Epsilon subdifferential method and integrability*

Matey Konstantinov

Faculty of Mathematics and Informatics

Sofia University

5, James Bourchier Blvd.

1164 Sofia, Bulgaria

e-mail: mbkonstant@fmi.uni-sofia.bg

Nadia Zlateva

Faculty of Mathematics and Informatics

Sofia University

5, James Bourchier Blvd.

1164 Sofia, Bulgaria

e-mail: zlateva@fmi.uni-sofia.bg

June 14, 2022

Abstract

We develop a novel variant of the epsilon subdifferential method and use it to give a new proof of Moreau-Rockafellar theorem that a proper lower semicontinuous convex function on Banach space is determined up to a constant by its subdifferential.

Key words: convex function, epsilon subdifferential, epsilon subdifferential method

AMS Subject Classification: 52A41, 49J53, 26E15, 47H05

*Research supported by the Scientific Fund of Sofia University under grant 80-10-40/22.3.2021.

1 Introduction

Epsilon Subdifferential Method is well known and widely used for minimizing convex functions, see e.g. [2, 3]. In this note we develop a novel Epsilon Subdifferential Method (ESM). Let us outline it here.

ESM applies to a given proper convex lower semicontinuous function $f : X \to \mathbb{R}\cup\{+\infty\}$, defined on a Banach space X, such that $0 = f(0) = \min_{x\in X} f(x)$ with fixed in advance parameters $\varepsilon > 0$ and $\delta \in (0, \varepsilon)$.

Starting at arbitrary $x_0 \in \operatorname{dom} f$, for $i = 0, 1, \dots$

- if $0 \in \partial_\varepsilon f(x_i)$, then STOP;
- if $0 \notin \partial_\varepsilon f(x_i)$, for
$$\varphi_{x_i}(K) := \inf_{x\in X} F_{x_i}(K, x),$$
where
$$F_{x_i}(K, x) := f(x) - f(x_i) + \varepsilon + K\|x - x_i\|,$$
find $K_i > 0$ such that $\varphi_{x_i}(K_i) = 0$. Take any x_{i+1} satisfying
$$0 \le f(x_{i+1}) - f(x_i) + \varepsilon + K_i\|x_{i+1} - x_i\| \le \delta.$$

In the finite dimensional case $\delta = 0$ works, and ESM is much more simple.

The immediate estimate for the number of iterations n is $n \le \text{const}\,\varepsilon^{-1}$. But when f satisfies $f(x) \ge c\|x\|$ for all $x \in X$ and some $c > 0$, the parameter δ is appropriately chosen, and the starting point $x_0 \in \operatorname{dom}\partial f$, then the number of iterations n of ESM has the more precise estimate $n \le \text{const}\,\varepsilon^{-\frac{1}{2}}$, see Lemma 3.3. The proof relies on Lemma 2.3. Note that in this case, $n\varepsilon \le \text{const}\,\varepsilon^{\frac{1}{2}}$ which yields that $n\varepsilon$ tends to 0 as ε tends to 0. This is the key argument in the presented here new proof of the famous Moreau-Rockafellar Theorem, see e.g. [10, 11]:

Theorem 1.1. *Let X be a Banach space. Let g and h be proper lower semicontinuous convex functions from X to $\mathbb{R}\cup\{+\infty\}$. If*
$$\partial g \subset \partial h, \tag{1.1}$$
then
$$h = g + \text{const}.$$

This result has numerous important implications, see e.g. Section 3 of Phelps' book [9].

Let us make a short historical overview. The integrability of the subdifferential of proper lower semicontinuous convex function on Hilbert space is stated and proved first by Moreau in [7] by using Moreau-Yosida regularisation. The proof also works in reflexive Banach space as mentioned at p. 87 of [8]. The first complete proof in Banach space – that of Rockafellar in [11] – uses strong duality arguments. Another approach is to approximate the directional derivative and to reduce to the one-dimensional case. The latter was taken by Rockafellar in his original proof in [10]. Though there are some gaps in this proof, Taylor [12] fills them and provides a different proof, cf. [4]. The idea of directional derivative approximation/one dimensional reduction is most clearly outlined in the proof of Thibault [13]. A different proof using the mean-value theorem of Zagrodny is due to Thibault and Zagrodny [14], see also [15]. In [16] the result is proved by using regularization (and approximation) techniques which was the initial idea of Moreau.

In [6] Ivanov and Zlateva give a proof similar to the proof of the classical calculus theorem that a monotone function is Riemann integrable which uses neither duality nor explicit one-dimensional arguments. The main step in their proof is to show directly that a proper lower semicontinuous convex function on Banach space differs by a constant from the *Rockafellar function* (see [1]) of its subdifferential, see [6, Theorem 1.2]. The proof relies on a technical lemma [6, Lemma 3.3] proved by Ekeland variational principle.

Here we use the novel Epsilon Subdifferential Method (ESM) to prove in a different way the following

Theorem 1.2 (Rockafellar [10, 11], see also [6] Theorem 1.2)**.** *Let $g : X \to \mathbb{R}\cup\{+\infty\}$ be a proper lower semicontinuous convex function. Let $\bar{x} \in \operatorname{dom} \partial g$ and $\bar{p} \in \partial g(\bar{x})$. Then for all $x \in X$*

$$g(x) = g(\bar{x}) + R_{\partial g,(\bar{x},\bar{p})}(x),$$

where

$$R_{\partial g,(\bar{x},\bar{p})}(x) := \sup\Big\{ \sum_{i=0}^{n-1} \langle q_{i+1}, x_i - x_{i+1}\rangle : \tag{1.2}$$

$$x_0 = x,\ x_n = \bar{x},\ q_n = \bar{p},\ q_i \in \partial g(x_i), n \in \mathbb{N}\Big\}.$$

A distinctive feature of the new proof here is that it reveals the relationship between a natural optimization method and Moreau-Rockafellar Theorem. By use of ESM the sequences realizing supremum in (1.2) are kind of constructed.

Thereafter, the proof of Theorem 1.1 continues exactly as in [6]. That is why, we only sketch it here: it readily follows that (1.1) implies

$$g(x) - g(\overline{x}) \leq h(x) - h(\overline{x})$$

for any $\overline{x} \in \operatorname{dom} \partial g$ and all $x \in X$. In particular, $g - h \equiv \text{const}$ on $\operatorname{dom} \partial g$. To conclude, we use lower semicontinuity of h and graphical density of points of subdifferentiability to g, i.e. that for any $\bar{x} \in \operatorname{dom} g$ and any $\varepsilon > 0$ there exists $x \in \operatorname{dom} \partial g$ such that $\|x - \bar{x}\| + |g(x) - g(\bar{x})| < \varepsilon$, see [5] and [4].

Let us also note that tools used in the proof had been known by 1970.

The rest of the paper is organized as follows. After a short Section 2 on notations and preliminaries, in Section 3 we dwell on some of the basic properties of the novel Epsilon Subdifferential Method (ESM). In the last Section 4 we give the proof of Theorem 1.2.

2 Preliminaries and notations

The notation used throughout the paper is standard. Usually $(X, \|\cdot\|)$ denotes a real Banach space, that is, complete normed space over $\mathbb{R}$. The dual space X^* of X is the Banach space of all continuous linear functionals p from X to $\mathbb{R}$. The natural norm of X^* is again denoted by $\|\cdot\|$. The value of $p \in X^*$ at $x \in X$ is denoted by $\langle p, x \rangle$.

The *effective domain* $\operatorname{dom} f$ of an extended real-valued function $f : X \to \mathbb{R} \cup \{+\infty\}$ is the set of points x where $f(x) \in \mathbb{R}$. The function f is *proper* if $\operatorname{dom} f \neq \varnothing$. It is *lower semicontinuous* if $f(\bar{x}) \leq \liminf_{x \to \bar{x}} f(x)$ for all $\bar{x} \in X$.

Let us recall that for $\varepsilon \geq 0$, the ε-subdifferential of a proper, convex and lower semicontinuous function $f : X \to \mathbb{R} \cup \{+\infty\}$ at $x \in \operatorname{dom} f$ is the set

$$\partial_\varepsilon f(x) = \{p \in X^* \,:\, -\varepsilon + \langle p, y - x \rangle \leq f(y) - f(x), \quad \forall y \in X\},$$

and $\partial_\varepsilon f = \varnothing$ on $X \setminus \operatorname{dom} f$. Of course, for $\varepsilon = 0$, $\partial_0 f(x)$ coincides with the subdifferential of f at x in the sense of Convex Analysis $\partial f(x)$. The *domain* $\operatorname{dom} \partial_\varepsilon f$ consists of all points $x \in X$ such that $\partial_\varepsilon f(x)$ is non-empty. But while $\partial f(x)$ could be empty, for $\varepsilon > 0$, the sets $\partial_\varepsilon f(x)$ are non-empty for any $x \in \operatorname{dom} f$. For any real numbers ε_1 and ε_2 such that $0 < \varepsilon_1 \le \varepsilon_2$ one has $\partial_{\varepsilon_1} f(x) \subset \partial_{\varepsilon_2} f(x)$ and $\partial f(x) = \bigcap_{\varepsilon > 0} \partial_\varepsilon f(x)$. Moreover, if $f, g : X \to \mathbb{R} \cup \{+\infty\}$ are two proper lower semicontinuous convex functions with $x \in \operatorname{dom} f \cap \operatorname{dom} g$ and one of them is continuous at x, then the following sum rule holds, see e.g. [15, Theorem 2.8.7],

$$\partial_\varepsilon (f+g)(x) = \bigcup \{\partial_{\varepsilon_1} f(x) + \partial_{\varepsilon_2} g(x) : \varepsilon_1 \ge 0,\ \varepsilon_2 \ge 0\ \varepsilon = \varepsilon_1 + \varepsilon_2\}.$$

We will use it further in its weaker form

$$\partial_\varepsilon (f+g)(x) \subset \partial_\varepsilon f(x) + \partial_\varepsilon g(x).$$

Brøndsted-Rockafellar Theorem saying that the graph of $\partial_\varepsilon f$ is close to the graph of ∂f is well known:

Theorem 2.1 (Brøndsted-Rockafellar [5])**.** *Let $f : X \to \mathbb{R} \cup \{+\infty\}$, be a proper, convex and lower semicontinuous function, let $\varepsilon > 0$ and $p \in \partial_\varepsilon f(x)$. Then there exists $q \in \partial f(z)$ such that*

$$\|z - x\| \le \sqrt{\varepsilon}, \text{ and } \|q - p\| \le \sqrt{\varepsilon}.$$

Another result of Brøndsted and Rockafellar [5] will also be used:

Proposition 2.2. *Let f be a proper lower semicontinuous convex function from a Banach space X into $\mathbb{R} \cup \{+\infty\}$. Then for all $x \in X$*

$$f(x) = \sup\{f(\bar{x}) + \bar{p}(x - \bar{x});\ (\bar{x}, \bar{p}) \in \operatorname{gph} \partial f\}. \tag{2.1}$$

To estimate the number of iteration of the novel ESM we prove the following lemma of its own interest.

Lemma 2.3. *Let $A > 0$, $B > 0$ and $\varepsilon > 0$ be real numbers. If there exist reals $a_1, \dots, a_n$ and $b_1, \dots, b_n$ satisfying the conditions*

$$a_i > 0,\ b_i > 0, \qquad i = 1, \dots, n, \tag{2.2}$$

$$a_i b_i \ge \varepsilon, \qquad i = 1, \dots, n, \tag{2.3}$$

$$\sum_{i=1}^{n} a_i \le A, \tag{2.4}$$

$$\sum_{i=1}^{n} b_i \le B, \tag{2.5}$$

then the inequality $n \leq \sqrt{\dfrac{AB}{\varepsilon}}$ *holds.*

Proof. From (2.2) and (2.3) follow the inequalities

$$b_i \geq \frac{\varepsilon}{a_i}, \quad i = 1, \ldots, n. \tag{2.6}$$

Summing the inequalities (2.6) for $i = 1, \ldots, n$ we get that by (2.5)

$$\sum_{i=1}^{n} \frac{\varepsilon}{a_i} \leq \sum_{i=1}^{n} b_i \leq B.$$

Hence, $\displaystyle\sum_{i=1}^{n} \frac{1}{a_i} \leq \frac{B}{\varepsilon}$ and, equivalently, $\dfrac{\varepsilon}{B} \leq \dfrac{1}{\displaystyle\sum_{i=1}^{n} \frac{1}{a_i}}$. Multiplying by n we get

$$\frac{n\varepsilon}{B} \leq \frac{n}{\displaystyle\sum_{i=1}^{n} \frac{1}{a_i}}. \tag{2.7}$$

From (2.4) it follows that

$$\frac{\displaystyle\sum_{i=1}^{n} a_i}{n} \leq \frac{A}{n}. \tag{2.8}$$

By (2.7), Cauchy inequality and (2.8) we get the following chain of inequalities

$$\frac{n\varepsilon}{B} \leq \frac{n}{\displaystyle\sum_{i=1}^{n} \frac{1}{a_i}} \leq \frac{\displaystyle\sum_{i=1}^{n} a_i}{n} \leq \frac{A}{n}$$

which yields that $n^2 \leq \dfrac{AB}{\varepsilon}$. Therefore, $n \leq \sqrt{\dfrac{AB}{\varepsilon}}$. □

3 Properties of the novel ESM

We have outlined the method in the introduction. In this section we will consider some of its properties. Throughout this section we work with a proper lower semicontinuous convex function $f : X \to \mathbb{R} \cup \{+\infty\}$, such that $\min_{x \in X} f(x) = f(0) = 0$, and fixed $\varepsilon > 0$, and $\varepsilon > \delta > 0$. Next lemma describes what happens at one of ESM iterations.

Lemma 3.1. *Let $x_0 \in \operatorname{dom} f$. The function $\varphi_{x_0} : \mathbb{R} \to \mathbb{R}$ defined by*

$$\varphi_{x_0}(K) := \inf_{x \in X} F_{x_0}(K, x),$$

where

$$F_{x_0}(K, x) := f(x) - f(x_0) + \varepsilon + K\|x - x_0\|,$$

is strictly monotone increasing and locally Lipschitz on $(0, \infty)$.

Assume in addition that $0 \notin \partial_\varepsilon f(x_0)$. Then

(i) there exists $K_0 > 0$ such that $\varphi_{x_0}(K_0) = 0$;

(ii) for any $x_1 \in X$ such that

$$0 \le f(x_1) - f(x_0) + \varepsilon + K_0\|x_1 - x_0\| \le \delta, \tag{3.1}$$

there is $p_1 \in \partial_\delta f(x_1)$ such that

$$K_0 \ge \|p_1\| - \delta, \tag{3.2}$$

and,

$$\langle p_1, x_1 - x_0 \rangle \le f(x_1) - f(x_0) + \varepsilon + \delta. \tag{3.3}$$

Moreover,

$$K_0 \le \min\{\|p\| : p \in \partial_\varepsilon f(x_0)\}, \tag{3.4}$$

and if $p_0 \in \partial_\delta f(x_0)$, then

$$\varepsilon \le (\|p_0\| - \|p_1\|)\|x_1 - x_0\| + \delta\left(2 + \frac{f(x_0)}{K_0}\right). \tag{3.5}$$

Proof. It is straightforward to prove the strict monotonicity and the local Lipschitz continuity of φ_{x_0} on $(0, \infty)$.

The existence of $K' > 0$ such that $\varphi_{x_0}(K') < 0$ follows by $0 \notin \partial_\varepsilon f(x_0)$, while the existence of $K'' > 0$ such that $\varphi_{x_0}(K'') > 0$ is a consequence of lower semicontinuity of f at x_0. Bolzano Theorem ensures the existence of $K_0 > 0$ with $\varphi_{x_0}(K_0) = 0$ which thanks to the strict monotonicity of φ_{x_0} on $(0, \infty)$ have to be unique.

Now, let $x_1 \in X$ be a point of δ-infimum of the function $F_{x_0}(K_0, \cdot)$, i.e. satisfying (3.1). Equivalently, $0 \in \partial_\delta \left(f(\cdot) - f(x_0) + \varepsilon + K_0 \| \cdot - x_0 \| \right)(x_1)$. By the weaker form of the sum rule for the δ-subdifferential, we have that there exist $p_1 \in \partial_\delta f(x_1)$, and $\xi_1 \in \partial_\delta K_0 \| \cdot - x_0 \| (x_1)$ such that $0 = p_1 + \xi_1$. Since $\xi_1 \in \partial_\delta K_0 \| \cdot - x_0 \| (x_1)$,

$$\langle \xi_1, x - x_1 \rangle \leq K_0 \|x - x_0\| - K_0 \|x_1 - x_0\| + \delta \leq K_0 \|x - x_1\| + \delta, \ \forall x \in X. \tag{3.6}$$

Hence,

$$|\langle \xi_1, x - x_1 \rangle| \leq K_0 \|x - x_1\| + \delta, \ \forall x \in X,$$

and (3.2) holds. From (3.6) it easily follows that

$$K_0 \|x_1 - x_0\| \leq \langle \xi_1, x_1 - x_0 \rangle + \delta = \langle p_1, x_0 - x_1 \rangle + \delta$$

which combined with the left inequality in (3.1) yields (3.3).

Take arbitrary $p \in \partial_\varepsilon f(x_0)$. By definition of the ε-subdifferential,

$$\langle p, x - x_0 \rangle \leq f(x) - f(x_0) + \varepsilon, \ \forall x \in X.$$

Hence,

$$\langle p, x - x_0 \rangle + K_0 \|x - x_0\| \leq f(x) - f(x_0) + \varepsilon + K_0 \|x - x_0\|, \ \forall x \in X,$$

$$\langle p, \frac{x - x_0}{\|x - x_0\|} \rangle + K_0 \leq \frac{f(x) - f(x_0) + \varepsilon + K_0 \|x - x_0\|}{\|x - x_0\|}, \ \forall x \in X, \ x \neq x_0,$$

$$K_0 + \inf_{x \in X, x \neq x_0} \langle p, \frac{x - x_0}{\|x - x_0\|} \rangle \leq \inf_{x \in X, x \neq x_0} \left(\frac{f(x) - f(x_0) + \varepsilon + K_0 \|x - x_0\|}{\|x - x_0\|} \right) = 0.$$

Finally, $K_0 \leq \|p\|$, and (3.4) holds.

Take any $p_0 \in \partial_\delta f(x_0)$. Using (3.2), and $\|x_1 - x_0\| \leq \dfrac{f(x_0)}{K_0}$ (which is an easy consequence of (3.1) and $\delta < \varepsilon$), we get that

$$
\begin{aligned}
\varepsilon &\leq f(x_0) - f(x_1) - K_0\|x_1 - x_0\| + \delta \\
&\leq \langle p_0, x_0 - x_1\rangle - \|p_1\|\|x_1 - x_0\| + \delta\|x_1 - x_0\| + 2\delta \\
&\leq \|p_0\|\|x_1 - x_0\| - \|p_1\|\|x_1 - x_0\| + \delta\left(\frac{f(x_0)}{K_0} + 2\right) \\
&= (\|p_0\| - \|p_1\|)\|x_1 - x_0\| + \delta\left(\frac{f(x_0)}{K_0} + 2\right),
\end{aligned}
$$

which is (3.5). The proof is then completed. □

In the context of the ESM, Lemma 3.1 ensures the existence of $K_i > 0$. As x_{i+1} can be taken any point of δ-minimum, i.e. such that

$$0 \leq f(x_{i+1}) - f(x_i) + \varepsilon + K_i\|x_{i+1} - x_i\| \leq \delta. \tag{3.7}$$

From the lemma we also have the existence of $p_{i+1} \in \partial_\delta f(x_{i+1})$ such that

$$K_i \geq \|p_{i+1}\| - \delta,\ i \geq 0, \tag{3.8}$$

$$\langle p_{i+1}, x_{i+1} - x_i\rangle \leq f(x_{i+1}) - f(x_i) + \varepsilon + \delta,\ i \geq 0, \tag{3.9}$$

as well as,

$$\varepsilon \leq (\|p_i\| - \|p_{i+1}\|)\|x_{i+1} - x_i\| + \delta\left(2 + \frac{f(x_i)}{K_i}\right),\ i \geq 1. \tag{3.10}$$

The next Lemma shows that ESM is finite.

Lemma 3.2. *ESM ends after a finite number of iterations n such that $n \leq \dfrac{f(x_0)}{\varepsilon - \delta} + 1$ at point x_{n-1} of ε-minimum of f.*

Proof. Let us assume the contrary, i.e. that the number of iterations satisfy $n > \dfrac{f(x_0)}{\varepsilon - \delta} + 1$ and fix such $n \in \mathbb{N}$. This means that ESM generates at least x_i, $i = 0, \ldots, n-1$ such that

$$0 \notin \partial_\varepsilon f(x_i), \quad i = 0, \ldots, n-2.$$

Then from (3.7) we will have that

$$f(x_i) - f(x_{i+1}) \geq \varepsilon + K_i\|x_{i+1} - x_i\| - \delta \geq \varepsilon - \delta > 0, \qquad i = 0, \ldots, n-2.$$

Summing the inequalities we obtain that

$$f(x_0) - f(x_{n-1}) = \sum_{i=0}^{n-2} (f(x_i) - f(x_{i+1})) \geq (n-1)(\varepsilon - \delta) > f(x_0),$$

hence $0 > f(x_{n-1})$ which contradicts to $f(x_{n-1}) \geq f(0) = 0$. □

It is possible to obtain a better estimate of the number of iteration for a strictly convex function with more precise choice of the parameter δ.

Lemma 3.3. *Let $f : X \to \mathbb{R}\cup\{+\infty\}$ be proper lower semicontinuous convex function satisfying $f(x) \geq 2c\|x\|$ for all $x \in X$ and some $c > 0$.*

Applied for f with $\varepsilon > 0$ and $\delta > 0$ such that

$$\delta \leq \frac{c}{2}, \quad \delta \leq 1, \quad \delta\left(1 + \frac{f(x_0)}{c}\right) \leq \frac{\varepsilon}{4}, \tag{3.11}$$

ESM ends after n iterations, and

$$\sum_{i=0}^{n-2} \|x_{i+1} - x_i\| \leq \frac{2f(x_0)}{c}. \tag{3.12}$$

Moreover, for the number of iterations n we have the estimation

$$n \leq 2\sqrt{\frac{f(x_0)(\|p_0\| + 1)}{c\varepsilon}} + 2, \tag{3.13}$$

where $p_0 \in \partial_\varepsilon f(x_0)$ is arbitrary.

Proof. Since $f(x) \geq 2c\|x\|$, it is easy to see that if $p \in \partial_\delta f(x)$, then

$$0 = f(0) \geq f(x) - \langle p, x\rangle - \delta \geq 2c\|x\| - \langle p, x\rangle - \delta$$

yields

$$\|p\|\|x\| \geq \langle p, x\rangle \geq 2c\|x\| - \delta. \tag{3.14}$$

We have three cases: (a) $\|p\| \geq 2c$; (b) $\|p\| < 2c$, and $\|x\| > \delta/c$, and (c) $\|p\| < 2c$, and $\|x\| \leq \delta/c$.

In case (b), by (3.14) we have

$$\|p\| \geq 2c - \frac{\delta}{\|x\|} > 2c - \frac{\delta c}{\delta} = c.$$

In case (c),

$$f(x) \leq \langle p, x\rangle + \delta \leq \|p\|\|x\| + \delta \leq 3\delta < \varepsilon,$$

and x should be a point of ε-minimum for f.

As x_i, $i = 0, \dots, n-2$, are not ε-minimum points for f, the latter implies, see (3.8), that

$$K_i \geq \|p_{i+1}\| - \delta \geq c - \delta \geq c - \frac{c}{2} = \frac{c}{2}.$$

To establish (3.12) we sum up inequalities (3.7) from 0 to $n-2$ to get that

$$f(x_{n-1}) - f(x_0) + (n-1)\varepsilon + \sum_{i=0}^{n-2} K_i\|x_{i+1} - x_i\| \leq (n-1)\delta.$$

Hence,

$$\sum_{i=0}^{n-2} K_i\|x_{i+1} - x_i\| + (n-1)(\varepsilon - \delta) \leq f(x_0) - f(x_{n-1}).$$

Since $K_i \geq \dfrac{c}{2}$ for all i in the above sum, and $\delta < \varepsilon$,

$$\frac{c}{2}\sum_{i=0}^{n-2} \|x_{i+1} - x_i\| \leq f(x_0),$$

and (3.12) holds.

Since $f(x_{i+1}) \leq f(x_i)$ for all i, see (3.7), we have that $f(x_i) \leq f(x_0)$ for all i. Using this and $K_i \geq \dfrac{c}{2}$ in (3.10) we obtain that

$$\varepsilon \leq (\|p_i\| - \|p_{i+1}\|)\|x_{i+1} - x_i\| + 2\delta\left(1 + \frac{f(x_0)}{c}\right),\ i \geq 1,$$

hence, having in mind the choice of δ,

$$\frac{\varepsilon}{2} \leq (\|p_i\| - \|p_{i+1}\|)\|x_{i+1} - x_i\|,\ i \geq 1. \tag{3.15}$$

To apply Lemma 2.3, set

$$a_i := \|p_i\| - \|p_{i+1}\|, \quad b_i := \|x_{i+1} - x_i\|, \quad i = 1, \dots, n-2.$$

From (3.15) we have that $a_i b_i \geq \varepsilon/2$, hence, $a_i > 0$, $b_i > 0$, $i = 1, \dots, n-2$. From (3.12)

$$\sum_{i=1}^{n-2} b_i \leq \frac{2f(x_0)}{c}.$$

On the other hand,

$$\sum_{i=1}^{n-2} a_i = \|p_1\| - \|p_{n-1}\| \leq \|p_1\| \leq K_0 + \delta \leq \|p_0\| + \delta \leq \|p_0\| + 1,$$

where $p_0 \in \partial_\varepsilon f(x_0)$ is arbitrary (see (3.4)).

Setting $A := \|p_0\| + 1$ and $B := \dfrac{2f(x_0)}{c}$ we have that the conditions of Lemma 2.3 hold. Hence,

$$n - 2 \leq \sqrt{\frac{2AB}{\varepsilon}} = 2\sqrt{\frac{f(x_0)(\|p_0\| + 1)}{c\varepsilon}}$$

and (3.13) holds. The proof is completed. □

Let us note that p_0 in (3.13) as an arbitrary element in $\partial_\varepsilon f(x_0)$ depends on ε. But when $x_0 \in \operatorname{dom} \partial f$, then p_0 could be taken in $\partial f(x_0)$ and in this case, the estimation (3.13) is of the type $n\sqrt{\varepsilon} \leq \text{const}$.

4 Proof of Theorem 1.2

We will establish first that $g(x) = g(\bar{x}) + R_{\partial g, (\bar{x}, \bar{p})}(x)$ for $x \in \operatorname{dom} \partial g$.

To prove that

$$g(x) - g(\bar{x}) \geq R_{\partial g, (\bar{x}, \bar{p})}(x) \tag{4.1}$$

is easy. Indeed, for any sequence $\{(x_i, q_i)\}_{i=1}^n \subset \operatorname{gph} \partial g$ with $x_0 = x$, $x_n = \bar{x}$, and $q_n = \bar{p}$, by the definition of subdifferential we have that

$$\langle q_{i+1}, x_i - x_{i+1} \rangle \leq g(x_i) - g(x_{i+1}), \quad i = 0, \dots, (n-1).$$

After summing these inequalities we immediately get

$$\sum_{i=0}^{n-1}\langle q_{i+1}, x_i - x_{i+1}\rangle \le g(x) - g(\bar{x})$$

and (4.1) follows.

To obtain that

$$g(x) - g(\bar{x}) \le R_{\partial g,(\bar{x},\bar{p})}(x)$$

it is enough for any fixed $\varepsilon' > 0$ to find a sequence $\{(x_i, q_i)\}_{i=1}^{n} \subset \operatorname{gph}\partial g$ such that $x_0 = x$, $x_n = \bar{x}$, $q_n = \bar{p}$, and

$$g(x) - g(\bar{x}) - \sum_{i=0}^{n-1}\langle q_{i+1}, x_i - x_{i+1}\rangle < \varepsilon'. \tag{4.2}$$

To this end we consider the function $f : X \to \mathbb{R} \cup \{+\infty\}$, defined as

$$f(x) := g(x + \bar{x}) - \langle \bar{p}, x\rangle - g(\bar{x}) + 2c\|x\|, \tag{4.3}$$

where $c \in (0, 1)$ is a fixed constant. It is easy to see that f is proper lower semicontinuous and convex, $f(0) = 0$, $0 \in \partial f(0)$, $f(x) \ge 2c\|x\|$ for all $x \in X$ and $\operatorname{dom}\partial f \equiv \operatorname{dom}\partial g - \bar{x}$. Set $x_0 := x$, take arbitrary $p_0 \in \partial g(x_0)$ and set $M := 4\left(\sqrt{\dfrac{f(x_0)(\|p_0\| + 1)}{c}} + 1\right)$.

Take $\varepsilon \in (0, c)$ such that $M\sqrt{\varepsilon} < \varepsilon'$ and then apply ESM for f with this ε and $\delta{>}0$ such that $\eta(\delta){<}\varepsilon/3$, where $\eta(\delta){:=}2\sqrt{\delta}\left(1{+}2c{+}\|p_0\|{+}\|\overline{p}\|{+}\dfrac{f(x_0)}{c}\right)$. It is easy to check that if δ is such that $\eta(\delta) < \varepsilon/3$, then δ satisfies (3.11). When such a δ is chosen, set $\eta := \eta(\delta)$.

Denote $y_0 := x - \bar{x}$. Observe that $p_0 \in \partial f(y_0)$.

Starting at $y_0' := y_0$ ESM generates a finite sequence $p_{i+1} \in \partial_\delta f(y_{i+1}')$, $i = 0, \dots, n-2$.

By the weaker version of the δ-subdifferential sum rule we have that

$$\partial_\delta f(\cdot) \subset \partial_\delta g(\cdot + \overline{x}) + \partial_\delta\langle -\bar{p}, \cdot\rangle + \partial_\delta 2c\|\cdot\|,$$

therefore,

$$p_{i+1} = q_{i+1}' - \overline{p}_{i+1} + \xi_{i+1}, \tag{4.4}$$

for some $q'_{i+1} \in \partial_\delta g(\cdot + \bar{x})(y'_{i+1})$, $\xi_{i+1} \in \partial_\delta 2c\|\cdot\|(y'_{i+1})$, and $\overline{p}_{i+1}$ such that $\|\overline{p}_{i+1} - \overline{p}\| \leq \delta$, $i = 0, \ldots, n-2$.

From (3.9) we have that

$$\langle p_{i+1}, y'_{i+1} - y'_i \rangle \leq f(y'_{i+1}) - f(y'_i) + \varepsilon + \delta, \quad i = 0, \ldots, n-2.$$

Summing these equalities and using that $\delta < \eta$, we get

$$\sum_{i=0}^{n-2} \langle p_{i+1}, y'_{i+1} - y'_i \rangle \leq f(y'_{n-1}) - f(y_0) + (n-1)(\varepsilon + \eta),$$

and from (4.4) we obtain that

$$\begin{aligned} \sum_{i=0}^{n-2} \langle q'_{i+1}, y'_{i+1} - y'_i \rangle \quad &\leq \quad \sum_{i=0}^{n-2} \langle \overline{p}_{i+1}, y'_{i+1} - y'_i \rangle + \sum_{i=0}^{n-2} \langle \xi_{i+1}, y'_i - y'_{i+1} \rangle \\ &+ \quad f(y'_{n-1}) - f(y_0) + (n-1)(\varepsilon + \eta). \end{aligned} \tag{4.5}$$

To estimate the right hand side of (4.5) we use, first, that

$$\begin{aligned} \sum_{i=0}^{n-2} \langle \overline{p}_{i+1}, y'_{i+1} - y'_i \rangle \quad &\leq \quad \langle \overline{p}, y'_{n-1} - y_0 \rangle + \delta \sum_{i=0}^{n-2} \|y'_{i+1} - y'_i\| \\ &\leq \quad \langle \overline{p}, y'_{n-1} - y_0 \rangle + 2\delta \frac{f(x_0)}{c} \\ &\leq \quad \langle \overline{p}, y'_{n-1} - y_0 \rangle + \eta, \end{aligned}$$

second, that $\xi_{i+1} \in \partial_\delta 2c\|\cdot\|(y'_{i+1})$, hence

$$\begin{aligned} \sum_{i=0}^{n-2} \langle \xi_{i+1}, y'_i - y'_{i+1} \rangle \quad &\leq \quad \sum_{i=0}^{n-2} \left(2c\|y'_i\| - 2c\|y'_{i+1}\| + \delta\right) \\ &= \quad 2c\|y_0\| - 2c\|y'_{n-1}\| + (n-1)\delta \leq 2c\|y_0\| + (n-1)\eta, \end{aligned}$$

and, third, that y'_{n-1} is an ε-minimum of f, hence $f(y'_{n-1}) \leq \varepsilon$.

Incorporating all these in (4.5) we obtain that

$$\sum_{i=0}^{n-2} \langle q'_{i+1}, y'_{i+1} - y'_i \rangle \leq \langle \bar{p}, y'_{n-1} - y_0 \rangle + 2c\|y_0\| - f(y_0) + \tag{4.6}$$

$$(n-1)(\varepsilon + 2\eta) + \varepsilon + \eta.$$

By Brøndsted-Rockafellar Theorem there exist $q_{i+1} \in \partial g(\overline{x} + y_{i+1})$ such that $\|q_{i+1} - q'_{i+1}\| \le \sqrt{\delta}$, and $\|y_{i+1} - y'_{i+1}\| \le \sqrt{\delta}$. Then

$$\langle q_{i+1}, y_{i+1} - y_i\rangle - \langle q'_{i+1}, y'_{i+1} - y'_i\rangle =$$

$$\langle q_{i+1} - q'_{i+1}, y_{i+1} - y_i\rangle + \langle q'_{i+1}, y_{i+1} - y_i - y'_{i+1} + y'_i\rangle \le$$
$$\|q_{i+1} - q'_{i+1}\|\|y_{i+1} - y_i\| + \|q'_{i+1}\|(\|y_{i+1} - y'_{i+1}\| + \|y_i - y'_i\|).$$

Since $\|p_{i+1}\| \le \|p_1\|$, $\forall i$, which follows from (3.10), and since $\|p_1\| \le \|p_0\|$, see (3.5), we easily derive that $\|q'_{i+1}\| \le 2\delta + 2c + \|\overline{p}\| + \|p_0\|$, $\forall i$. Using the latter and $\|y_{i+1} - y_i\| \le 2\sqrt{\delta} + \|y'_{i+1} - y'_i\|$ we obtain that

$$\langle q_{i+1}, y_{i+1} - y_i\rangle - \langle q'_{i+1}, y'_{i+1} - y'_i\rangle \le$$

$$\sqrt{\delta}(2\sqrt{\delta} + \|y'_{i+1} - y'_i\|) + 2\sqrt{\delta}(2\delta + 2c + \|\overline{p}\| + \|p_0\|) \le \eta + \sqrt{\delta}\|y'_{i+1} - y'_i\|.$$

Hence,

$$\sum_{i=0}^{n-2}\langle q_{i+1}, y_{i+1} - y_i\rangle - \sum_{i=0}^{n-2}\langle q'_{i+1}, y'_{i+1} - y'_i\rangle \le$$

$$(n-1)\eta + \sqrt{\delta}\sum_{i=0}^{n-2}\|y'_{i+1} - y'_i\| \le (n-1)\eta + 2\sqrt{\delta}\frac{f(x_0)}{c} \le (n-1)\eta + \eta.$$

Using this in (4.6), as well as $\sqrt{\delta}\|\overline{p}\| \le \eta$, and $\eta \le \varepsilon/3$, we get

$$\begin{aligned}\sum_{i=0}^{n-2}\langle q_{i+1}, y_{i+1} - y_i\rangle &\le \langle \bar{p}, y_{n-1} - y_0\rangle + 2c\|y_0\| - f(y_0)\\ &\quad +(n-1)(\varepsilon + 3\eta) + \varepsilon + 2\eta + \sqrt{\delta}\|\overline{p}\|\\ &\le \langle \bar{p}, y_{n-1} - y_0\rangle + 2c\|y_0\| - f(y_0) + 2n\varepsilon. \quad (4.7)\end{aligned}$$

But

$$f(y_0) = f(x - \bar{x}) = g(x) - \langle \bar{p}, y_0\rangle - g(\bar{x}) + 2c\|y_0\|,$$

see (4.3), which combined with (4.7) yields

$$\sum_{i=0}^{n-2}\langle q_{i+1}, y_{i+1} - y_i\rangle \le \langle \bar{p}, y_{n-1}\rangle + g(\overline{x}) - g(x) + 2n\varepsilon. \quad (4.8)$$

Now, let us denote $x_{i+1} := y_{i+1} + \bar{x}$, $i = 0, \ldots, n-2$. Then $q_{i+1} \in \partial g(x_{i+1})$, and $x_i - x_{i+1} = y_i - y_{i+1}$, $i = 0, \ldots, n-2$.

Setting $x_n = \bar{x}$ $y_n = 0$, and $q_n = \bar{p}$ from (4.8) we obtain that

$$\begin{aligned}
g(x)-g(\bar{x})-\sum_{i=0}^{n-1}\langle q_{i+1}, x_i-x_{i+1}\rangle &= g(x)-g(\bar{x})-\sum_{i=0}^{n-1}\langle q_{i+1}, y_i-y_{i+1}\rangle \\
&\leq \langle q_n, y_n-y_{n-1}\rangle + \langle \bar{p}, y_{n-1}\rangle + 2n\varepsilon \\
&= 2n\varepsilon \quad (\text{since } y_n = 0 \text{ and } q_n = \bar{p}) \\
&\leq 4\left(\sqrt{\frac{f(x_0)(\|p_0\|+1)}{c\varepsilon}}+1\right)\varepsilon \quad (\text{by (3.13)}) \\
&\leq 4\left(\sqrt{\frac{f(x_0)(\|p_0\|+1)}{c}}+1\right)\sqrt{\varepsilon} = M\sqrt{\varepsilon} \\
&< \varepsilon',
\end{aligned}$$

and (4.2) follows.

So far we have shown that $g(x) = g(\bar{x}) + R_{\partial g,(\bar{x},\bar{p})}(x)$ for $x \in \operatorname{dom}\partial g$.

Now, fix any $x \in X$ and a real number r such that $r < g(x)$. By Proposition 2.2 we can find $(y,p) \in \operatorname{gph}\partial g$ such that $r < g(y) + \langle p, x-y\rangle$.

Since $y \in \operatorname{dom}\partial g$ for a fixed $\varepsilon > 0$ we find a sequence $\{(x_i,q_i)\}_{i=2}^n \in \operatorname{gph}\partial g$ with $x_1 = y$, $x_n = \bar{x}$ and $q_n = \bar{p}$ such that

$$g(y) - g(\bar{x}) - \sum_{i=1}^{n-1}\langle q_{i+1}, x_i - x_{i+1}\rangle < \varepsilon.$$

Then,

$$r < g(\overline{x}) + \langle p, x-y\rangle + \sum_{i=1}^{n-1}\langle q_{i+1}, x_i - x_{i+1}\rangle + \varepsilon = g(\overline{x}) + \sum_{i=0}^{n-1}\langle q_{i+1}, x_i - x_{i+1}\rangle + \varepsilon,$$

where $q_1 := p$.

Since $r < g(x)$ and $\varepsilon > 0$ were arbitrary, the proof is completed. □

Acknowledgements. The authors are very grateful to Dr. Milen Ivanov, who had the idea of the novel epsilon subfifferential method and how to use it to prove Rockafellar's formula. His support in obtaining and shaping the results was invaluable.

References

[1] S. Bartz, H. H. Bauschke, J. M. Borwein, S. Reich, X. Wang, *Fitzpatrick functions, cyclic monotonicity and Rockafellar's antiderivative*, Nonlinear Analysis, Theory Methods and Applications **66**, 2007, No. 5, 1198–1223.

[2] D. Bertsekas, S. Mitter, *Steepest Descent For Optimization Problems With Nondifferentiable Cost Functionals*, In: Proc. 5th Annual Princeton Conference on Information Sciences and Systems, Princeton, N.J., 1971, 347–351.

[3] D. Bertsekas, S. Mitter, *A Descent Numerical Method for Optimization Problems with Nondifferentiable Cost Functionals*, SIAM Journal on Control **11**, 1973, No. 4, 637–652.

[4] J. M. Borwein, *A note on ε-subgradients and maximal monotonicity*, Pac. J. Math. **103**, 1982, 307–314.

[5] A. Brøndsted, R. T. Rockafellar, *On the subdifferentiability of convex functions*, Proc. Am. Math. Soc. **16** (1965), 605–611.

[6] M. Ivanov, N. Zlateva, *A new proof of the integrability of the subdifferential of a convex function on a Banach space*, Proc. Amer. Math. Soc. **136**, 2008, No 5, 1787–1793.

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

[8] J.-J. Moreau, *Fonctionnelles convexes*, Séminaire sur les équations aux dérivées partielles, Collège de France (1966–1967).

[9] R. R. Phelps, *Convex functions, monotone operators and differentiability*, 2nd ed. Lecture Notes in Mathematics. 1364. Berlin: Springer-Verlag (1993).

[10] R. T. Rockafellar, *Characterization of the subdifferentials of convex functions*, Pac. J. Math. **17**, 1966, 497–510.

[11] R. T. Rockafellar, *On the maximal monotonicity of subdifferential mappings*, Pac. J. Math. **33**, 1970, 209–216.

[12] P. D. Taylor, *Subgradients of a convex function obtained from a directional derivative*, Pac. J. Math. **44**, 1973, 739–747.

[13] L. Thibault, *Limiting convex subdifferential calculus with applications to integration and maximal monotonicity of subdifferential*, In: Constructive, experimental, and nonlinear analysis. Selected papers of a workshop, Limoges, France, September 22–23, 1999 (M. Théra ed.), Providence, RI: American Mathematical Society (AMS), publ. for the Canadian Mathematical Society. CMS Conf. Proc. **27**, 2000, 279–289.

[14] L. Thibault, D. Zagrodny, *Integration of subdifferentials of lower semicontinuous functions on Banach spaces*, J. Math. Anal. Appl. **189**, 1995, No. 1, 33–58.

[15] C. Zalinescu, *Convex analysis in general vector spaces*, Singapore: World Scientific (2002).

[16] N. Zlateva, *Integrability through infimal regularization*, Compt. rend. Acad. bulg. Sci. **68**, 2015, No. 5, 551–560.