The Equivalence of Several Basic Theorems for Subdifferentials

Qiji J. Zhu
Department of Mathematics and Statistics
Western Michigan University
Kalamazoo, MI 49008
e-mail:zhu@math-stat.wmich.edu

Abstract. Several different basic properties are used for developing a system of calculus for subdifferentials. They are a nonlocal fuzzy sum rule in [5, 25], a multidirectional mean value theorem in [7, 8], local fuzzy sum rules in [14, 15] and an extremal principle in [19, 21]. We show that all these basic results are equivalent and discuss some interesting consequences of this equivalence.

Keywords: Subdifferentials, mean value inequalities, local fuzzy sum rules, nonlocal fuzzy sum rules, extremal principles and Asplund spaces.

AMS (1991) subject classification: Primary 26B05.
1 Introduction

Smooth subdifferentials play important roles in nonsmooth analysis for two reasons. They characterize many important generalized differentials and results in terms of smooth subdifferentials often require very little technical assumptions. Currently there are several ways of developing a set of basic theorems for subdifferentials so that they can be conveniently applied to a wide range of problems. The difference lies on the starting point. Borwein, Treiman and Zhu [5] use the nonlocal fuzzy sum rule [25]; Clarke, Ledyeav, Stern and Wolenski [8] starts with the Clarke-Ledyeav multidirectional inequality [7]; Ioffe [15] begins with the local fuzzy sum rule and Mordukhovich and Shao [21] choose the Kruger-Mordukhovich extremal principle [17]. The main purpose of this note is to observe that all the aforementioned basic results are equivalent.

Combining this equivalent result and Fabian’s results on the equivalence of the fuzzy sum rule of Fréchet subdifferential for Lipschitz functions and the Asplund property [1] of a Banach space yields that the nonlocal fuzzy sum rule [25] and the Clarke-Ledyeav mean value inequality [7] for the Fréchet subdifferential also equivalent to the Asplund property of a Banach space which expands the list of equivalent characterizations of Asplund spaces given by Mordukhovich and Shao in [20]. Also, this equivalent result shows that several versions of the local fuzzy sum rules with different assumptions on the summand functions (see, e.g. [2, 6, 10, 14, 15, 16]) are actually equivalent.

We state our main results and comment on some of their applications in the next section. The proofs are contained in Section 3.

2 The Main Result and some of its Consequences

2.1 The main result

It turns out that the relations among the basic results discussed in the introduction become transparent when they are applied in product spaces. For any positive integer $N$ and Banach space $X$ we will use $X^N$ to denote the Euclidean product space of $N$ copies of $X$. Our main result is

**Theorem 2.1** Let $X$ be a Banach space. Then the following are equivalent:

a. for any positive integer $N$, $X^N$ has a $\beta$-local Lipschitz fuzzy sum rule;

b. for any positive integer $N$, $X^N$ has a $\beta$-nonlocal fuzzy sum rule;

c. for any positive integer $N$, $X^N$ has a $\beta$-multidirectional mean value inequality;

d. for any positive integer $N$, $X^N$ has a $\beta$-local fuzzy sum rule;

e. for any positive integer $N$, $X^N$ has a $\beta$-extremal principle.
Since [a]-[c] in Theorem 2.1 are equivalent we will say that a Banach space $X$ is $\beta$-subdifferential compatible if any of them holds for $X$. Now let us make precise the concepts used in Theorem 2.1.

Let $X$ be a Banach space with closed unit ball $B_X$ and with continuous real dual $X^*$. For a point $x \in X$ and a set $Y \subset X$, we denote $d(x,Y) := \inf\{\|x - y\| : y \in Y\}$ and $[x,Y] := \{x + t(y - x) : y \in Y, t \in [0,1]\}$. The diameter of points $x_n, n = 1, ..., N$ in $X$ is defined by $\text{diam}(x_1, ..., x_N) := \max\{\|x_n - x_m\| : n, m = 1, ..., N\}$. We use $\tilde{R}$ to denote the extended real line $\tilde{R} = \mathbb{R} \cup \{+\infty\}$. Let us recall the definition of smooth subdifferentials and normal cones.

**Definition 2.2** Let $f : X \rightarrow \tilde{R}$ be a lower semicontinuous function and $S$ a closed subset of $X$. We say $f$ is $\beta$-subdifferentiable and $x^*$ is a $\beta$-subderivative of $f$ at $x$ if there exists a locally Lipschitz $\beta$-smooth function $g$ such that $\nabla g(x) = x^*$ and $f - g$ attains a local minimum at $x$. We denote the set of all $\beta$-subderivatives of $f$ at $x$ by $D_\beta f(x)$. We define the $\beta$-normal cone of $S$ at $x$ to be $N_\beta(x,S) := D_\beta \delta_S(x)$ where $\delta_S$ is the indicator function of $S$ defined by $\delta_S(x) = 0$ for $x \in S$ and $+\infty$ otherwise.

In Definition 2.2 “$\beta$-smooth” represents various smoothness concepts of different strengths. It can be any of the bornological smoothness or s-Hölder smoothness concepts. A bornology $\beta$ of a Banach space $X$ is a family of closed bounded and centrally symmetric subsets of $X$ whose union is $X$, which is closed under multiplication by scalars and is directed upwards (that is, the union of any two members of $\beta$ is contained in some member of $\beta$). We will denote by $X^*_\beta$ the dual space of $X$ endowed with the topology of uniform convergence on $\beta$-sets. Given a function $g$ on $X$, we say that $g$ is $\beta$-differentiable at $x$ and has a $\beta$-derivative $\nabla g(x)$ if $g(x)$ is finite and

$$t^{-1}(g(x + tu) - g(x) - t\langle \nabla g(x), u \rangle) \rightarrow 0$$

as $t \rightarrow 0$ uniformly in $u \in V$ for every $V \in \beta$. We say that a function $g$ is $\beta$-smooth at $x$ if $\nabla g : X \rightarrow X^*_\beta$ is continuous in a neighborhood of $x$. Some well known examples are: Gateaux smooth ($\beta = G$): the function $g$ is Gateaux differentiable and $\nabla g$ is weak-star continuous in a neighborhood of $x$; Fréchet smooth ($\beta = F$): the function $g$ is Fréchet differentiable and $\nabla g$ is norm continuous in a neighborhood of $x$. Similarly, for $s \in (0,1]$, we say $g$ is $s$-Hölder smooth ($\beta = H(s)$) if the function $g$ is $s$-Hölder differentiable and $\nabla g$ is norm continuous in a neighborhood of $x$. In a Hilbert space, $D_{H(1)}$ is the proximal subderivative [22]. We refer to [4, 9, 23] for details. By shifting a constant we may always assume that the $\beta$-smooth function $g$ in Definition 2.2 satisfies $g(x) = f(x)$.

Definitions below are motivated by [5, 25] (the nonlocal sum rule), [7, 25] (the multidirectional mean value inequality), [2, 6, 13, 14, 15, 16] (the local fuzzy sum rule and the local Lipschitz fuzzy rule) and [17, 18, 20] (the extremal principle).
Definition 2.3 We say that $X$ has a $\beta$-nonlocal fuzzy sum rule if for lower semicontinuous functions $f_1, \ldots, f_N : X \to \hat{R}$ bounded below and any $\varepsilon > 0$, there exist $x_n, n = 1, \ldots, N$ and $x_n^* \in D_\beta f_n(x_n)$ satisfying \( \text{diam}(x_1, \ldots, x_N) < \varepsilon \),

\[
\text{diam}(x_1, \ldots, x_N) \cdot \max(\|x_1^*\|, \ldots, \|x_N^*\|) < \varepsilon, \tag{1}
\]

and

\[
\sum_{n=1}^{N} f_n(x_n) < \liminf_{\eta \to 0} \{ \sum_{n=1}^{N} f_n(x_n) : \text{diam}(x_1, \ldots, x_N) \leq \eta \} + \varepsilon \tag{2}
\]

such that

\[
0 \in \sum_{n=1}^{N} x_n^* + \varepsilon B_{X^*}.
\]

Definition 2.4 We say that $X$ has a multidirectional mean value theorem if for any nonempty, closed and convex subset $Y$ of $X$, any element $x \in X$ and any lower semicontinuous function $f : X \to \hat{R}$ bounded below on a neighborhood of $[x, Y]$ and

\[
r < \liminf_{\eta \to 0} \inf_{y \in Y + \eta B_X} f(y) - f(x),
\]

given $\varepsilon > 0$, there exist $z \in [x, Y] + \varepsilon B$ and $z^* \in D_\beta f(z)$ such that

\[
r < \langle z^*, y - x \rangle + \varepsilon \|y - x\| \quad \text{for all } y \in Y.
\]

Further, we can choose $z$ to satisfy

\[
f(z) < \liminf_{\eta \to 0, \|z\| + \eta B_X} f + |r| + \varepsilon.
\]

Definition 2.5 We say that $X$ has a local fuzzy sum rule if, for lower semicontinuous functions $f_1, \ldots, f_N : X \to \hat{R}$ such that $\sum_{n=1}^{N} f_n$ attains a local minimum at $x$ and satisfies, for some $\eta > 0$

\[
\inf_{y \in x + \eta B_X} \sum_{n=1}^{N} f_n(y) \leq \liminf_{\varepsilon \to 0} \{ \sum_{n=1}^{N} f_n(x_n) : \|x_n - x_m\| \leq \varepsilon, x_n, x_m \in x + \eta B_X, n, m = 1, \ldots, N \}, \tag{3}
\]

and any $\varepsilon > 0$, there exist $x_n, n = 1, \ldots, N$ satisfying $(x_n, f_n(x_n)) \in (x, f_n(x)) + \varepsilon B_{X \times \hat{R}}$ and $x_n^* \in D_\beta f_n(x_n)$ such that

\[
0 \in \sum_{n=1}^{N} x_n^* + \varepsilon B_{X^*}.
\]

Definition 2.6 We say that $X$ has a Lipschitz local fuzzy sum rule if condition (3) in Definition 2.5 is replaced by all but one of the functions are locally Lipschitz at $x$.  

4
Remark 2.7 Obviously condition (3) is satisfied if all but one of the functions $f_n$ are uniformly continuous in a neighborhood of $x$. In particular, a local fuzzy sum rule implies a local Lipschitz fuzzy sum rule.

Let $S_1$ and $S_2$ be closed sets in a Banach space $X$ and let $x \in S_1 \cap S_2$. Then $x$ is called a locally extremal point of the set system $\{S_1, S_2\}$ if there are a neighborhood $U$ of $x$ and a sequence $\{a_k\} \subset X$, such that $a_k \to 0$ and

$$(S_1 - a_k) \cap S_2 \cap U = \emptyset \forall k = 1, 2, \ldots$$ (4)

Definition 2.8. We say that $X$ has a $\beta$-extremal principle if, for every local extremal point $x$ of closed sets $S_1$ and $S_2$ in $X$ and any $\varepsilon > 0$ there exist $x_i \in S_i \cap (x + \varepsilon B_X)$ and $\xi_i \in N_{\beta}(x_i; S_i) + \varepsilon B_{X^*}$, $i = 1, 2$, such that

$$\|\xi_1\| = \|\xi_2\| = 1 \text{ and } \xi_1 + \xi_2 = 0.$$ (5)

2.2 A sufficient condition

Note that if $X$ has a $\beta$-smooth Lipschitz bump function then so does $X^N$ for each positive integer. It is known that for a Banach space with a $\beta$-smooth Lipschitz bump function a $\beta$-local fuzzy sum rule holds (see Borwein and Ioffe [2] and the refinement in Borwein and Zhu [6]), a $\beta$-nonlocal fuzzy sum rule and a $\beta$-multidirectional mean value inequality holds (see [25]). Thus, we have

Theorem 2.9 Let $X$ be a Banach space with a $\beta$-smooth bump function. Then $X$ is $\beta$ subdifferential compatible.

2.3 Characterizations for Asplund spaces

In this subsection when refer to a Fréchet-subdifferential we use the following limiting definition. Let $f : X \to \bar{R}$. The limiting Fréchet-subdifferential $f$ at $x \in \text{dom}(f)$ is the set

$$\{\xi \in X^* : \liminf_{\|h\| \to 0} \frac{f(x + h) - f(x) - \langle \xi, h \rangle}{\|h\|} \geq 0\}.$$ When $X$ has a Fréchet-smooth bump function this limiting definition and the viscosity Fréchet-subdifferential definition we used so far are the same [9] but they may differ in general. We note that although our equivalent result is proved for viscosity subdifferentials the proof is valid for the limiting Fréchet-subdifferential with obvious modifications. In [13] Fabian proved that a Banach space is Asplund if and only if $X$ has a Fréchet Lipschitz fuzzy sum rule. In [20] Mordukhovich and Shao showed that an extremal principle in terms of the Fréchet normal cone characterizes Asplund spaces. Observe if $X$ is Asplund then so is $X^N$ for each positive integer $N$. Combining Fabian’s and Mordukhovich and Shao’s results and Theorem 2.1 we obtain $X$ is Asplund if and only if $X$ is Fréchet subdifferential compatible. That is:
Theorem 2.10 Let $X$ be a Banach space. Then the following are equivalent:

a. $X$ is an Asplund space;

b. for any positive integer $N$, $X^N$ is a Asplund space;

c. for any positive integer $N$, $X^N$ has a Fréchet-local Lipschitz fuzzy sum rule;

d. for any positive integer $N$, $X^N$ has a Fréchet-nonlocal fuzzy sum rule;

e. for any positive integer $N$, $X^N$ has a Fréchet-multidirectional mean value inequality;

f. for any positive integer $N$, $X^N$ has a Fréchet-local fuzzy sum rule;

g. for any positive integer $N$, $X^N$ has a Fréchet-extremal principle;

Remark 2.11 Theorem 2.10 give three new characterizations d., e. and f. of Asplund spaces. It shows that for the purpose of discussing the Fréchet-subdifferential with a variational analysis approach the usual assumption that the underlying Banach space admit an equivalent Fréchet-smooth norm or a Fréchet-smooth bump function (associated with the use of smooth variational principles, see [4, 9]) can be weakened to Asplund property (every continuous convex function is densely Fréchet differentiable). We should note that the smooth variational principles still play a crucial role here because the genes of the proofs of these equivalent relations is the smooth variational principle combined with a separable reduction method in Fabian [13]. In [21] Mordukhovich and Shao discussed in detail calculus rules for a coderivative in Asplund spaces under certain additional constraint qualification conditions required by the limiting process in defining the coderivative. The equivalence theorem in this section shows that parallel calculus rules for Fréchet-subdifferentials and Fréchet-coderivatives can be established in Asplund spaces without qualification conditions.

2.4 Local fuzzy sum rules

Local fuzzy sum rules are very useful in discussing optimization problems and as a basic result for deriving other calculus rules for subdifferentials. They have been the focus for many researches. Ioffe first proved a local fuzzy sum rule in finite dimensional spaces in [14]. Infinite dimensional generalizations are discussed in [15] under the assumption that all but one summand functions are locally Lipschitz. The local Lipschitz assumption is weakened to uniform continuity by Deville and Haddad in [10]. Borwein and Ioffe [2] and Ioffe and Rockafellar [16] weakened the requirement further to a sequential form of the uniform lower semicontinuity condition. Then the slightly weaker topological form of the uniform lower semicontinuity (3) was used in Borwein and Zhu [6] where the fuzzy sum rule was also refined to contain an estimate of the sizes of the subderivatives involved. It is somewhat surprising that according to the equivalence theorem in Section 2 all these different versions of the fuzzy sum rules are equivalent. On the other hand examples in Deville and Ivanov [11] and
Vanderwerff and Zhu [24] show that local sum rule does not hold for lower semicontinuous functions without additional assumption while an example in [26] shows that the sufficient conditions mentioned above are not necessary.

3 Proof of Theorem 2.1

We will show (a) \( \Rightarrow \) (b) \( \Rightarrow \) (c) \( \Rightarrow \) (d) \( \Rightarrow \) (e) \( \Rightarrow \) (a). The implication (b) \( \Rightarrow \) (c) has been established in [25]; for proximal subderivatives the implication (c) \( \Rightarrow \) (d) has been established in [8] and for Fréchet subdifferentials implications (d) \( \Rightarrow \) (e) \( \Rightarrow \) (a) has been established in [20] in Asplund spaces. The methods for proving those special cases apply to the general case with minor modifications. The missing link is (a) \( \Rightarrow \) (b) which we establish below. For completeness we also include the proofs for the general case of the other implications.

3.1 (a) \( \Rightarrow \) (b)

Consider extended-valued function \( f \) and Lipschitz function \( \varphi \) on \( X^N \) defined by

\[
f(y) = f(y_1, \ldots, y_N) := \sum_{n=1}^{N} f_n(y_n)
\]

and

\[
\varphi(y) = \varphi(y_1, \ldots, y_N) := \sum_{n,m=1}^{N} \|y_n - y_m\|.
\]

Define, for any real number \( i > 0 \),

\[
w_i(y) := f(y) + i\varphi(y)
\]

and \( M_i := \inf w_i \). Then \( M_i \) is an increasing sequence. Note that, for each \( i \),

\[
M_i \leq \inf \left\{ f(y) + i\varphi(y) : \text{diam}(y_1, \ldots, y_N) \leq \eta \right\} \\
\leq \inf \left\{ f(y) + iN^2\eta : \text{diam}(y_1, \ldots, y_N) \leq \eta \right\} \\
\leq \inf \left\{ f(y) : \text{diam}(y_1, \ldots, y_N) \leq \eta \right\} + iN^2\eta
\]

Taking limits when \( \eta \to 0 \) we obtain

\[
M_i \leq \liminf_{\eta \to 0} \{ f(y) : \text{diam}(y_1, \ldots, y_N) \leq \eta \}.
\]

Thus, \( M_i \) converges. Let \( M = \lim_{i \to \infty} M_i \). For each \( i \), applying the Ekeland variational principle [12] to function \( w_i \), we obtain that there exist \( u^i = (u^i_1, \ldots, u^i_N) \) such that

\[
w_i(y) + \varepsilon \|y - u^i\|_{X^N} = f(y) + i\varphi(y) + \varepsilon \|y - u^i\|_{X^N}
\]
attains a minimum at \( y = u^i \) and
\[
w_i(u^i) < \inf w_i + 1/i < M + 1/i. \tag{8}
\]
Note that both \( i\varphi \) and \( \varepsilon \| \cdot \| \) are Lipschitz. By the Lipschitz fuzzy sum rule (a) there exist \( x^i, z^i \in u^i + (\varepsilon/i^2)B_{X^N} \) with \( f(x^i) \leq f(u^i) + \varepsilon/i^2 \) such that
\[
0 \in D_\beta f(x^i) + D_\beta(i\varphi)(z^i) + (\varepsilon/N)B_{X^N^*}. \tag{9}
\]
Let \( \xi^i = (\xi_1^i, ..., \xi_N^i) \in D_\beta f(x^i) \) and \( \zeta^i = (\zeta_1^i, ..., \zeta_N^i) \in D_\beta(i\varphi)(z^i) \) satisfy
\[
0 \in \xi^i + \zeta^i + (\varepsilon/N)B_{X^N^*}. \tag{10}
\]
Then \( \xi_n^i \in D_\beta f_n(x_n^i) \). Note that, for any \( h \in X \) and \( t > 0 \) we have \( \varphi(z_1^i + th, ..., z_N^i + th) = \varphi(z_1^i, ..., z_N^i) \). Therefore, \( \zeta^i = (\zeta_1^i, ..., \zeta_N^i) \in D_\beta(i\varphi)(z^i) \) implies that, for any \( h \in X \),
\[
\langle \zeta_1^i + ... + \zeta_N^i, h \rangle \leq \lim_{t \to 0^+} \frac{i\varphi(z_1^i + th, ..., z_N^i + th) - i\varphi(z_1^i, ..., z_N^i)}{t} = 0,
\]
that is,
\[
\zeta_1^i + ... + \zeta_N^i = 0.
\]
Thus,
\[
\xi_1^i + ... + \xi_N^i \in \varepsilon B_{X^N^*}.
\]
By the definition of \( M_i \) we have
\[
M_{i/2} \leq w_{i/2}(u^i) = w_i(u^i) - \frac{i}{2} \varphi(u^i) \leq M_i + \frac{1}{i} - \frac{i}{2} \varphi(u^i). \tag{11}
\]
Rewriting (11) as
\[
i\varphi(u^i) \leq 2(M_i - M_{i/2} + \frac{1}{i})
\]
yields
\[
\lim_{i \to \infty} \frac{1}{i} \sum_{n,m=1}^N \| u_n^i - u_m^i \| = 0. \tag{12}
\]
Since \( \| x^i - u^i \| \leq \varepsilon/i^2 \) we also have
\[
\lim_{i \to \infty} \frac{1}{i} \sum_{n,m=1}^N \| x_n^i - x_m^i \| = 0. \tag{13}
\]
Therefore,
\[
\lim_{i \to \infty} \text{diam}(x_1^i, ..., x_N^i) = 0.
\]
Since \( \varphi \) is a Lipschitz function with a Lipschitz constant \( N^2 \), \( \| \xi^i \| \leq iN^2 \). It follows from (13) that
\[
\lim_{i \to \infty} \text{diam}(x_1^i, \ldots, x_N^i) \cdot \max(\| \xi_1^i \|, \ldots, \| \xi_N^i \|) = 0.
\]
Thus,
\[
M \leq \liminf_{\eta \to 0} \left\{ \sum_{n=1}^N f_n(x_n) : \text{diam}(x_1, \ldots, x_N) \leq \eta \right\}
\]
\[
\leq \liminf_{n \to \infty} \sum_{i=1}^N f_n(x_n^i) \leq \liminf_{n \to \infty} \sum_{n=1}^N f_n(u_n^i)
\]
\[
= \liminf_{n \to \infty} w_i(u^i) \leq M,
\]
that is
\[
M = \liminf_{\eta \to 0} \left\{ \sum_{n=1}^N f_n(x_n) : \text{diam}(x_1, \ldots, x_N) \leq \eta \right\}.
\]
It remains to take \( x_n = x_n^i \) and \( \xi_n = \xi_n^i \), \( n = 1, \ldots, N \) for a sufficiently large \( i \).

3.2 (b) \( \Rightarrow \) (c)

1. A special case. We begin by considering the special case when
\[
\lim_{\eta \to 0} \inf_{y \in Y + \eta B_X} f(y) > f(x) \quad \text{and} \quad r = 0.
\]

Let \( \tilde{f} = f + \delta_{[x,Y]+\eta B_X} \). Then \( \tilde{f} \) is bounded below on \( X \). Fix a \( \tilde{h} \in (0, h/2) \) such that
\[
\inf_{y \in Y + 2\tilde{h} B_X} f(y) > f(x).
\]
Without loss of generality we may assume that
\[
\varepsilon < \min \left\{ \inf_{y \in Y + 2\tilde{h} B_X} f(y) - f(x), \tilde{h} \right\}.
\]

Applying (b) to \( f_1 = \tilde{f} \) and \( f_2 = \delta_{[x,Y]} \) we obtain that there exist \( z, u \) with \( \| z - u \| < \varepsilon \),
\[
z^* \in D_\beta \tilde{f}(z) = D_\beta f(z) \quad \text{and} \quad u^* \in N_\beta(u, [x,Y])
\]
satisfying
\[
\max(\| z^* \|, \| u^* \|) \cdot \| z - u \| < \varepsilon
\]
and
\[
f(z) < \lim_{\eta \to 0} \inf_{y \in [x,Y]+\eta B_X} f + \varepsilon \leq f(x) + \varepsilon
\]
such that
\[
\| z^* + u^* \| < \varepsilon.
\]
Since \([x, Y]\) is convex \(N_\beta(u, [x, Y])\) coincides with the normal cone of \([x, Y]\) at \(u\) in the sense of convex analysis. Thus, \(u^* \in N_\beta(u, [x, Y])\) implies that
\[
\langle u^*, w - u \rangle \leq 0, \quad \forall w \in [x, Y].
\] (17)
Combining (16) and (17) yields
\[
0 < \langle z^*, w - u \rangle + \varepsilon \|w - u\|, \quad \forall w \in [x, Y] \setminus \{u\}. \quad (18)
\]
Moreover, we must have \(d(u, Y) \geq \tilde{h}\) for otherwise we would have \(d(z, Y) \leq 2\tilde{h}\) and \(f(z) \geq \inf_{y \in Y + 2\tilde{h}B_x} f(y) > f(x) + \varepsilon\) which contradicts (15). Let \(u = x + \tilde{r}(\tilde{y} - x)\). Then \(\tilde{h} \leq \|u - \tilde{y}\| = (1 - \tilde{r})\|x - \tilde{y}\|\) implies \(1 - \tilde{r} > 0\). Clearly \(x \not\in Y\). For any \(y \in Y\) set \(w = y + \tilde{r}(\tilde{y} - y) \neq u\) in (18) yields
\[
0 < \langle z^*, y - u \rangle + \varepsilon \|y - u\|, \quad \forall y \in Y. \quad (19)
\]

2. The general case. We now turn to the general case. Consider \(X \times R\) with the euclidean product norm. Take an \(\varepsilon' \in (0, \varepsilon/2)\) small enough so that
\[
\lim_{\eta \to 0} \inf_{y \in Y + \eta B_x} f(y) - f(x) > r + \varepsilon'
\]
and define \(F(z, t) := f(z) - (r + \varepsilon')t\). Obviously \(F\) is lower semicontinuous on \(X \times R\) and is bounded below on \([x, 0), Y \times \{1\}] + hB_{X \times R}\). Moreover,
\[
\lim_{\eta \to 0} \inf_{Y \times \{1\} + \eta B_{X \times R}} F = \lim_{\eta \to 0} \inf_{Y + \eta B_x} f - (r + \varepsilon') > f(x) = F(x, 0).
\]
Since \(X \times R\) is a subspace of \(X \times X\) we can apply the special case proved above with \(f, x\) and \(Y\) replaced by \(F, (x, 0)\) and \(Y \times \{1\}\) to conclude that there exist \((z, s) \in [(x, 0), Y \times \{1\}] + \varepsilon B_{X \times R}\), \(z^* \in D_\beta f(z)\) satisfying
\[
f(z) - (r + \varepsilon's) < \lim_{\eta \to 0} \inf_{(w, t) \in [(x, 0), Y \times \{1\}] + \eta B_{X \times R}} (f(w) - (r + \varepsilon')t) + \varepsilon'
\]
i.e.,
\[
f(z) < \lim_{\eta \to 0} \inf_{(w, t) \in [(x, 0), Y \times \{1\}] + \eta B_{X \times R}} (f(w) - (r + \varepsilon')(t - s)) + \varepsilon'
\leq \lim_{\eta \to 0} \inf_{[x, Y] + \eta B_x} f + |r| + \varepsilon
\]
such that, for all \(y \in Y\),
\[
0 < \langle z^*, y - x \rangle - (r + \varepsilon') + \varepsilon' \sqrt{\|y - x\|^2 + 1}
\leq \langle z^*, y - x \rangle - r + \varepsilon'\|y - x\|
\leq \langle z^*, y - x \rangle - r + \varepsilon\|y - x\|.
\]
This completes the proof. ■
3.3 (c) ⇒ (d)

Let \( y = (y_1, \ldots, y_N) \in X^N \), \( f(y) := f_1(y_1) + f_2(y_2) + \ldots + f_N(y_N) : X^N \to \bar{R} \). Let \( \varepsilon \) be an arbitrary positive number. Since \( f_n \) are lower semicontinuous there exists a positive number \( \eta < \varepsilon \) such that \( u \in x + \eta B_X \) implies that \( f(u) > f(x) - \varepsilon/2N \). Let \( \eta < \eta/4N \) be as in the uniform lower semicontinuity condition (3) and \( Y := \{(y_1, y_2, \ldots, y_N) : y_n \in x + \eta B_X, n = 1, 2, \ldots, N\} \). Then

\[
\lim_{h \to 0} \inf_{u \in Y + h B_X} f(u) - f(\bar{x}) \geq 0
\]

where \( \bar{x} = (x, x, \ldots, x) \). Note that \( [\bar{x}, Y] \subset Y \). Applying the multidirectional mean value theorem (c) with \( \varepsilon' = \varepsilon \min(1/4N, \eta/2) \) and \( r = -\eta^2 \) yields that there exist \( z = (x_1, x_2, \ldots, x_N) \in Y + \varepsilon' B_X \) and \( z^* = (x_1^*, \ldots, x_N^*) \in D_B f(z) \), i.e., \( x_n^* \in D_B f_n(x_n) \) such that

\[
-\eta^2 - \varepsilon' \|y - \bar{x}\|_{X^N} < \langle z^*, y - \bar{x} \rangle, \quad \forall y \in Y
\]

and

\[
f(z) < \lim_{h \to 0} \inf_{u \in Y + h B_X} f(u) + \varepsilon'.
\]

Observing that \( \lim_{h \to 0} \inf_{u \in Y + h B_X} f(u) + \varepsilon' < f(\bar{x}) + \varepsilon' \varepsilon/4N \) and \( f_n(x_n) > f_n(x) - \varepsilon/4N \) we arrive at \( (x_n, f_n(x_n)) \in (x, f_n(x)) + \varepsilon B_{X \times R} \). For any \( w \in \eta B_X \), let \( y = (x + w, \ldots, x + w) \) in (20) we obtain

\[
-\eta^2 - \varepsilon' \sqrt{N} \|w\| < \langle \sum_{n=1}^{N} x_n^*, w \rangle.
\]

It follows that

\[
\| \sum_{n=1}^{N} x_n^* \| < \varepsilon' \sqrt{N} + \eta \leq \varepsilon.
\]

3.4 (d) ⇒ (e)

Let \( x \in S_1 \cap S_2 \) be an extremal point of \((S_1, S_2)\). Then there exists a neighborhood \( U \) of \( x \) such that, for any \( \varepsilon > 0 \), there exists an \( a \in X \) with \( \|a\| < \varepsilon^2/4 \) and \((S_1 + a) \cap S_2 = \emptyset\). We may choose a smaller \( \varepsilon \) if necessary to ensure that \( x + \frac{\varepsilon}{2} B_X \subset U \).

Consider function \( f : X^2 \to \bar{R} \) defined by

\[
f(x_1, x_2) := \|x_1 - x_2 + a\| + \delta_{(S_1 \cap U) \times (S_2 \cap U)}(x_1, x_2).
\]

Then

\[
\|a\| = f(x, x) < \frac{\varepsilon^2}{4} \leq \inf f + \frac{\varepsilon^2}{4}.
\]

By the Ekeland variational principle [12] there exists a point \((y_1, y_2) \in (x, x) + \frac{\varepsilon}{2} B\) such that

\[
(x_1, x_2) \to f(x_1, x_2) + \frac{\varepsilon}{2} \| (x_1, x_2) - (y_1, y_2) \|
\]

11
attains a minimum at \((x_1, x_2) = (y_1, y_2)\). Note that we must have \(y_1 \in S_1\) and \(y_2 \in S_2\). Therefore, \(\|y_1 - y_2 + a\| > 0\).

Since both \(\|x_1 - x_2 + a\|\) and \(\frac{\delta}{2}\|x_1, x_2\| - (y_1, y_2)\) are Lipschitz these two functions together with \(\delta_{(S_1, S_2)}(x_1, x_2)\) satisfy the condition of the local fuzzy sum rule (see Remark 2.7). Applying the local fuzzy sum rule (d) to the above functions, there exist \((x_1', x_2') \in (y_1, y_2) + \frac{\delta}{2} B \subset (x, x) + \varepsilon B\) such that

\[
0 \in D_{\beta}(\|x_1' - x_2' + a\|) + N_{\beta}(x_1', S_1) \times N_{\beta}(x_2', S_2) + \varepsilon B_{(X^2)^*}.
\]

We may choose \((x_1', x_2')\) closer to \((y_1, y_2)\) if necessary to make \(\|x_1' - x_2' + a\| > 0\). Then it is easy to deduce that all the elements of \(D_{\beta}(\|x_1' - x_2' + a\|)\) is of the form \((\xi, -\xi)\) with \(\|\xi\| = 1\). Thus, there exists \(\xi\) in the unit ball of \(X^*\) with

\[
(\xi, -\xi) \in N_{\beta}(x_1', S_1) \times N_{\beta}(x_2', S_2) + \varepsilon B_{(X^2)^*}.
\]

It remains to take \(\xi_1 = \xi\) and \(\xi_2 = -\xi\). \(\blacksquare\)

### 3.5 \((e) \Rightarrow (a)\)

Note that we need only to prove the Lipschitz local fuzzy sum rule for two functions because the general result can be deduced then by induction. Let \(f_1\) be a Lipschitz function with rank \(L\) and \(f_2\) be a lower semicontinuous function. Assume, without loss of generality that \(f_1 + f_2\) attains a minimum at 0 and \(f_1(0) = f_2(0) = 0\). Define

\[
S_1 := \{(x, \mu) \in X \times R : f_1(x) \leq \mu\}
\]

and

\[
S_2 := \{(x, \mu) \in X \times R : f_2(x) \leq -\mu\}.
\]

Then it is easy to check that \((0, 0)\) is an extremal point for \((S_1, S_2)\). For any \(\varepsilon > 0\), let \(\varepsilon' = \min\{\varepsilon/4(1+L^2), 1/4(1+L)\}\). By the extremal principle (e), there exist \((x_i, \mu_i) \in S_i \cap \varepsilon B_{X \times R}\), and \((\xi_i, -\lambda_i) \in X^* \times R, i = 1, 2\) such that

\[
(\xi_i, -\lambda_i) \in N_{\beta}((x_i, \mu_i), S_i),
\]

\[
\|(\xi_i, -\lambda_i)\| \geq 1 - \varepsilon' \geq 1/2, i = 1, 2
\]

and

\[
|\lambda_1 + \lambda_2| < \varepsilon' \quad \|\xi_1 + \xi_2\| < \varepsilon'.
\]

Let \(g\) be a \(\beta\)-smooth function on \(X \times R\) such that \(\delta_{S_1} - g\) attains a minimum 0 at \((x_1, \mu_1)\) and \(\nabla g(x_1, \mu_1) = (\xi_1, -\lambda_1)\). Note that we must have \(\delta_{S_1}(x_1, \mu_1) = 0\) so that \(g(x_1, \mu_1) = 0\). Since \(f_1\) is Lipschitz with rank \(L\), for any \(h \in X\) with \(\|h\| = 1\) we have \(\delta_{S_1}(x_1 + th, \mu_1 + tL) = 0\) for \(t > 0\) sufficiently small. Therefore,

\[
\frac{g(x_1 + th, \mu_1 + tL) - g(x_1, \mu_1)}{t} \leq 0.
\]

12
Taking limits as $t \to 0$ we arrive at
\[
\langle \xi, h \rangle - \lambda L \leq 0, \ \forall \|h\| = 1.
\] (22)

Combining (21) and (22) we get $\lambda_1 > \frac{1}{2(1+L)} > 0$ and $\lambda_2 < \varepsilon' - \lambda_1 < -\frac{1}{4(1+L)} < 0$ which forces $\mu_1 = f_1(x_1)$ and $\mu_2 = f_2(x_2)$. By Proposition 2.3 in [3] we have $x_1^* := \xi_1/\lambda_1 \in D_\beta f_1(x_1)$ and $x_2^* := -\xi_2/\lambda_2 \in D_\beta f_2(x_2)$. Then, observing that $\|x_1^*\| \leq L$ we obtain
\[
\|x_1^* + x_2^*\| = \left\| \frac{\xi_1}{\lambda_1} - \frac{\xi_2}{\lambda_2} \right\| = \left\| \frac{\xi_1(\lambda_1 + \lambda_2)}{\lambda_1 \lambda_2} - \frac{\xi_1 + \xi_2}{\lambda_2} \right\|
\leq \|x_1^*\| \cdot \left| \frac{\lambda_1 + \lambda_2}{\lambda_2} \right| + \frac{\|\xi_1 + \xi_2\|}{|\lambda_2|} \leq 4L(1+L)\varepsilon' + 4(1+L)\varepsilon' = 4(1+L)^2\varepsilon' < \varepsilon.
\]

References


