilon_i > 0$ and $r_i > 0$, then f is para-prox-regular at $\bar{x}$ for $\bar{v}$ with parameters $\epsilon = \min_i\{\epsilon_i, \frac{\bar{\lambda}_i\epsilon_i}{2}\}$ and $r = \max_i\{\frac{3\bar{\lambda}_i r_i}{2}\}$, with compatible parametrization by λ at $\bar{\lambda}$.

Proof. Let $A(\bar{x},\bar{\lambda}) = \{i : f(\bar{x},\bar{\lambda}) = \bar{\lambda}_i f_i(\bar{x})\}$ be the active set of indices. Then $\partial_x f(\bar{x},\bar{\lambda}) = \operatorname{conv}_{i\in A(\bar{x},\bar{\lambda})}\{\bar{\lambda}_i \nabla f_i(\bar{x})\}$ [6, Theorem 2.8.2], so every $\bar{v} \in \partial_x f(\bar{x},\bar{\lambda})$ has the form $\sum_{i\in A(\bar{x},\bar{\lambda})} \theta_i \bar{\lambda}_i v_i$ for some $v_i \in \partial f_i(\bar{x})$ and some set of non-negative $\{\theta_i\}_{i\in A(\bar{x},\bar{\lambda})}$ such that $\sum_{i\in A(\bar{x},\bar{\lambda})} \theta_i = 1$. Corollary 5.4 tells us that the $\lambda_i f_i$ are para-prox-regular at $(\bar{x},\bar{\lambda}_i)$ for all $\bar{\lambda}_i v_i \in \partial\bar{\lambda}_i f_i(\bar{x})$. In particular, for each i we can take $\hat{\epsilon}_i = \min\{\epsilon_i, \frac{\bar{\lambda}_i\epsilon_i}{2}\}$ and $\hat{r}_i = \frac{3\bar{\lambda}_i r_i}{2}$ to get

$$\lambda_i f_i(x') \geq \lambda_i f_i(x) + \langle \lambda_i \hat{v}_i, x' - x\rangle - \frac{\hat{r}_i}{2}|x' - x|^2 \tag{43}$$

for $x' \neq x$, $|x' - \bar{x}| < \hat{\epsilon}_i$, $|x - \bar{x}| < \hat{\epsilon}_i$, $|\lambda_i f_i(x) - \bar{\lambda}_i f(\bar{x})| < \hat{\epsilon}_i$, $\lambda_i\hat{v}_i \in \partial\lambda_i f_i(x)$, $|\lambda_i\hat{v}_i - \bar{\lambda}_i v_i| < \hat{\epsilon}_i$, and $|\lambda_i - \bar{\lambda}_i| < \hat{\epsilon}_i$. This is a system of m inequalities, so using $\epsilon = \min_i\{\hat{\epsilon}_i\}$ and $r = \max_i\{\hat{r}_i\}$ we can say that for all $i \in \{1,2,\ldots,m\}$

$$\lambda_i f_i(x') \geq \lambda_i f_i(x) + \langle \lambda_i \hat{v}_i, x' - x\rangle - \frac{r}{2}|x' - x|^2 \tag{44}$$

for all $x' \neq x, |x'-\bar{x}| < \epsilon, |x-\bar{x}| < \epsilon, |\lambda_i - \bar{\lambda}_i| < \epsilon, \lambda_i\hat{v}_i \in \partial\lambda_i f_i(x), |\lambda_i\hat{v}_i - \bar{\lambda}_i v_i| < \epsilon$, $|\lambda_i f_i(x) - \bar{\lambda}_i f_i(\bar{x})| < \epsilon$ (with r and ϵ no longer dependent on i). Since $f(x',\lambda) \geq \lambda_i f_i(x')$ for all i, we can replace the left-hand side of inequality (44) with $f(x',\lambda)$ to get

$$f(x',\lambda) \geq \lambda_i f_i(x) + \langle \lambda_i \hat{v}_i, x' - x\rangle - \frac{r}{2}|x' - x|^2 \tag{45}$$

for all $x' \neq x, |x' - \bar{x}| < \epsilon, |x - \bar{x}| < \epsilon, |\lambda_i - \bar{\lambda}_i| < \epsilon, \lambda_i \hat{v}_i \in \partial \lambda_i f_i(x), |\lambda_i \hat{v}_i - \bar{\lambda}_i v_i| < \epsilon$, $|\lambda_i f_i(x) - \bar{\lambda}_i f_i(\bar{x})| < \epsilon$. Let x', x, λ, v be such that $x' \neq x$, $|x' - \bar{x}| < \epsilon$, $|x - \bar{x}| < \epsilon$, $v \in \partial_x f(x, \lambda)$, $|v - \bar{v}| < \epsilon$, $|\lambda - \bar{\lambda}| < \epsilon$, $|f(x, \lambda) - f(\bar{x}, \bar{\lambda})| < \epsilon$. Let $A(x, \lambda) = \{i : f(x, \lambda) = \lambda_i f_i(x)\}$. Then $\partial_x f(x, \lambda) = \text{conv}_{i \in A(x,\lambda)}\{\lambda_i \partial f_i(x)\}$. So $v \in \partial_x f(x, \lambda)$ has the form $\sum_{i \in A(x,\lambda)} \theta_i \lambda_i \hat{v}_i$ for some $\hat{v}_i \in \partial f_i(x)$ and some non-negative $\{\theta_i\}_{i \in A(x,\lambda)}$ such that $\sum_{i \in A(x,\lambda)} \theta_i = 1$. We can also substitute $f(x, \lambda)$ on the right-hand side of inequality (45) for those $i \in A(x, \lambda)$ to get

$$f(x', \lambda) \geq f(x, \lambda) + \langle \lambda_i \hat{v}_i, x' - x \rangle - \frac{r}{2}|x' - x|^2. \tag{46}$$

Multiplying (46) by θ_i and adding over $i \in A(x, \lambda)$ yields

$$\begin{aligned} &\sum_{i \in A(x,\lambda)} \theta_i f(x', \lambda) \\ &\geq \sum_{i \in A(x,\lambda)} \theta_i f(x, \lambda) + \left\langle \sum_{i \in A(x,\lambda)} \theta_i \lambda_i \hat{v}_i, x' - x \right\rangle - \sum_{i \in A(x,\lambda)} \theta_i \frac{r}{2}|x' - x|^2 \end{aligned}$$

which simplifies to

$$f(x', \lambda) \geq f(x, \lambda) + \langle v, x' - x \rangle - \frac{r}{2}|x' - x|^2. \tag{47}$$

Therefore f is para-prox-regular at $\bar{x}$ for $\bar{v}$ with compatible parametrization by λ at $\bar{\lambda}$, with parameters ϵ and r. □

6. Open Questions

There are two areas of further research that are made apparent by this work. First, our results concerning weighted parametrized sums and weighted parametrized max functions rely on starting with a finite number of prox-regular functions. A natural direction to continue with the development of these theorems would be to explore the case where the family of functions is infinite. Second, all the results found in this paper are for functions in finite-dimensional space. A natural extension of this research would be to reexamine these results in a Hilbert or Banach space.

Acknowledgements. The authors thank the anonymous referee and our colleagues, H. Bauschke, P. Loewen, and X. Wang, for constructive feedback during the final stages of revision of this paper.

References

[1] H. Bauschke, Y. Lucet, M. Trienis: How to transform one convex function continuously into another, SIAM Rev. 50(1) (2008) 115–132.

[2] H. Bauschke, E. Matoušková, S. Reich: Projection and proximal point methods: convergence results and counterexamples, Nonlinear Anal. 56(5) (2004) 715–738.

[3] F. Bernard, L. Thibault: Prox-regularity of functions and sets in Banach spaces, Set-Valued Anal. 12(1–2) (2004) 25–47.

[4] S. Boralugoda, R. A. Poliquin: Local integration of prox-regular functions in Hilbert spaces, J. Convex Analysis 13(1) (2006) 27–36.

[5] M. Bounkhel, A. Jofré: Subdifferential stability of the distance function and its applications to nonconvex economies and equilibrium, J. Nonlinear Convex Anal. 5(3) (2004) 331–347.

[6] F. H. Clarke: Optimization and Nonsmooth Analysis, 2nd Ed., Classics in Applied Mathematics 5, Society for Industrial and Applied Mathematics, Philadelphia (1990).

[7] F. H. Clarke, R. J. Stern, P. R. Wolenski: Proximal smoothness and the lower-C^2 property, J. Convex Analysis 2(1–2) (1995) 117–144.

[8] A. Eberhard, R. Wenczel: A study of tilt-stable optimality and sufficient conditions, Nonlinear Anal. 75(3) (2012) 1260–1281.

[9] W. Hare, C. Sagastizábal: A redistributed proximal bundle method for nonconvex optimization, SIAM J. Optim. 20(5) (2010) 2442–2473.

[10] W. L. Hare: A proximal average for nonconvex functions: a proximal stability perspective, SIAM J. Optim. 20(2) (2009) 650–666.

[11] W. L. Hare: A proximal method for identifying active manifolds, Comput. Optim. Appl. 43(2) (2009) 295–306.

[12] W. L. Hare, R. A. Poliquin: Prox-regularity and stability of the proximal mapping, J. Convex Analysis 14(3) (2007) 589–606.

[13] A. B. Levy: Calm minima in parameterized finite-dimensional optimization, SIAM J. Optim. 11(1) (2000) 160–178.

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

[15] A. S. Lewis, D. R. Luke, J. Malick: Local linear convergence for alternating and averaged nonconvex projections, Found. Comput. Math. 9(4) (2009) 485–513.

[16] A. Moudafi: An algorithmic approach to prox-regular variational inequalities, Appl. Math. Comput. 155(3) (2004) 845–852.

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

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

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

[20] R. A. Poliquin, R. T. Rockafellar: A calculus of prox-regularity, J. Convex Analysis 17(1) (2010) 203–210.

[21] R. T. Rockafellar, R. J. Wets: Variational Analysis, Springer, Berlin (1998).