2020

Conditions for the existence, identification and calculus rules of the threshold of prox-boundedness

Chayne Planiden
University of Wollongong, chayne@uow.edu.au

Conditions for the existence, identification and calculus rules of the threshold of prox-boundedness

Abstract

2020, Springer-Verlag GmbH Germany, part of Springer Nature. This note advances knowledge of the threshold of prox-boundedness of a function; an important concern in the use of proximal point optimization algorithms and in determining the existence of the Moreau envelope of the function. In finite dimensions, we study general prox-bounded functions and then focus on some useful classes such as piecewise functions and Lipschitz continuous functions. The thresholds are explicitly determined when possible and bounds are established otherwise. Some calculus rules are constructed; we consider functions with known thresholds and find the thresholds of their sum and composition.

Disciplines

Engineering | Science and Technology Studies

Conditions for the existence, identification and calculus rules of the threshold of prox-boundedness

C. Planiden[*]

April 9, 2020

Abstract

This note advances knowledge of the threshold of prox-boundedness of a function; an important concern in the use of proximal point optimization algorithms and in determining the existence of the Moreau envelope of the function. In finite dimensions, we study general prox-bounded functions and then focus on some useful classes such as piecewise functions and Lipschitz continuous functions. The thresholds are explicitly determined when possible and bounds are established otherwise. Some calculus rules are constructed; we consider functions with known thresholds and find the thresholds of their sum and composition.

Keywords: Fenchel conjugate, infimal convolution, Lipschitz continuous, Moreau envelope, Moreau–Yosida regularization, piecewise function, prox-bounded, proximal mapping, regularization, threshold.

Mathematics Subject Classification: Primary 49A52; Secondary 46N10, 90C26

1 Introduction

The Moreau envelope function is a particular infimal convolution that first came about in the 1960s [32]. Given a function f on a finite-dimensional space, the Moreau envelope of f employs a nonnegative parameter r and is denoted $e_r f$:

$$e_r f(x) = \inf_{y \in \mathbb{R}^n} \left\{ f(y) + \frac{r}{2} \|y - x\|^2 \right\}. \tag{1.1}$$

It is a well-established, regularizing function that has many desirable properties when f has reasonable structure [38, 40]. The set of all solution points to (1.1) is known as the *proximal mapping* of f, denoted by Prox_f. The proximal mapping is a key component of many optimization algorithms, such as the proximal point method and its variants [5, 8, 13, 16, 20, 21, 37]. Because of the

---

[*]School of Mathematics and Applied Statistics, University of Wollongong, Wollongong, NSW, 2522, Australia. chayne@uow.edu.au

above and other nice features, the Moreau envelope and proximal mapping have been thoroughly researched and applied to many situations in the convex [10, 17, 24, 29, 33, 35] and nonconvex [22, 25, 26, 28, 30, 34, 39] settings.

For a function f, its Moreau envelope may or may not exist, or may exist only for certain x and/or certain r. If there exists a point x such that $e_r f(x) \in \mathbb{R}$ for some $r \geq 0$, we say that f is *prox-bounded*. If a function is not prox-bounded, then its Moreau envelope does not exist anywhere, for any choice of r. In that case, for any r fixed we have $e_r f(x) \equiv -\infty$, which is not a proper function (see [38]) and has empty domain.

It is the parameter r that is of primary interest in this work. There are many theoretically proved-convergent proximal algorithms (see [1, 2, 7, 9, 11, 12] and the references therein), but in practice, it has been observed that the initial choice of r and the manner of adjusting it as the algorithm runs are of critical importance, in order to obtain reliable performance [6, 14, 15, 23, 27, 36]. We explore the *threshold of prox-boundedness* of f: the infimum of the set of $r \geq 0$ such that $e_r f$ exists at at least one point.

In [19], the class of piecewise linear-quadratic (PLQ) functions on $\mathbb{R}^n$ was examined in the context of prox-boundedness. The main result of [19] is a theorem that explicitly identifies the threshold and the domain of the Moreau envelope of any finite-dimensional PLQ function. In that setting, the threshold is $\max r_i$, where r_i is the threshold of f_i for each i. One of the aims of the present work is to generalize that result in two aspects: where the functions f_i are not necessarily linear nor quadratic and the domains $\text{dom} f_i$ are not necessarily polyhedral. The main question on which we focus is this: what are the minimal conditions needed on f_i in order to be sure that the threshold of the piecewise function is $\max r_i$? We establish bounds and illustrate several counterexamples for functions with conditions that one might suspect sufficient to guarantee prox-boundedness, but are not. Under certain conditions, the threshold can be determined exactly.

The second focus of this work is the establishing of calculus rules for thresholds of prox-bounded functions. We explore classes of functions with known thresholds and study the conditions needed to determine the threshold of their sum and composition, or to produce an upper bound when an exact result cannot be found. By making use of the Fenchel conjugate representation of the Moreau envelope and some other previously-established properties and characterizations of prox-bounded functions, we determine sufficient conditions for the existence of a sum rule and a composition rule for thresholds of prox-boundedness.

The remainder of this work is organized as follows. Section 2 presents the notation used throughout and several definitions and known facts regarding prox-bounded functions and their thresholds. In Section 3, we examine the family of piecewise functions and determine the minimal conditions needed for finding the threshold. An example of what can go wrong when these conditions are not met is provided. Section 4 is dedicated to forming calculus rules for the threshold of the sum and the composition of prox-bounded functions. Section 5 offers concluding remarks and suggests interesting areas of further research in this vein.

2 Preliminaries

Throughout this paper, we work in finite-dimensional space $\mathbb{R}^n$, endowed with inner product $\langle \cdot, \cdot \rangle$ and induced norm $\| \cdot \|$. The set $\mathbb{R}^n \cup \{+\infty\}$ is denoted by $\overline{\mathbb{R}}$. We generally conform to the notation used in [38], including the terms proper, lower semicontinuous (lsc), Lipschitz continuous and indicator function defined therein. We denote the domain and image of a function f by $\operatorname{dom} f$ and $\operatorname{Im} f$, respectively.

Definition 2.1. *For a function $f : X \rightarrow Y$, X is called the* domain *of f and Y is called the* codomain *of f. The domain of f is the set $\operatorname{dom} f = X = \{x : f(x) \in Y\}$. The* image *of f is the set $\operatorname{Im} f = \{y \in Y : f(x) = y$ for some $x \in X\}$.*

Definition 2.2. *For any set S, the* indicator function *ι_S is defined by*

$$\iota_S(x) = \begin{cases} 0, & \text{if } x \in S, \\ \infty, & \text{otherwise.} \end{cases}$$

Definition 2.3. *For proper functions $f_i : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$, $i \in \{1, 2, \ldots, m\}$, the* piecewise function *$f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ is defined by*

$$f(x) = \begin{cases} f_1(x), & x \in S_1, \\ \quad\quad \vdots \\ f_m(x), & x \in S_m, \end{cases}$$

where $\bigcup_i S_i = \mathbb{R}^n$ and $S_i \cap \operatorname{int} S_j = \varnothing$ for $i \neq j$.

Definition 2.4. *A function $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ is* prox-bounded *if there exists $r \geq 0$ such that $e_r f(x) > -\infty$ for some $x \in \mathbb{R}^n$. The infimum of all such r is called the* threshold of prox-boundedness *of f.*

Our interest in this work is to identify thresholds of prox-boundedness of functions and the conditions for their existence. To that end, we list the following results from previous works.

Fact 2.5. [19, Lemma 2.4] *Let $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ be proper and lsc. Then f is bounded below if and only if its threshold $\bar{r} = 0$ and $\operatorname{dom} e_{\bar{r}} f = \mathbb{R}^n$.*

Fact 2.6. [38, Theorem 1.25] *Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be proper, lsc and prox-bounded with threshold $\bar{r}$. Then for all $r > \bar{r}$, $\operatorname{dom} e_r f = \mathbb{R}^n$.*

Fact 2.7. [38, Example 3.28] *Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be such that*

$$\liminf_{\|x\| \to \infty} \frac{f(x)}{\|x\|} > -\infty.$$

Then f is prox-bounded with threshold $\bar{r} = 0$.

Fact 2.8. [38, Theorem 2.26] *Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be proper, lsc and convex. Then f is prox-bounded with threshold $\bar{r} = 0$.*

3

Note that by definition of prox-regular, if $r < \bar{r}$, then $e_r f(x) = -\infty$ for all $x \in \mathbb{R}^n$. By Fact 2.6, if $r > \bar{r}$, then $e_r f(x) > -\infty$ for all $x \in \mathbb{R}^n$. At the threshold itself, however, there is no such universal behaviour of the Moreau envelope. Depending on the nature of f, $\operatorname{dom} e_{\bar{r}} f$ can be empty, full-domain or a proper subset of $\mathbb{R}^n$, even for very simple functions (see [19, Examples 2.5–2.7]). This is partly why the choice of initial prox-parameter r and the manner in which it changes are so crucial in many minimization algorithms; vastly different proximal behaviour is possible with distinct values of r. In the next section, we explore existence conditions for the threshold of prox-boundedness of piecewise functions.

3 Existence of thresholds of prox-boundedness

The conditions for prox-boundedness in the case of PLQ functions was thoroughly examined in [19]. Now we move beyond that class of functions. We concentrate primarily on piecewise functions as defined in Definition 2.3, as that is a natural extension to what has been accomplished already. We begin by establishing the fact that Lipschitz continuous functions are prox-bounded. The definition of the little-o set is required for the proof.

Definition 3.1. *Let $f, g : \mathbb{R}^n \to \overline{\mathbb{R}}$. We say that f is* little-o *of g, written $f \in o(g)$, if*

$$\lim_{\|x\| \to \infty} \frac{f(x)}{g(x)} = 0.$$

That is, $f \in o(g)$ if asymptotically g grows in magnitude faster than f.

Proposition 3.2. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper and globally K-Lipschitz. Then f is prox-bounded with threshold $\bar{r} = 0$.*

Proof. Let $r > 0$ and $x \in \mathbb{R}^n$. We have

$$f(y) + \frac{r}{2}\|x - y\|^2 = f(y) - f(x) + \frac{r}{2}\|x - y\|^2 + f(x)$$
$$\geq -K\|x - y\| + \frac{r}{2}\|x - y\|^2 + f(x).$$

Thus,
$$e_r f(x) \geq \inf_{y \in \mathbb{R}^n} \left\{ \frac{r}{2}\|x - y\|^2 - K\|x - y\| \right\} + f(x) > -\infty,$$

since the quadratic function $y \mapsto \frac{r}{2}\|x - y\|^2 - K\|x - y\|$ is bounded below. Since this holds true for every $r > 0$, we conclude that $\bar{r} = 0$. $\qquad\square$

Now we focus on the family of piecewise functions and say what we can about their thresholds. Henceforth, any mention of a piecewise function refers to a function defined as in Definition 2.3.

Proposition 3.3. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a proper, lsc, piecewise function. Then f is prox-bounded if and only if $f_i + \iota_{S_i}$ is prox-bounded for each i.*

Proof. ($\Rightarrow$) Let f be prox-bounded with threshold $\bar{r}$. Suppose that there exists j such that $f_j + \iota_{S_j}$ is not prox-bounded. Fix $r > \bar{r}$, so that $e_r f(x) \in \mathbb{R}$ for all x (Fact 2.6). By definition of $e_r f$, for $\bar{x}$ fixed we have

$$
\begin{aligned}
e_r f(\bar{x}) &= \inf_y \left\{ f(y) + \frac{r}{2} \|y - \bar{x}\|^2 \right\} \\
&= \min \left[\inf_y \left\{ f_1(y) + \iota_{S_1}(y) + \frac{r}{2} \|y - \bar{x}\|^2 \right\}, \ldots, \inf_y \left\{ f_m(y) + \iota_{S_m}(y) + \frac{r}{2} \|y - \bar{x}\|^2 \right\} \right]. \quad (3.1)
\end{aligned}
$$

Since $f_j + \iota_{s_j}$ is not prox-bounded, we have

$$
\inf_y \left\{ f_j(y) + \iota_{S_j}(y) + \frac{r}{2} \|y - \bar{x}\|^2 \right\} = -\infty,
$$

hence $e_r f(\bar{x}) = -\infty$. This is a contradiction to the fact that $e_r f(\bar{x}) \in \mathbb{R}$. Therefore, $f_i + \iota_{S_i}$ is prox-bounded for all $i \in \{1, 2, \ldots, m\}$.

($\Leftarrow$) Let $f_i + \iota_{S_i}$ be prox-bounded with threshold r_i for each $i \in \{1, 2, \ldots, m\}$. Let $\bar{r} = \max r_i$, and choose any $r > \bar{r}$. Then $r > r_i$ for all i, so that $e_r f_i(x) \in \mathbb{R}$ for all i and for all x. Then each infimum in (3.1) is a real number, hence the minimum of (3.1) exists. Therefore, $e_r f(\bar{x}) \in \mathbb{R}$, and we have that f is prox-bounded. $\qquad\square$

Theorem 3.4. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a proper, lsc, piecewise function and let each $f_i + \iota_{S_i}$ be prox-bounded with threshold r_i. Then the threshold of f is $\bar{r} = \max r_i$.*

Proof. Choose any $r > \max r_i$. By Proposition 3.3, there exists $\bar{x}$ such that $e_r f(\bar{x}) \in \mathbb{R}$. Since r is arbitrary, we have that f is prox-bounded for all $r > \max r_i$. Hence, $\bar{r} \leq \max r_i$. Now choose any $r < \max r_i$. Then there exists $j \in \{1, 2, \ldots, m\}$ such that $r < r_j$. By definition of prox-boundedness, $e_r \left(f_j + \iota_{S_j} \right)(x) = -\infty$ for all x. The Moreau envelope of f is the expression of (3.1), whose minimand contains at least one instance of $-\infty$ due to $f_j + \iota_{S_j}$. Hence, $e_r f(x) = -\infty$ for all x, and we have that $\bar{r} \geq \max r_i$. Therefore, $\bar{r} = \max r_i$. $\qquad\square$

We have our first results for piecewise functions. However, the result of Theorem 3.4 is weakened if the term ι_{S_i} is removed from the statement and we require f_i itself to be prox-bounded, as Theorem 3.5 shows.

Theorem 3.5. *For $i \in \{1, 2, \ldots, m\}$, let $f_i : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper, lsc and prox-bounded with threshold r_i. With these f_i, define a piecewise function f as per Definition 2.3. Then f is prox-bounded with threshold $\bar{r} \leq \max r_i$.*

Proof. Let $r > r_i$. Since f_i is prox-bounded with threshold r_i, there exists $\bar{x} \in \mathbb{R}^n$ such that $e_r f_i(\bar{x}) > -\infty$. Since

$$
e_r f_i(\bar{x}) = \inf_{y \in \mathbb{R}^n} \left\{ f_i(y) + \frac{r}{2} \|y - \bar{x}\|^2 \right\} \leq \inf_{y \in S_i} \left\{ f_i(y) + \frac{r}{2} \|y - \bar{x}\|^2 \right\} = e_r(f_i + \iota_{S_i})(\bar{x}),
$$

we have that $e_r(f_i + \iota_{S_i})(\bar{x}) > -\infty$. Hence, $f_i + \iota_{S_i}$ is prox-bounded with threshold $\tilde{r}_i \leq r_i$. This is true for all $i \in \{1, 2, \ldots, m\}$, so Theorem 3.4 applies and we have $\bar{r} = \max \tilde{r}_i \leq \max r_i$. $\qquad\square$

5

The best we can do is an upper bound in this case. One might hope to establish a lower bound for $\bar{r}$ as well, such as $\min r_i$. However, this cannot be done in the general setting of Theorem 3.5. Example 3.6 illustrates why not.

Example 3.6. *Let $f_1, f_2 : \mathbb{R} \to \mathbb{R}$ be defined by*

$$f_1(x) = \begin{cases} x^2, & x < 0, \\ -x^2, & x \geq 0, \end{cases} \qquad f_2(x) = \begin{cases} -x^2, & x < 0, \\ x^2, & x \geq 0. \end{cases}$$

Define

$$f(x) = \begin{cases} f_1(x), & x < 0, \\ f_2(x), & x \geq 0. \end{cases}$$

Then the thresholds of f_1 and f_2 are $r_1 = r_2 = 2$ and the threshold of f is $\bar{r} = 0$.

Proof. Considering f_1 first, we define

$$\varphi_r(y) = f_1(y) + \frac{r}{2}|y - x|^2 = \begin{cases} \left(1 + \frac{r}{2}\right) y^2 - rxy + \frac{r}{2}x^2, & y < 0, \\ \left(-1 + \frac{r}{2}\right) y^2 - rxy + \frac{r}{2}x^2, & y \geq 0. \end{cases}$$

For any $r > 2$, both pieces of φ_r are strictly convex quadratic. Thus, $e_r f_1(x) = \inf \phi_r(y) > -\infty$ and $r_1 \leq 2$. For any $r < 2$, the second piece of φ_r is concave quadratic, so $e_r f_1(x) = -\infty$ for all x and $r_1 \geq 2$. Therefore, $r_1 = 2$. Similarly, $r_2 = 2$. But $f(x) = x^2$ has $\bar{r} = 0$ by Fact 2.5. $\square$

It is equally simple to construct an example where $\bar{r} = \max r_i$ for a piecewise function, for instance

$$f(x) = \begin{cases} f_1(x), & x \geq 0, \\ f_2(x), & x < 0 \end{cases}$$

where f_1, f_2 are defined in Example 3.6. In that case, $f(x) = -x^2$ and $\bar{r} = 2 = \max\{r_1, r_2\}$. So we cannot do better than bounding $\bar{r}$ from above in this most general setting. Furthermore, one can obtain a prox-bounded function from the sum of two functions that are not prox-bounded. For instance, $f_1(x) = x^3$ and $f_2(x) = -x^3$ are not prox-bounded, but their sum is the constant function zero, with threshold zero. The next section considers more specific cases of both the sum and the composition of functions, where we can make some tighter conclusions about exact thresholds.

4   Calculus of the threshold of prox-boundedness

In this section, we consider the thresholds of the sum and the composition of prox-bounded functions. The functions here are no longer (necessarily) piecewise functions, as they were in the previous section. The following definition and facts will be useful.

Definition 4.1 (Fenchel conjugate). *For any function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, the* Fenchel conjugate *of f is the function $f^* : \mathbb{R}^n \to \overline{\mathbb{R}}$ defined by*

$$f^*(y) = \sup_{x \in \mathbb{R}^n} \{\langle y, x \rangle - f(x)\}.$$

The Fenchel conjugate and the Moreau envelope enjoy a beautiful equivalence, as the following fact states.

Fact 4.2. [34, Proposition 2.1] *For any proper function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$,*

$$e_r f(x) = \frac{r}{2} \|x\|^2 - g_r^*(rx),$$

where $g_r(x) = f(x) + \frac{r}{2}\|x\|^2$.

Fact 4.3. [38, Exercise 1.24] *For a proper, lsc function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, the following are equivalent:*

(i) *f is prox-bounded;*

(ii) *f majorizes a quadratic function;*

(iii) *there exists $r \in \mathbb{R}$ such that $f + \frac{r}{2}\| \cdot \|^2$ is bounded below;*

(iv) *$\liminf\limits_{\|x\| \to \infty} \frac{f(x)}{\|x\|^2} > -\infty$.*

If $\hat{r}$ is the infimum of all r for which (iii) holds, the limit in (iv) is $-\frac{\hat{r}}{2}$ and the threshold for f is $\bar{r} = \max\{0, \hat{r}\}$.

First, we address the quadratic function mentioned in Fact 4.3(ii). Rockafellar and Wets state that it exists, but give no details as to its form. Lemma 4.4 below describes the curvature that such a quadratic function must have, in terms of the threshold.

Lemma 4.4. *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper, lsc and prox-bounded with threshold $\bar{r}$. Then for any $r > \bar{r}$, there exists $m_r = m(r) \in \mathbb{R}$ such that f is bounded below by the quadratic function $q_r : \mathbb{R}^n \to \mathbb{R}$, $q_r = -\frac{r}{2}\| \cdot \|^2 + m_r$. Furthermore, for any $r < \bar{r}$ there does not exist $m_r \in \mathbb{R}$ such that f is bounded below by q_r. Therefore, the smallest possible curvature of a quadratic function that minorizes f is $\frac{\bar{r}}{2}$, which occurs in the case $\lim\limits_{r \searrow \bar{r}} m_r > -\infty$.*

Proof. Let $r > \bar{r}$. By Fact 4.3(iii), we have that $f + \frac{r}{2}\| \cdot \|^2$ is bounded below, i.e., there exists $m_r \in \mathbb{R}$ such that for all $x \in \operatorname{dom} f$, $f(x) + \frac{r}{2}\|x\|^2 \geq m_r$. Rearranging, we have

$$f(x) \geq -\frac{r}{2}\|x\|^2 + m_r \quad \forall x \in \operatorname{dom} f. \tag{4.1}$$

Since (4.1) is true for all $r > \bar{r}$, we have

$$\lim_{r \searrow \bar{r}} f(x) \geq \lim_{r \searrow \bar{r}} \left(-\frac{r}{2}\|x\|^2 + m_r \right)$$

$$f(x) \geq -\frac{\bar{r}}{2}\|x\|^2 + \lim_{r \searrow \bar{r}} m_r.$$

7

Therefore, if the above limit exists, we have that f is bounded below by $q_{\bar{r}}$. If the limit is $-\infty$, then f is bounded below by q_r for all $r > \bar{r}$, but not by $q_{\bar{r}}$. Now suppose that for some $r < \bar{r}$, there exists m_r such that (4.1) holds. Then we have that $f + \frac{\hat{r}}{2}\|\cdot\|^2$ is bounded below for some $\hat{r} \in (r, \bar{r})$, which by Fact 4.3(iii) and the postamble contradicts the fact that $\bar{r}$ is the threshold of f. Therefore, $\frac{\bar{r}}{2}$ is the minimum curvature possible of a quadratic minorant of f. $\qquad\square$

Remark 4.5. *It is easy to see that some nonconvex functions, such as concave quadratic functions, are indeed bounded below by $q_{\bar{r}}$. However, there are others that are not. For instance, the function $f : \mathbb{R} \to \mathbb{R}$, $f(x) = -|x|$ has threshold zero and is not bounded below by q_0 (i.e. constant function), though it is bounded below by q_r for any $r > 0$. It is also interesting to note that this can happen with convex functions as well. Nonconstant affine functions, which are convex and have threshold zero, are not bounded below by q_0.*

Now we focus on the threshold of the sum of two prox-bounded functions. As in the case of the piecewise function with prox-bounded pieces, we will find that an exact threshold cannot be obtained and we settle for an upper bound in the general case. If certain restrictions are put on one or both of the functions, an exact threshold can be determined.

Lemma 4.6. *Let $f_1, f_2 : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper, lsc and prox-bounded with respective thresholds r_1, r_2. If $f_1 \leq f_2$, then for any $r > \max\{r_1, r_2\}$, we have*

$$-\infty < e_r f_1 \leq e_r f_2.$$

Proof. We have $-\infty < e_r f_1$ by Fact 2.6. Defining $g_1, g_2 : \mathbb{R}^n \to \overline{\mathbb{R}}$ by $g_1(x) = f_1(x) + \frac{r}{2}\|x\|^2$ and $g_2(x) = f_2(x) + \frac{r}{2}\|x\|^2$, we have $g_1 \leq g_2$. By [4, Proposition 13.14(ii)], $g_1^*(rx) \geq g_2^*(rx)$. By Fact 4.2, we have

$$\begin{aligned}
e_r f_1(x) &= \frac{r}{2}\|x\|^2 - g_1^*(rx), \\
&\leq \frac{r}{2}\|x\|^2 - g_2^*(rx), \\
&= e_r f_2(x). \qquad\qquad\square
\end{aligned}$$

Corollary 4.7. *Let $f_1, f_2 : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper, lsc and prox-bounded with respective thresholds r_1, r_2. If $f_1 \leq f_2$, then $r_1 \geq r_2$.*

Proposition 4.8. *Let $f_1, f_2 : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper, lsc and prox-bounded with respective thresholds r_1, r_2. Define $f(x) = (f_1 + f_2)(x)$. Then f is prox-bounded with threshold $\bar{r} \leq r_1 + r_2$. Moreover, if f_2 is bounded, then $\bar{r} = r_1$.*

Proof. The first part of this proposition appears as part of [18, Lemma 2.4], but we provide a full

proof for the sake of completeness. For any $\varepsilon > 0$, we have

$$e_{r_1+r_2+2\varepsilon} f(x) = \inf_y \left\{ f_1(y) + f_2(y) + \frac{r_1 + r_2 + 2\varepsilon}{2} \|y - x\|^2 \right\}$$

$$= \inf_y \left\{ \left[f_1(y) + \frac{r_1 + \varepsilon}{2} \|y - x\|^2 \right] + \left[f_2(y) + \frac{r_2 + \varepsilon}{2} \|y - x\|^2 \right] \right\}$$

$$\geq \inf_y \left\{ f_1(y) + \frac{r_1 + \varepsilon}{2} \|y - x\|^2 \right\} + \inf_y \left\{ f_2(y) + \frac{r_2 + \varepsilon}{2} \|y - x\|^2 \right\}$$

$$= e_{r_1+\varepsilon} f_1(x) + e_{r_2+\varepsilon} f_2(x) > -\infty \ \forall x \in \mathbb{R}^n .$$

So we have $e_{(r_1+r_2)+2\varepsilon} f(x) > -\infty$ for all $x \in \mathbb{R}^n$ and for any $\varepsilon > 0$, which by Fact 2.6 tells us that f is prox-bounded and $\bar{r} \leq r_1 + r_2$. Now suppose that f_2 is bounded. Since f_2 is bounded below, we have that $r_2 = 0$ by Fact 2.5. Hence, $\bar{r} \leq r_1 + 0 = r_1$. Since f_2 is bounded above, there exists $M \in \mathbb{R}$ such that $M \geq f_2(x) \ \forall x$. Suppose that $r_1 > 0$. (Otherwise, trivially $\bar{r} = 0 = r_1$.) Then for any $r \in (0, r_1)$, we have

$$e_r f(x) = \inf_y \left\{ f_1(y) + f_2(y) + \frac{r}{2} \|y - x\|^2 \right\},$$

$$\leq \inf_y \left\{ f_1(y) + M + \frac{r}{2} \|y - x\|^2 \right\},$$

$$= M + \inf_y \left\{ f_1(y) + \frac{r}{2} \|y - x\|^2 \right\},$$

$$= M + e_r f_1(x) = -\infty.$$

Hence, $\bar{r} \geq r_1$. Therefore, $\bar{r} = r_1$. $\qquad \square$

Corollary 4.9. *Let $f_1, f_2 : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper, lsc and prox-bounded with threshold 0. Then $(f_1 + f_2)(x)$ is prox-bounded with threshold 0.*

The very strong condition of f_2 being bounded above and below in Proposition 4.8 can be relaxed slightly, as the corollary below indicates, with the same proof as the proposition.

Corollary 4.10. *Let $f_1, f_2 : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper, lsc and prox-bounded with respective thresholds r_1, r_2. Define $f(x) = (f_1 + f_2)(x)$. If $r_2 = 0$ and f_2 is bounded above, then $\bar{r} = r_1$.*

Proposition 4.11. *Let $f_1 : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper and lsc. Let $f_2 : \mathbb{R}^n \to \overline{\mathbb{R}}$ be an affine function. Then $f_1 + f_2$ is prox-bounded with threshold r_1 if and only if f_1 is prox-bounded with threshold r_1.*

Proof. We see in [3, Lemma 3.6] that the Moreau envelope of f can be expressed as the sum of a quadratic function and a Moreau envelope of f_1 only, with x plus a constant as the argument. Therefore, $e_r f$ exists if and only if $e_r f_1$ exists and we have that f has the same threshold as f_1. $\quad \square$

Proposition 4.11 invites another slight relaxation of the condition on f_2 in Proposition 4.8.

Corollary 4.12. *Let $f_1, f_2 : \mathbb{R}^n \to \overline{\mathbb{R}}$ be prox-bounded with respective thresholds r_1, r_2. Define $f = f_1 + f_2$. If $r_2 = 0$ and f_2 is majorized by an affine function, then f is prox-bounded with threshold $\bar{r} = r_1$.*

Now we move on to sufficient conditions for a composition rule. This is a difficult issue; one can construct examples of composition where the resulting threshold is any nonnegative number one desires, or even nonexistent. The following simple examples demonstrate.

Example 4.13. *For $a, b \geq 0$, define $f_1, f_2 : \mathbb{R} \to \mathbb{R}$,*

$$f_1(x) = -bx, \qquad f_2(x) = -\frac{a}{2}x^2.$$

Then the threshold of $f_1 \circ f_2$ is $\bar{r}_{12} = 0$, while the threshold of $f_2 \circ f_1$ is $\bar{r}_{21} = ab^2$.

Proof. Since $(f_1 \circ f_2)(x) = \frac{ab}{2}x^2$ is bounded below, by Fact 2.5 we have $\bar{r}_{12} = 0$. We find that $(f_2 \circ f_1)(x) = -\frac{ab^2}{2}x^2$, so by the same method as the proof of Example 3.6 we have $\bar{r}_{21} = ab^2$. $\quad\square$

Example 4.13 shows that with two basic prox-bounded functions it is possible to obtain a threshold for the composition that is any particular nonnegative number, by making appropriate choices of a and b. The next example shows that we can just as easily use two prox-bounded functions to construct a function that is not prox-bounded.

Example 4.14. *Define $f_1(x) = x^2$, $f_2(x) = -x^2$. Then $f_1 \circ f_2$ has threshold $\bar{r}_{12} = 0$, while $f_2 \circ f_1$ is not prox-bounded.*

Proof. Since $(f_1 \circ f_2)(x) = x^4$ is bounded below, by Fact 2.5 we have that $\bar{r}_{12} = 0$. We find that $(f_2 \circ f_1)(x) = -x^4$, so by Fact 4.3(ii), $f_2 \circ f_1$ is not prox-bounded. $\quad\square$

Furthermore, one can compose two functions that are not prox-bounded to form a function that is prox-bounded. For instance on $\mathbb{R}^+$, $f_1(x) = -x^3$ and $f_2(x) = \ln x$ are not prox-bounded, yet they yield the composition $(f_1 \circ f_2)(x) = -\ln^3 x$, which is minorized by $-x^2$ and thus prox-bounded by Fact 4.3(ii). So what can we say about the thresholds of the composition of prox-bounded functions? As in the case of the sum of prox-bounded functions, if we restrict ourselves to certain classes of functions, we can make some conclusions. We start by listing a known fact that is used in the proof of the subsequent proposition.

Fact 4.15. [18, Lemma 2.4] *Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper, lsc and prox-bounded with threshold $\bar{r}$. Then for any $\lambda \geq 0$, λf is prox-bounded with threshold $\lambda \bar{r}$.*

Proposition 4.16. *Let $f_1 : \mathbb{R}^m \to \overline{\mathbb{R}}$ and $f_2 : \mathbb{R}^n \to \mathbb{R}^m$ be such that $\operatorname{Im} f_2 \subseteq \operatorname{dom} f_1$. Define $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, $f = f_1 \circ f_2$. If f_1, f_2 are Lipschitz continuous, then the threshold of f is $\bar{r} = 0$.*

Proof. Let f_1 be K_1-Lipschitz and f_2 be K_2-Lipschitz. Then

$$\|f_1(f_2(y)) - f_1(f_2(x))\| \leq K_1 \|f_2(y) - f_2(x)\| \leq K_1 K_2 \|y - x\|,$$

which says that $f_1 \circ f_2$ is $K_1 K_2$-Lipschitz. By Proposition 3.2, $\bar{r} = 0$. $\quad\square$

Proposition 4.17. *Let $f_1, f_2 : \mathbb{R} \to \overline{\mathbb{R}}$ be prox-bounded with respective thresholds $\bar{r}_1, \bar{r}_2$. Define $f : \mathbb{R} \to \overline{\mathbb{R}}$, $f = f_1 \circ f_2$.*

(i) *If f_1 is affine: $f_1(x) = ax + b$ with $a \in \mathbb{R}_+$ and $b \in \mathbb{R}$, then the threshold of f is $\bar{r} = a\bar{r}_2$.*

(ii) *If f_2 is affine: $f_2(x) = ax + b$ with $a, b \in \mathbb{R}$, then the threshold of f is $\bar{r} = a^2\bar{r}_1$.*

Proof. (i) We have $(f_1 \circ f_2)(x) = af_2(x) + b$. This yields $\bar{r} = a\bar{r}_2$ by Fact 4.15 and the fact that the vertical shift by the constant value b has no impact on the threshold.

(ii) We have $(f_1 \circ f_2)(x) = f_1(ax + b)$. Let $r > \bar{r}_1$. Then by Fact 4.3(iii), $f_1 + \frac{r}{2}\|\cdot\|^2$ is bounded below. We have, for some $m \in \mathbb{R}$ and for all $x \in \operatorname{dom} f_1$,

$$f_1(x) + \frac{r}{2}\|x\|^2 \geq m,$$

$$f_1(ax + b) + \frac{r}{2}\|ax + b\|^2 \geq m,$$

$$f_1(ax + b) + ar\langle x, b \rangle + \frac{a^2 r}{2}\|x\|^2 \geq m - \frac{r}{2}\|b\|^2. \tag{4.2}$$

Hence, $f_1(ax + b) + ar\langle x, b \rangle + \frac{a^2 r}{2}\|x\|^2$ is bounded below. Since (4.2) is true for any arbitrary $r > \bar{r}_1$, it is true for all $r > \bar{r}_1$. By an identical argument, for any $r < \bar{r}_1$ we have that $f_1(ax + b) + ar\langle x, b \rangle + \frac{a^2 r}{2}\|x\|^2$ is not bounded below. Thus, $\bar{r}_1$ is the infimum of all r such that (4.2) is true. By Fact 4.3, the threshold of $f_1(a \cdot + b) + ar\langle \cdot, b \rangle$ is $a^2\bar{r}_1$. By Proposition 4.11, we conclude that the threshold of $f_1 \circ f_2$ is $a^2\bar{r}_1$. $\qquad \square$

So we have that (on $\mathbb{R}^n$) if both f_1, f_2 are Lipschitz continuous functions, or (on $\mathbb{R}$) if one of f_1, f_2 is affine, then the threshold of the composition can be determined exactly. It is clear from Examples 4.13 and 4.14 that if one of f_1, f_2 is quadratic, chaos ensues. So far, it does not seem that other standard properties such as convexity and boundedness are any more promising in forming composition rules, not even in providing an upper bound for the threshold. We leave the further development of properties of the threshold of prox-boundedness to future consideration.

5 Conclusion and future work

The threshold of prox-boundedness of the objective function of a minimization problem is an important value to take into consideration when implementing optimization algorithms. In this work, we have determined the threshold of Lipschitz functions and bounds on the threshold of piecewise functions. We established properties of thresholds of the sum and the composition of functions under certain conditions and shown that when we do not have these conditions, functions can be constructed so that the threshold of the sum or composition is any nonnegative number.

This paper is the first step in determining thresholds for larger classes of functions, with the long-term goal of improving the efficiency of optimization routines that are based in the proximal point algorithm. At the moment, the conditions imposed are quite heavy; the search continues for other well-behaved functions whose thresholds can be identified or at least bounded. The work done here regarding piecewise functions, together with the results of [19] on PLQ functions, should open the way for exploration of thresholds of composition classes such as the fully subamenable functions of [31]. Such functions are an extension of fully amenable functions [38], respect a chain rule and are suitable for use in constrained composite modelling and optimization applications.

Acknowledgement

References

[1] M. Annergren, A. Hansson, and B. Wahlberg. An ADMM algorithm for solving l_1-regularized MPC. In *2012 IEEE 51st Conference on Decision and Control*, pages 4486–4491. IEEE, 2012.

[2] H. Attouch and J. Peypouquet. Convergence of inertial dynamics and proximal algorithms governed by maximally monotone operators. *Math. Program.*, 174(1-2, Ser. B):391–432, 2019.

[3] M. Bačak, J. Borwein, A. Eberhard, and B. Mordukhovich. Infimal convolutions and Lipschitzian properties of subdifferentials for prox-regular functions in Hilbert spaces. *J. Convex Anal.*, 17(3-4):737–763, 2010.

[4] H. Bauschke and P. Combettes. *Convex analysis and monotone operator theory in Hilbert spaces*. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC. Springer, New York, 2011.

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

[6] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Found. Trend. Mach. Learn.*, 3(1):1–122, 2011.

[7] L. Briceno-Arias, P. Combettes, J.-C. Pesquet, and N. Pustelnik. Proximal algorithms for multicomponent image recovery problems. *J. Math. Imag. Vision*, 41(1-2):3–22, 2011.

[8] R. Burachik and A. Iusem. A generalized proximal point algorithm for the nonlinear complementarity problem. *RAIRO Oper. Res.*, 33(4):447–479, 1999.

[9] P. Calamai and J. Moré. Projected gradient methods for linearly constrained problems. *Math. Program.*, 39(1):93–116, 1987.

[10] Y. Chen, C. Kan, and W. Song. The Moreau envelope function and proximal mapping with respect to the Bregman distances in Banach spaces. *Vietnam J. Math.*, 40(2-3):181–199, 2012.

[11] M. Dao and M. Tam. Union averaged operators with applications to proximal algorithms for min-convex functions. *J. Optim. Theory Appl.*, 181(1):61–94, 2019.

[12] J. Douglas and H. Rachford. On the numerical solution of heat conduction problems in two and three space variables. *Trans. Amer. Math. Soc.*, 82(2):421–439, 1956.

[13] J. Eckstein and D. Bertsekas. On the Douglas–Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Math. Program.*, 55(3, Ser. A):293–318, 1992.

[14] C. Fougner and S. Boyd. Parameter selection and preconditioning for a graph form solver. In *Emerging Applications of Control and Systems Theory*, pages 41–61. Springer, 2018.

[15] A. Fuduli and M. Gaudioso. Tuning strategy for the proximity parameter in convex minimization. *J. Optim. Theory Appl.*, 130(1):95–112, 2006.

[16] M. Fukushima and H. Mine. A generalized proximal point algorithm for certain nonconvex minimization problems. *Internat. J. Systems Sci.*, 12(8):989–1000, 1981.

[17] O. Güler. On the convergence of the proximal point algorithm for convex minimization. *SIAM J. Control Optim.*, 29(2):403–419, 1991.

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

[19] W. Hare and C. Planiden. Thresholds of prox-boundedness of PLQ functions. *J. Convex Anal.*, 23(3), 2014.

[20] W. Hare and C. Planiden. Computing proximal points of convex functions with inexact subgradients. *Set-Valued and Variational Analysis*, 26(3):469–492, 2018.

[21] W. Hare, C. Planiden, and C. Sagastizábal. A derivative-free VU-algorithm for convex finite-max problems. *arXiv preprint arXiv:1903.11184*, 2019.

[22] W. Hare and C. Sagastizábal. Computing proximal points of nonconvex functions. *Math. Program.*, 116(1-2, Ser. B):221–258, 2009.

[23] B. He, H. Yang, and S. Wang. Alternating direction method with self-adaptive penalty parameters for monotone variational inequalities. *J. Optim. Theory Appl.*, 106(2):337–356, 2000.

[24] J.-B. Hiriart-Urruty and H. Le. From Eckart and Young approximation to Moreau envelopes and vice versa. *RAIRO Oper. Res.*, 47(3):299–310, 2013.

[25] A. Kaplan and R. Tichatschke. Proximal point methods and nonconvex optimization. *J. Global Optim.*, 13(4):389–406, 1998. Workshop on Global Optimization (Trier, 1997).

[26] I. Kecis and L. Thibault. Moreau envelopes of s-lower regular functions. *Nonlinear Anal.*, 127:157–181, 2015.

[27] Y. Li, H. Zhang, Z. Li, and H. Gao. Proximal gradient method with automatic selection of the parameter by automatic differentiation. *Optim. Methods Softw.*, 33(4-6):708–717, 2018.

[28] Y. Lucet. Fast Moreau envelope computation. I. Numerical algorithms. *Numer. Algorithms*, 43(3):235–249 (2007), 2006.

[29] B. Martinet. Régularisation d'inéquations variationnelles par approximations successives. *Rev. Française Informat. Recherche Opérationnelle*, 4(Ser. R-3):154–158, 1970.

[30] R. Mifflin and C. Sagastizábal. VU-smoothness and proximal point results for some nonconvex functions. *Optim. Methods Softw.*, 19(5):463–478, 2004.

[31] A. Mohammadi, B. Mordukhovich, and M. Sarabi. Variational analysis of composite models with applications to continuous optimization. *arXiv preprint arXiv:1905.08837*, 2019.

[32] J.-J. Moreau. Proximité et dualité dans un espace Hilbertien. *Bull. Soc. Math. France*, 93:273–299, 1965.

[33] C. Planiden and X. Wang. Strongly convex functions, Moreau envelopes, and the generic nature of convex functions with strong minimizers. *SIAM J. Optim.*, 26(2):1341–1364, 2016.

[34] C. Planiden and X. Wang. Epi-convergence: the Moreau envelope and generalized linear-quadratic functions. *J. Optim. Theory Appl.*, 177(1):21–63, 2018.

[35] C. Planiden and X. Wang. Proximal mappings and Moreau envelopes of single-variable convex piecewise cubic functions and multivariable gauge functions. In *Nonsmooth optimization and its applications*, volume 170 of *Internat. Ser. Numer. Math.*, pages 89–130. Birkhäuser/Springer, Cham, 2019.

[36] P. Rey and C. Sagastizábal. Dynamical adjustment of the prox-parameter in bundle methods. *Optimization*, 51(2):423–447, 2002.

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

[38] R. Rockafellar and R. Wets. *Variational analysis*. Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]. Springer-Verlag, Berlin, 1998.

[39] W. Sun, R. Sampaio, and M. Candido. Proximal point algorithm for minimization of DC function. *J. Comput. Math.*, 21(4):451–462, 2003.

[40] K. Yosida. *Functional analysis*. Die Grundlehren der Mathematischen Wissenschaften, Band 123. Academic Press, Inc., New York; Springer-Verlag, Berlin, 1965.