Journal of Convex Analysis
Volume 1 (1994), No.1, 47–60

Differentiability of Convex Functions on a Banach Space with Smooth Bump Function[1]

Li Yongxin, Shi Shuzhong
Nankai Institute of Mathematics
Tianjin, 300071, People's Repubilic of China

Received 15 November 1993

In this paper, we prove that if a Banach space X admits a Lipschitz β-smooth bump function, then any continuous convex function on a convex open subset Ω of X is β-differentiable in a dense G_δ subset of Ω.

In 1990, Preiss, Phelps and Namioka [15] proved that if a real Banach space X admits a β-smooth norm, then every continuous convex function f on a convex open subset Ω of X is generically β-differentiable; i.e., there exists a dense G_δ subset G of Ω such that f is β-differentiable at each point of G. This result with Haydon's counter example ([9], [10], see also [7], Chap. VII,) finally resolves the "generalized" famous Asplund conjecture (see Asplund [1] and Day [4], but only for Fréchet- and Gâteaux-differentiability), which asks whether the existence of an equivalent β-smooth norm is equivalent to the generic β-differentiablility of any continuous convex function.

The first partial answer to the Asplund conjecture was obtained in 1976 by Ekeland–Lebourg [8] who showed that if a Banach space X admits a Fréchet-smooth bump function, then X is an Asplund space, i.e., every continuous convex function on a convex open subset of X is generically Fréchet-differentiable. Since Haydon has constructed a Banach space, which admits a Lipschitz Fréchet-smooth (even C^1) bump function, but does not admit any equivalent Gâteaux-smooth norm, the necessary part of Asplund conjecture is false.

However, Haydon's counter example does not negate the following "modified" Asplund conjecture: it is whether the existence of a Lipschitz β-smooth bump function is equivalent to the generic β-differentiablility of any continuous convex function (cf. [7], Problem II.1). In this paper, we will prove the sufficient part of this "modified" Asplund conjecture: if a Banach space X admits a Lipschitz β-smooth bump function, then every continuous convex function on a convex open subset of X is generically β-differentiable.

The weaker result, where instead of a dense G_δ subset one has only a dense subset, was proved independently by Deville–Godefroy–Zizler [5], [6] and Li–Shi [12]. [5] also announced the main result of this paper for the case of the Gâteaux-differentiability with a sketch of proof. Our contribution is to give a different and general proof.

Our method is similar to that in [15] and in [12]. We adapt the construction used in [15]

[1] Research supported by the National Natural Science Foundation of China.

ISSN 0944-6532 / $ 2.50

and also invoke a Banach–Mazur game to discuss the generic properties. An essential difference between Preiss–Phelps–Namioka's proof and ours is that for a maximal monotone mapping $T: X \to 2^{X^*}$, we do not seek a pair (x_0, e_0) satisfying

$$\forall x^* \in T(x_0), \qquad \frac{\langle x^*, e_0\rangle}{\|e_0\|} = \sup_{e\in X\setminus\{0\}} \frac{\langle x^*, e\rangle}{\|e\|} = \|x^*\|^* = \text{const}, \tag{1}$$

but seek a pair (x_0, e_0) such that

$$\forall x^* \in T(x_0), \qquad \frac{\langle x^*, e_0\rangle}{\rho(e_0)} = \sup_{e\in X} \frac{\langle x^*, e\rangle}{\rho(e)} := \rho^*(x^*) = \text{const}, \tag{2}$$

where ρ^* is a "topologically equivalent" gauge function (positive sublinear function) of the norm on X^* and it is generated by a "β-well function" ρ on X, which is induced by a Lipschitz β-smooth bump function (see (5)). In particular, it makes the process of searching for pairs (x_0, e_0) easier than that in [15] and a subset of any β-neighbourhood of $x_0^* \in X^*$ may be characterized by ρ and ρ^* (see Proposition 6). We use also a method by which we proved a generalization of Ekeland's ε-variational principle and of its Borwein–Preiss smooth variant in [12].

Let us begin with some concepts about β-differentiability. All of them can be found in [14] or [2].

Let X be a real Banach space. A **bornology** of X, denoted by β, is a family of bounded subsets of X which forms a covering of X, i.e., $\bigcup_{S\in\beta} S = X$. An (extended real-valued) function $f: X \to \mathbb{R} \cup \{\pm\infty\}$ is **β-superdifferentiable** at $x \in X$ if f is finite at x and there is an $x^* \in X^*$ such that for any $S \in \beta$ and $t > 0$,

$$\begin{cases} f(x+th) - f(x) \le \langle x^*, th\rangle + t\varepsilon(t,h) \\ \text{with } \varepsilon(t,h) \to 0 \text{ as } t \to 0 \text{ uniformly with respect to } h \in S. \end{cases} \tag{3}$$

Such an x^* is called a **β-superderivative** of f at x and all β-superderivatives of f at x form a subset $\partial^\beta f(x)$ of X^*. For convenience, we put $\partial^\beta f(x) = \emptyset$ if $f(x) = -\infty$ and say that if $f(x) = +\infty$, then f is β-superdifferentiable at x and $\partial^\beta f(x) = X^*$. Similarly, we can define **β-subdifferentiable**, **β-subderivative** and $\partial_\beta f(x)$. If f is both β-superdifferentiable and β-subdifferentiable at x, then f is **β-differentiable** at x and $\partial^\beta f(x) = \partial_\beta f(x)$ is necessarily a singleton; its unique element is the **β-derivative** of f at x, denoted by $\nabla_\beta f(x)$. Moreover, if f is β-differentiable at each point of a subset D, we say that f is **β-smooth** in D. But a β-smooth equivalent norm means that this norm is β-smooth away from the origin.

A bornology β of X induces a topological vector structure on X^*, which is generated by the zero-neighbourhood base $\{D_{S,\varepsilon}\}$, where

$$\forall S \in \beta,\ \varepsilon > 0, \qquad D_{S,\varepsilon} = \{\, y^* \in X^* \mid \forall h \in S,\ \langle y^*, h\rangle < \varepsilon \,\}. \tag{4}$$

We denote X^* equipped with this topological vector structure by (X^*, τ_β), which is a locally convex space. When β is the collection of all bounded subsets of X, τ_β becomes the strong topology of X^*, and when β is the collection of all finite subsets of X, τ_β is reduced to the w^*-topology of X^*.

Let $\Omega \subset X$ be an open set. A set-valued mapping $T\colon \Omega \to 2^{X^*}$ is said to be **β-continuous** at x ([15]) if $T(x)$ is a singleton and upper semicontinuous at x from $(X, \|\cdot\|)$ to (X^*, τ_β); i.e., for any $S \in \beta$ and $\varepsilon > 0$, there is an open ball $B_{\delta,x} \subset \Omega$ centered at x such that $T(B_{\delta,x}) \subset T(x) + D_{S,\varepsilon}$. Notice that the notion used here is not the notion of continuity requiring upper and lower semicontinuity. Since $T(x)$ is required to be a singleton, the lower semicontinuity at x is a direct consequence of the upper one. But generally the image of T at x is not required to be a singleton.

Proposition 1. ([14], [15]) *A continuous convex function on a convex open subset Ω of X is β-differentiable at $x \in \Omega$ if and only if its subdifferential mapping $\partial f\colon X \to 2^{X^*}$ is β-continuous at x.*

Recall that the subdifferential mapping of a lower semicontinuous convex function on a Banach space X is maximal monotone ([14]); we reduce our problem to showing the generic β-continuity of a maximal monotone mapping on X.

The basic assumption of this paper is as follows:

$$(H) \quad \begin{cases} \text{There is a Lipschitz } \beta\text{-smooth bump function } \nu\colon X \to \mathbb{R}_+;\text{ i.e.} \\ \nu \text{ is nonconstant, Lipschitz, } \beta\text{-smooth and zero outside some ball.} \end{cases}$$

Proposition 2. *If a Banach space X satisfies (H), then one has:*

$$(H') \quad \begin{cases} \text{There exists a Lipschitz } \beta\text{-superdifferentiable function} \\ \mu\colon X \to [0,1] \text{ such that } \mu(0) = 0 \textit{ and } \Big[\|x\| \ge 1\Big] \Longrightarrow \Big[\mu(x) = 1\Big]. \end{cases}$$

Proof. Set $\alpha := \sup_{x\in X} \nu(x)$ and $R > \sup\{\|x\| \mid \nu(x) > 0\}$. Then, $\alpha > 0$ and $R > 0$. Without loss of generality, we can suppose that $\nu(0) \ge \alpha/2$ and $R = 1$. Otherwise, we replace ν by $\nu_1(x) = \nu(2Rx + x_0)$, where x_0 satisfies $\nu(x_0) \ge \alpha/2$. Take an increasing function $h \in C^\infty(\mathbb{R}_+; [0,1])$ such that

$$h(t) = \begin{cases} 1, & \text{when } t \ge \alpha/2; \\ 0, & \text{when } t = 0. \end{cases}$$

Then the function $\mu(x) = 1 - h(\nu(x))$ satisfies (H'). □

Actually, we need only a Lipschitz β-superdifferentiable function μ such as in (H') instead of a Lipschitz β-differentiable bump function (H). It means that we need only the β-subdifferentiability of ν in (H).

We say that ρ is a **β-well function** on X if $\rho\colon X \to [1, +\infty]$ is β-superdifferentiable and continuous, and satisfies

$$\rho(0) < +\infty \ \text{ and } \Big[\|x\| \ge 1\Big] \Longrightarrow \Big[\rho(x) = +\infty\Big], \tag{5}$$

where, according to our convention, if $\rho(x) = +\infty$, then ρ is considered to be β-superdifferentiable at x; and the continuity of ρ is for the norm topology on X and the usual topology on $[1, +\infty]$.

Proposition 3. *If (H') holds, then there exists a β-well function on X.*

Proof. Take a real function $g: [0, 1) \to [1, +\infty)$ which is an increasing C^∞-function with $g(0) = 1$ and $\lim_{t\to 1} g(t) = +\infty$. Set $g(1) := +\infty$. Then, the composite function of g with μ in (H'), $\rho_0 = g \circ \mu: X \to [1, +\infty]$ is a β-well function on X. In fact, we need only to show that ρ_0 is β-superdifferentiable at $x \in X$ with $\mu(x) < 1$. Take $S \in \beta$ with $M := \sup_{h\in S} \|h\| < +\infty$. Since μ is Lipschitz and β-superdifferentiable, and g is differentiable in $[0, 1)$ with $g'(\mu(x)) \geq 0$, we have that for $t > 0$,

$$\begin{aligned} & g(\mu(x + th)) - g(\mu(x)) \\ &= g'(\mu(x))(\mu(x + th) - \mu(x)) + o(|\mu(x + th) - \mu(x)|) \\ &\leq g'(\mu(x))(\langle x^*, th\rangle + t\varepsilon_1(t, h)) + LMt\varepsilon_2(t, h) \qquad \text{for any } x^* \in \partial^\beta \mu(x) \\ &= \langle g'(\mu(x))x^*, th\rangle + t\varepsilon_0(t, h), \end{aligned}$$

where L is a Lipschitz constant of μ and $\varepsilon_i(t, h)$, $i = 0, 1, 2$, tend to 0 as $t \to 0$ uniformly with respect to $h \in S$. So ρ_0 is β-superdifferentiable at x. □

The β-well function ρ that we will construct to get (2) will be of the form of ρ_∞ in the following proposition.

Proposition 4. *Let ρ_0 be a β-well function on X, μ be defined as in (H'), $\mu_n(x) = \mu(nx)/2^n$, $n = 1, 2, \ldots$, and $\{e_n\}_{n=1}^\infty \subset X$ be a sequence. Then*

$$\rho_n(x) := \rho_0(x) + \sum_{k=1}^{n} \mu_k(x - e_k), \quad n = 1, 2, \ldots, \tag{6}$$

and

$$\rho_\infty(x) := \rho_0(x) + \sum_{n=1}^{\infty} \mu_n(x - e_n), \tag{7}$$

are all β-well functions on X.

Proof. In fact, $\mu_n: X \to [0, 1/2^n]$ satisfies:

$$\begin{cases} \text{i) } \mu_n(0) = 0 \text{ and } \Big[\|x\| \geq 1/n\Big] \Longrightarrow \Big[\mu_n(x) = 1/2^n\Big]; \\ \text{ii) } \mu_n \text{ is } \beta\text{-superdifferentiable everywhere;} \\ \text{iii) } \mu_n \text{ is Lipschitz of rank } nL/2^n, \text{ where } L \text{ is a Lipschitz constant of } \mu. \end{cases} \tag{8}$$

Hence, from (6), ρ_n is obviously a β-well function. On the other hand, $\sum_{n=1}^\infty \mu_n(x - e_n)$ converges uniformly, and so, is continuous in x. Since μ_n is Lipschitz of rank $nL/2^n$, we have that

$$\begin{aligned}&\sum_{n=1}^{\infty}\mu_n(x-e_n+th)-\sum_{n=1}^{\infty}\mu_n(x-e_n)\\ \le &\sum_{n=1}^{N}\mu_n(x-e_n+th)-\sum_{n=1}^{N}\mu_n(x-e_n)+tL\|h\|\sum_{n=N+1}^{\infty}n/2^n.\end{aligned}$$

As any finite sum of β-superdifferentiable functions is also β-superdifferentiable, we can use the Weierstrass M-test to show that $\sum_{n=1}^{\infty}\mu_n(x-e_n)$ is β-superdifferentiable too and so is ρ_∞. □

Now for a β-well function ρ, we define a gauge function $\rho^*: X^* \to \mathbb{R}$ on X^* by

$$\forall x^* \in X^*, \qquad \rho^*(x^*) := \sup_{e\in X}\frac{\langle x^*, e\rangle}{\rho(e)}. \tag{9}$$

Obviously, for two β-well functions ρ_0 and ρ_1,

$$\Big[\rho_1 \ge \rho_0\Big] \Rightarrow \Big[\rho_1^* \le \rho_0^*\Big]. \tag{10}$$

Proposition 5. *For a gauge function ρ^* on X^* defined by (9) there exists $\varepsilon_0 \in (0,1)$ such that*

$$\forall x^* \in X^*, \qquad (1-\varepsilon_0)\|x^*\|^* \le \rho^*(x^*) \le \|x^*\|^*. \tag{11}$$

In particular, if ρ is symmetric, then ρ^* is an equivalent norm on $(X^*, \|\cdot\|^*)$.

Proof. Since $\rho(x) = +\infty$ whenever $\|x\| \ge 1$ and $\rho \ge 1$, we obtain that

$$\begin{aligned}\rho^*(x^*) &= \sup_{e\in X}\frac{\langle x^*, e\rangle}{\rho(e)} = \sup_{\|e\|\le 1}\frac{\langle x^*, e\rangle}{\rho(e)}\\ &\le \sup_{\|e\|\le 1}\langle x^*, e\rangle = \|x^*\|^*.\end{aligned}$$

On the other hand, since ρ is continuous and $\rho(0) \in [1,+\infty)$, there exists a $\delta \in (0, 2\rho(0))$ such that $\rho(x) \in [1, 2\rho(0)]$ whenever $x \in B_\delta = \{\, x \in X \mid \|x\| \le \delta \,\}$. Hence,

$$\begin{aligned}\rho^*(x^*) &= \sup_{e\in X}\frac{\langle x^*, e\rangle}{\rho(e)} \ge \sup_{e\in B_\delta}\frac{\langle x^*, e\rangle}{\rho(e)}\\ &\ge \frac{1}{2\rho(0)}\sup_{e\in B_\delta}\langle x^*, e\rangle \ge \frac{\delta}{2\rho(0)}\|x^*\|^*,\end{aligned}$$

i.e., (11) holds for $1-\varepsilon_0 = \delta/2\rho(0)$. □

Our main result is based on the following two simple facts.

Proposition 6. *Let ρ be a β-well function on X, $e_0 \in X$ with $\rho(e_0)$ finite. If there exists $x_0^* \in X^*$ such that*

$$c := \frac{\langle x_0^*, e_0\rangle}{\rho(e_0)} = \sup_{e\in X} \frac{\langle x_0^*, e\rangle}{\rho(e)} = \rho^*(x_0^*) > 0, \tag{12}$$

then

(i) *ρ is β-differentiable at e_0 and*

$$x_0^* = c\nabla_\beta \rho(e_0); \tag{13}$$

(ii) *for any $S \in \beta$ and any $\varepsilon > 0$, there exists $\delta > 0$ such that*

$$D(\rho, e_0, x_0^*, \delta) := \left\{ x^* \in X^* \,\middle|\, c - \delta < \frac{\langle x^*, e_0\rangle}{\rho(e_0)} \le \rho^*(x^*) < c + \delta \right\} \tag{14}$$

$$\subset x_0^* + D_{S,\varepsilon}. \tag{15}$$

Proof. (i) By (12), we have

$$\forall e \in X, \qquad c\rho(e) \ge \langle x_0^*, e\rangle.$$

Joining up with (12), it follows

$$\forall e \in X, \qquad c\rho(e) - c\rho(e_0) \ge \langle x_0^*, e - e_0\rangle.$$

Hence, $x_0^* \in c\partial_\beta \rho(e_0)$. On the other hand, from the definition of β-well function (5), $\partial^\beta \rho(e_0)$ is nonempty, and then, $\partial^\beta \rho(e_0)$ must be a singleton. Thus, (13) holds.

(ii) Suppose that $S \in \beta$ and $\varepsilon > 0$. From (13), we can take a fixed $t_0 > 0$ such that

$$\forall h \in S, \qquad \rho(e_0 + t_0 h) \le \rho(e_0) + \frac{1}{c}\langle x_0^*, t_0 h\rangle + \frac{\varepsilon t_0}{2c}. \tag{16}$$

If $x^* \in D(\rho, e_0, x_0^*, \delta)$ with a sufficiently small $\delta > 0$, then from (9), the definition (14) of $D(\rho, e_0, x_0^*, \delta)$ and (16), we have that

$$\begin{aligned} \langle x^*, e_0 + t_0 h\rangle &\le \rho^*(x^*)\rho(e_0 + t_0 h) \le (c+\delta)\rho(e_0 + t_0 h) \\ &\le (c+\delta)\rho(e_0) + \frac{c+\delta}{c}\langle x_0^*, t_0 h\rangle + \frac{c+\delta}{2c}\varepsilon t_0, \end{aligned}$$

and then,

$$\begin{aligned} \langle x^* - x_0^*, t_0 h\rangle &\le (c+\delta)\rho(e_0) - \langle x^*, e_0\rangle + \frac{\delta}{c}\langle x_0^*, t_0 h\rangle + \frac{c+\delta}{2c}\varepsilon t_0 \\ &\le (c+\delta)\rho(e_0) - (c-\delta)\rho(e_0) + \frac{\delta}{c}\langle x_0^*, t_0 h\rangle + \frac{c+\delta}{2c}\varepsilon t_0. \end{aligned}$$

Hence, for a sufficiently small $\delta > 0$, we obtain that

$$\forall h \in S, \qquad \langle x^* - x_0^*, h\rangle \le \frac{2\delta}{t_0}\rho(e_0) + \frac{\delta}{c}\langle x_0^*, h\rangle + \frac{c+\delta}{2c}\varepsilon < \varepsilon;$$

i.e., (15) holds. □

To obtain the desired generic properties, we have to introduce a Banach–Mazur game ([11]) on an open subset Ω of X. Let W be a subset of Ω. A **Banach–Mazur game** with the objective subset W on Ω is a nested open subset (of Ω) sequence $U_1 \supset V_1 \supset U_2 \supset V_2 \supset \cdots \supset U_n \supset V_n \supset \cdots$, where U_n and V_n are chosen by player A and player B alternatively. We say that B is the **winner** if $\bigcap_{n=1}^{\infty} V_n \subset W$, where $\bigcap_{n=1}^{\infty} V_n$ is not necessarily nonempty. A **strategy** for B is a sequence of maps f_n such that for any n, f_n is defined for every $U_1, U_2, \ldots, U_n$ chosen by A, and $f_n(U_1, U_2, \ldots, U_n)$ is an open subset contained in U_n. A strategy for B is a **winning strategy** if for any of A's choices of $\{U_n\}$ such that $U_1 \supset f_1(U_1) \supset U_2 \supset f_2(U_1, U_2) \supset \cdots \supset U_n \supset f_n(U_1, U_2, \cdots, U_n) \supset \cdots$, B is the winner.

Proposition 7. ([11]) *(Banach–Mazur Game Theorem) The objective subset W contains a dense G_δ subset of Ω if and only if B has a winning strategy.*

Finally, we need the following properties of maximal monotone mappings $T: X \to 2^{X^*}$.

Proposition 8. *Let $T: X \to 2^{X^*}$ be a maximal monotone mapping. If $\Omega = \operatorname{int}\{ x \in X \mid T(x) \neq \emptyset \}$ is nonempty, then*

(i) *T is locally bounded on Ω;*

(ii) *T is upper semicontinuous from $(X, \|\cdot\|)$ to (X^*, w^*) on Ω;*

(iii) *let $e \in X$, U be an open subset of Ω and $\bar{x} \in U$; if for some $\bar{x}^* \in T(\bar{x})$ and some $b \in \mathbb{R}$, $\langle \bar{x}^*, e \rangle > b$, then there exists an open subset $V \subset U$ such that*

$$\forall x^* \in T(V), \qquad \langle x^*, e \rangle > b. \tag{17}$$

Proof. i) and ii) are well-known ([14]). iii) has been posed in [15] in a little weaker form. We show iii) as follows. Take $t > 0$ such that $\bar{x} + te \in U$. By the monotonicity of T for any $x^* \in T(\bar{x} + te)$ we have $\langle x^*, e \rangle \geq \langle \bar{x}^*, e \rangle > b$, i.e., $T(\bar{x} + te) \subset \{ y^* \in X^* \mid \langle y^*, e \rangle > b \}$, which means that $\{ y^* \in X^* \mid \langle y^*, e \rangle > b \}$ is a w^*-open neighborhood of $T(\bar{x} + te)$. From ii), there exists an open neighbourhood V of $\bar{x} + te$, $V \subset U$, such that $T(V) \subset \{ y^* \in X^* \mid \langle y^*, e \rangle > b \}$. □

Now we deal with our main theorem.

Theorem 1. *Assume that X admits a Lipschitz β-differentiable bump function, i.e., (H) holds. If $T: X \to 2^{X^*}$ is a maximal monotone mapping on X and $\Omega := \operatorname{int}\{ x \in X \mid T(x) \neq \emptyset \}$ is nonempty, then T is β-continuous in a dense G_δ subset of Ω.*

Proof. By using Proposition 7, we suppose that player A and player B are playing a Banach–Mazur game on Ω with the objective subset

$$W = \{ x \in \Omega \mid T \text{ is } \beta\text{-continuous at } x. \} \tag{18}$$

It means that given A's choice of $\{U_n\}$, we will find B's choice $\{V_n\}$ such that $\bigcap_{n=1}^{\infty} V_n \subset W$; i.e., if $y_\infty \in \bigcap_{n=1}^{\infty} V_n$, then T is a singleton and β-continuous at y_∞. For proving that

$T(y_\infty)$ is a singleton, we will construct a β-well function of the form ρ_∞ in Proposition 3 and show that for $e_\infty = \lim_{n\to\infty} e_n$, we have

$$\forall y^*_\infty \in T(y_\infty), \quad \frac{\langle y^*_\infty, e_\infty\rangle}{\rho_\infty(e_\infty)} = \rho^*_\infty(y^*_\infty) = \text{const},$$

from which, by using Proposition 6 i), it follows $y^*_\infty = \text{const}\nabla_\beta\rho_\infty(e_\infty)$. And then, by using Proposition 6 ii), we will show that for any $S \in \beta$ and any $\varepsilon > 0$, there exists a sufficient large n such that $T(V_n) \subset T(y_\infty) + D_{S,\varepsilon}$, which implies the β-continuity of T at y_∞.

We begin this process. Firstly, we take a β-well function ρ_0 on X, the existence of which is due to Proposition 2 and 3. ρ^*_0 is generated as in (9) by ρ_0 and due to Proposition 5, we may suppose that $\varepsilon_0 \in (0,1)$ satisfies

$$\forall x^* \in X^*, \qquad (1-\varepsilon_0)\|x^*\|^* \le \rho^*_0(x^*) \le \|x^*\|^*. \tag{19}$$

For any U_1 chosen by A, due to Proposition 8 i), B can choose $V_1 \subset U_1$ such that T is bounded in V_1. Then the playing is going on with $U_2 \subset V_1$. Set

$$s_0 = \sup_{x^*\in T(U_2)} \rho^*_0(x^*) < +\infty. \tag{20}$$

Without loss of generality, we suppose that $s_0 > 0$; otherwise, the proof is trivial. From (9) and (20), there exist $\bar{x} \in U_2$, $\bar{x}^* \in T(\bar{x})$ and $e_1 \in X$ such that

$$\langle \bar{x}^*, e_1\rangle > \rho_0(e_1)(1-\varepsilon_0)s_0. \tag{21}$$

Due to iii) of Proposition 8, B can choose $V_2 \subset U_2$ such that

$$\forall x^* \in T(V_2), \qquad \langle x^*, e_1\rangle > \rho_0(e_1)(1-\varepsilon_0)s_0. \tag{22}$$

For any U_3 chosen by A with $U_3 \subset V_2$, define

$$D_1 = \left\{ e \in X \;\middle|\; \sup_{x^*\in T(U_3)} \langle x^*, e\rangle \ge \rho_0(e)(1-\varepsilon_0)s_0 \right\}. \tag{23}$$

Then (22) implies $e_1 \in D_1$. In addition, because of the boundedness of $T(U_3)$, $x \mapsto \sup_{x^*\in T(U_3)}\langle x^*, x\rangle$ is continuous, hence, D_1 is closed.

Denote

$$\rho_1(x) = \rho_0(x) + \mu_1(x - e_1), \tag{24}$$

where μ_1 is defined as in Proposition 4. Set

$$s_1 = \sup_{x^*\in T(U_3)} \rho^*_1(x^*). \tag{25}$$

Then from $U_3 \subset U_2$ and (10), $s_1 \le s_0$ and from (22), (24) and (25), we deduce that

$$\forall x^* \in T(U_3), \qquad (1-\varepsilon_0)s_0 < \frac{\langle x^*, e_1\rangle}{\rho_0(e_1)} = \frac{\langle x^*, e_1\rangle}{\rho_1(e_1)} \le s_1 \le s_0. \tag{26}$$

Take ε_1 so small that

$$\varepsilon_1 \in (0, (1-\varepsilon_0)^2/2^2) \quad \text{and} \quad (1-\varepsilon_0)s_0 < (1-\varepsilon_1)s_1. \tag{27}$$

By the same reasoning, B can choose $V_3 \subset U_3$ and e_2 such that

$$\forall x^* \in T(V_3), \qquad \langle x^*, e_2\rangle > \rho_1(e_2)(1-\varepsilon_1)s_1. \tag{28}$$

For any U_4 chosen by A with $U_4 \subset V_3$, define

$$D_2 = \Big\{ e \in X \;\Big|\; \sup_{x^* \in T(U_4)} \langle x^*, e\rangle \geq \rho_1(e)(1-\varepsilon_1)s_1 \Big\}. \tag{29}$$

Then $e_2 \in D_2$, $D_2 \subset D_1$ and D_2 is closed.

Now for $\rho_{n-1}, s_{n-1}, \varepsilon_{n-1}, e_n, V_{n+1}$, U_{n+2} and D_n given as above, $n = 1, 2, \ldots$, set

$$\rho_n(x) = \rho_{n-1}(x) + \mu_n(x - e_n), \tag{30}$$
$$s_n = \sup_{x^* \in T(U_{n+2})} \rho_n^*(x^*). \tag{31}$$

Then

$$\forall x^* \in T(U_{n+2}), \qquad (1-\varepsilon_{n-1})s_{n-1} < \frac{\langle x^*, e_n\rangle}{\rho_{n-1}(e_n)} = \frac{\langle x^*, e_n\rangle}{\rho_n(e_n)} \leq s_n \leq s_{n-1}. \tag{32}$$

Take ε_n so small that

$$\varepsilon_n \in (0, (1-\varepsilon_0)^2/2^{n+1}) \quad \text{and} \quad (1-\varepsilon_{n-1})s_{n-1} < (1-\varepsilon_n)s_n. \tag{33}$$

B can choose e_{n+1} and $V_{n+2} \subset U_{n+2}$ such that

$$\forall x^* \in T(V_{n+2}), \qquad \langle x^*, e_{n+1}\rangle > \rho_n(e_{n+1})(1-\varepsilon_n)s_n. \tag{34}$$

For any U_{n+3} chosen by A with $U_{n+3} \subset V_{n+2}$, set

$$D_{n+1} = \Big\{ e \in X \;\Big|\; \sup_{x^* \in T(U_{n+3})} \langle x^*, e\rangle \geq \rho_n(e)(1-\varepsilon_n)s_n \Big\}. \tag{35}$$

Then $e_{n+1} \in D_{n+1}$, $D_{n+1} \subset D_n$ and D_{n+1} is closed.

As the game continues, B obtains sequences $\{\rho_{n-1}\}_{n=1}^\infty$, $\{s_{n-1}\}_{n=1}^\infty$, $\{\varepsilon_{n-1}\}_{n=1}^\infty$, $\{e_n\}_{n=1}^\infty$, $\{U_n\}_{n=1}^\infty$, $\{V_n\}_{n=1}^\infty$ and $\{D_n\}_{n=1}^\infty$.

For any $x_n \in D_{n+1}$, from (35), there exists $x_n^* \in T(U_{n+3})$ satisfying

$$\frac{\langle x_n^*, x_n\rangle}{\rho_{n-1}(x_n) + \mu_n(x_n - e_n)} \geq (1-\varepsilon_n)s_n > (1-\varepsilon_{n-1})s_{n-1}$$

or

$$\rho_{n-1}(x_n) + \mu_n(x_n - e_n) < \frac{\langle x_n^*, x_n\rangle}{(1-\varepsilon_{n-1})s_{n-1}}. \tag{36}$$

On the other hand, since $x_n^* \in T(U_{n+3}) \subset T(U_{n+1}) \subset T(U_2)$, by (31), we have that

$$\frac{\langle x_n^*, x_n\rangle}{\rho_{n-1}(x_n)} \le s_{n-1} \qquad \text{or} \qquad \rho_{n-1}(x_n) \ge \frac{\langle x_n^*, x_n\rangle}{s_{n-1}}. \tag{37}$$

Obviously, $\|x_n\| < 1$; otherwise $\rho_{n-1}(x_n) = +\infty$, which contradicts (36). So, from (19) and (20), it follows

$$\langle x_n^*, x_n\rangle \le \|x_n^*\|^* \le \rho_0^*(x_n^*)/(1-\varepsilon_0) \le s_0/(1-\varepsilon_0). \tag{38}$$

Therefore, from (36), (37), (33), (27) and (38), it follows that

$$\mu_n(x_n - e_n) \le \frac{\langle x_n^*, x_n\rangle}{(1-\varepsilon_{n-1})s_{n-1}} - \frac{\langle x_n^*, x_n\rangle}{s_{n-1}} \le \frac{\varepsilon_{n-1}\langle x_n^*, x_n\rangle}{(1-\varepsilon_{n-1})s_{n-1}} \le \frac{\varepsilon_{n-1}}{(1-\varepsilon_0)^2} < \frac{1}{2^n}.$$

Hence, by (8) i), $\|x_n - e_n\| < 1/n$. Thus, $\{D_n\}_{n=1}^\infty$ is a nested sequence of closed subsets with diameters decreasing to 0 and hence there exists a unique $e_\infty \in X$ such that

$$\{e_\infty\} = \bigcap_{n=1}^{\infty} D_n. \tag{39}$$

Now we show that T is β-continuous at any $y_\infty \in \bigcap_{n=1}^\infty V_n$. In fact, for any $y_\infty^* \in T(y_\infty) \subset T(V_{n+2})$, from (34), we have that

$$\langle y_\infty^*, e_{n+1}\rangle \ge \rho_n(e_{n+1})(1-\varepsilon_n)s_n. \tag{40}$$

By (32) and (33), s_n converges to some $s_\infty > 0$. By using Proposition 4, $\rho_\infty(x) = \rho_0(x) + \sum_{k=1}^\infty \mu_k(x - e_k)$ is a β-well function on X and $\rho_n \to \rho_\infty$ uniformly on $\{x \in X \mid \|x\| < 1\}$. Joining up with $e_n \to e_\infty$ and $\varepsilon_n \to 0$, from (40), we deduce that

$$\langle y_\infty^*, e_\infty\rangle \ge \rho_\infty(e_\infty)s_\infty. \tag{41}$$

On the other hand, for any n,

$$\frac{\langle y_\infty^*, e_\infty\rangle}{\rho_\infty(e_\infty)} \le \rho_\infty^*(y_\infty^*) \le \rho_n^*(y_\infty^*) \le s_n.$$

then we also have

$$\frac{\langle y_\infty^*, e_\infty\rangle}{\rho_\infty(e_\infty)} \le \rho_\infty^*(y_\infty^*) \le s_\infty. \tag{42}$$

Thus, (41) and (42) ensure that

$$\rho_\infty^*(y_\infty^*) = \frac{\langle y_\infty^*, e_\infty\rangle}{\rho_\infty(e_\infty)} = s_\infty. \tag{43}$$

From Proposition 6 i), $y_\infty^* = s_\infty \nabla_\beta \rho_\infty(e_\infty)$, and it follows that $T(y_\infty)$ is a singleton.

Moreover, for any given $\delta > 0$, we have that for n sufficiently large,

$$\forall y^* \in T(V_{n+2}) \subset T(U_{n+2}), \qquad \rho^*_\infty(y^*) \le \rho^*_n(y^*) \le s_n < s_\infty + \delta. \tag{44}$$

At the same time, from (34), we have

$$\forall y^* \in T(V_{n+2}), \quad \frac{\langle y^*, e_{n+1}\rangle}{\rho_n(e_{n+1})} > (1-\varepsilon_n)s_n \tag{45}$$

and noting (by (31)) that $\|e_\infty\| < 1$ and $\rho_n \ge 1$,

$$\begin{aligned}
&\left|\frac{\langle y^*, e_\infty\rangle}{\rho_\infty(e_\infty)} - \frac{\langle y^*, e_{n+1}\rangle}{\rho_n(e_{n+1})}\right| \\
&\le \left|\langle y^*, e_\infty\rangle \left(\frac{1}{\rho_\infty(e_\infty)} - \frac{1}{\rho_n(e_{n+1})}\right)\right| + \left|\frac{1}{\rho_n(e_{n+1})}\langle y^*, e_\infty - e_{n+1}\rangle\right| \\
&\le \|y^*\|^* \left(\left|\frac{1}{\rho_\infty(e_\infty)} - \frac{1}{\rho_n(e_{n+1})}\right| + \|e_\infty - e_{n+1}\|\right) \\
&\le \frac{s_0}{1-\varepsilon_0}\left(\left|\frac{1}{\rho_\infty(e_\infty)} - \frac{1}{\rho_n(e_{n+1})}\right| + \|e_\infty - e_{n+1}\|\right).
\end{aligned}$$

When n tends to $+\infty$, the last two terms above tend to 0. Hence, joining up with (45), we have also that for n sufficiently large,

$$\frac{\langle y^*, e_\infty\rangle}{\rho_\infty(e_\infty)} \ge s_\infty - \delta. \tag{46}$$

Due to Proposition 6 ii), from (44) and (46), we deduce that for any $S \in \beta$ and $\varepsilon > 0$, there exists a sufficiently large n such that

$$T(V_{n+2}) \subset T(y_\infty) + D_{S,\varepsilon},$$

i.e., T is β-continuous at y_∞. Thus, $\bigcap_{n=1}^\infty V_n \subset W$. □

Corollary. *If a Banach space X admits a Lipschitz β-differentiable bump function, i.e., (H) holds, then every continuous convex function on a convex open subset $\Omega \subset X$ is β-differentiable in a dense G_δ subset of Ω.*

Remark 1. Recall that a map $\phi\colon \Omega \to 2^{X^*}$, where Ω may be any Hausdorff space, is said to be a **w^*-usco map** if ϕ is nonempty compact convex valued and upper semicontinuous for the w^*-topology of X^*. A w^*-usco map $\phi\colon \Omega \to 2^{X^*}$ is said to be **minimal** if there is no w^*-usco map $\psi\colon \Omega \to 2^{X^*}$ such that the graph of ψ is a proper subset of the graph of ϕ ([15]). A maximal monotone mapping on X is an example of minimal w^*-usco map if we take $\Omega := \operatorname{int}\{\, x \in X \mid T(x) \ne \emptyset \,\}$. As in [15], we can use the same method to generalize Theorem 1 as follows.

Theorem 2. *Assume that X admits a Lipschitz β-differentiable bump function, i.e., (H) holds. If Ω is a Baire space and $\phi\colon \Omega \to 2^{X^*}$ is a minimal w^*-usco map, then ϕ is β-continuous in a dense G_δ subset of Ω.*

Remark 2. In the proof of Theorem 1, our aim is to show (43); that is, to show that there exists e_∞ such that for any $y^*_\infty \in T(e_\infty)$ and any $e \in X$,

$$\frac{1}{\langle y^*_\infty, e\rangle}\left(\rho_0(e) + \sum_{n=1}^{\infty}\mu_n(e - e_n)\right) \geq \frac{1}{\langle y^*_\infty, e_\infty\rangle}\left(\rho_0(e_\infty) + \sum_{n=1}^{\infty}\mu_n(e_\infty - e_n)\right)$$

It is similar to the generalization of Ekeland's ε-variational principle in our paper [11], and its proof is also as in [11]. So, combining a Banach–Mazur game, it is possible to prove a more general ε-variational principle. We will discuss this problem in another paper.

Remark 3. The only difference between (H') and the existence of a β-smooth norm is the convexity. In fact, we have the following proposition.

Proposition 9. *Let X be a Banach space. If the following assumption holds:*

$$(H'_c)\quad \begin{cases} \textit{There exists a (Lipschitz) } \beta\textit{-superdifferentiable function} \\ \mu: X \to [0,1] \ \textit{ such that } \mu(0) = 0, \ \Big[\|x\| \geq 1\Big] \Longrightarrow \Big[\mu(x) = 1\Big] \\ \textit{and for some } \alpha \in (0, 1/2),\ \mu \textit{ is convex on } \{x \in X \mid \mu(x) \leq 2\alpha\}. \end{cases}$$

Then there exists a β-smooth equivalent norm on X.

Proof. Without loss of generality, we can assume that μ is symmetric. Otherwise, consider $(\mu(x) + \mu(-x))/2$ instead of μ.

Define the Minkowski function $p: X \to \mathbb{R}$ associated to the open convex subset $\Omega := \{ x \in X \mid \mu(x) < \alpha \}$, i.e.,

$$p(x) = \inf\{\lambda > 0 \mid x \in \lambda\Omega\}.$$

Since Ω is nonempty, symmetric and bounded, p is an equivalent norm on X. We prove that p is β-smooth away from the origin.

Let $G := \partial\Omega$ be the boundary of Ω. It is sufficient to show that p is β-smooth at each point of G, or, due to Proposition 1, that ∂p is β-continuous on G.

Let $x_0 \in G$. Firstly, we notice that μ is convex and continuous on $\overline{\Omega}$. So, μ is subdifferentiable in the sense of convex analysis at x_0, and hence is β-differentiable at x_0. We show that

$$\langle \nabla_\beta \mu(x_0), x_0\rangle > 0. \tag{47}$$

Otherwise, $\mu(0) - \mu(x_0) \geq \langle \nabla_\beta\mu(x_0), -x_0\rangle \geq 0$, which contradicts the fact that $\mu(0) = 0 < \alpha = \mu(x_0)$.

Secondly, we show that $\partial p(x_0)$ is a singleton. In fact, it is well-known that for the indicator function $\delta_{\overline{\Omega}}: X \to \mathbb{R} \cup \{+\infty\}$, defined by

$$\delta_{\overline{\Omega}}(x) := \begin{cases} 0, & \text{if } x \in \overline{\Omega}, \\ +\infty, & \text{otherwise}, \end{cases}$$

we have that

$$\partial\delta_{\overline{\Omega}}(x_0) = \bigcup_{\lambda \geq 0} \lambda\partial p(x_0) = \bigcup_{\lambda\geq 0}\lambda\nabla_\beta\mu(x_0),$$

(see, for instance, [3], pp. 55–56.) and it follows that

$$\partial p(x_0) \subset \bigcup_{\lambda \geq 0} \lambda \nabla_\beta \mu(x_0). \tag{48}$$

On the other hand, since p is a norm on X, we have that

$$\forall x^* \in \partial p(x_0), \qquad \langle x^*, x_0 \rangle = p(x_0) = 1. \tag{49}$$

Therefore, from (47), (48) and (49), we deduce that

$$\partial p(x_0) = \left\{ \frac{\nabla_\beta \mu(x_0)}{\langle \nabla_\beta \mu(x_0), x_0 \rangle} \right\}.$$

i.e., the Gâteaux derivative of p at x_0 is

$$\nabla p(x_0) = \frac{\nabla_\beta \mu(x_0)}{\langle \nabla_\beta \mu(x_0), x_0 \rangle}. \tag{50}$$

Finally, we show that ∇p is β-continuous on G. In fact, from (50), it is easy to see that in general,

$$\forall x \neq 0, \qquad \nabla p(x) = p(x) \frac{\nabla_\beta \mu(x/p(x))}{\langle \nabla_\beta \mu(x/p(x)), x/p(x) \rangle}.$$

So, for any $x_0 \in G$ and any x, $h \in X$, we have that

$$\begin{aligned} &\langle \nabla p(x) - \nabla p(x_0), h \rangle \\ =& \left\langle \frac{p(x)\nabla_\beta \mu(x/p(x))}{\langle \nabla_\beta \mu(x/p(x)), x/p(x) \rangle} - \frac{\nabla_\beta \mu(x_0)}{\langle \nabla_\beta \mu(x_0), x_0 \rangle}, h \right\rangle \\ \leq& \left| \frac{p(x) - 1}{\langle \nabla_\beta \mu(x/p(x)), x/p(x) \rangle} \langle \nabla_\beta \mu(x/p(x)), h \rangle \right| \\ +& \left| \frac{1}{\langle \nabla_\beta \mu(x/p(x)), x/p(x) \rangle} - \frac{1}{\langle \nabla_\beta \mu(x_0), x_0 \rangle} \right| |\langle \nabla_\beta \mu(x/p(x)), h \rangle| \\ +& \frac{1}{\langle \nabla_\beta \mu(x_0), x_0 \rangle} \langle \nabla_\beta \mu(x/p(x)) - \nabla_\beta \mu(x_0), h \rangle. \end{aligned} \tag{51}$$

Take any $S \in \beta$. Then, $\sup_{h \in S} \|h\| < +\infty$. Since $\nabla_\beta \mu$ is locally bounded and $\|\cdot\|$-w^*-continuous over $\{x \,|\, \mu(x) < 2\alpha\}$ and $\langle \nabla_\beta \mu(x_0), x_0 \rangle > 0$, it is easy to deduce that the function $x \mapsto 1/\langle \nabla_\beta \mu(x/p(x)), x/p(x) \rangle$ is continuous at x_0. Hence, joining up with the continuity of p, when $x \to x_0$, the first and second added terms of (51) can be converges to 0 uniformly on $h \in S$. Moreover, due to the β-continuity of $\nabla_\beta \mu$ over $\{x \,|\, \mu(x) < 2\alpha\}$, the third one also converges to 0.

Thus, ∇p is β-continuous on G. □

Given a C^1 bump function, we can apply Leduc's method to construct a C^1 positive homogeneous function which is different from an equivalent norm only in the subadditivity

([14], Theorem 3.6). This together with Proposition 9 tells us that the essential obstruction to the transition from a smooth bump function to a smooth norm is exactly the lack of convexity.

This is the reason why, in our proof of Theorem 1, we have avoided the use of the subadditivity and the subdifferentiability of norm; in particular, we did not use Cauchy sequences, but used sequences of nested closed subsets to show the convergence of $\{e_n\}$.

Acknowledgment. The authors are grateful to two anonymous referees for their useful comments and suggestions. In particular, they show us that our main result for the case of the Gâteaux-differentiability has been announced by [5]. The first author wishes to thank Professor Phelps for his kind letter with helpful comments and a copy of [5]. The second author wishes to thank Professor Godefroy for sending the new book [7] to China.

References

[1] E. Asplund: Fréchet differentiability of convex functions, Acta Math., Vol. 121, 1968, 31–47

[2] J.M. Borwein, D. Preiss: A smooth variational principle with applications to subdifferentiability and to differentiability of convex functions Trans. Amer. Math. Soc., Vol. 303, 1987, 517–527

[3] F.H. Clarke: Optimization and Nonsmooth Analysis, Wiley–Interscience, New York, 1983

[4] M. M. Day: Normed Linear Spaces, 3rd ed. Springer–Verlag, Berlin–Heidelberg–New York, 1973

[5] R. Deville, G. Godefroy, V. Zizler: Un principe variationnel utilisant des fonctions bosses C. R. Acad. Sci. Paris Vol. 312, série I, 1991, 281–286

[6] R. Deville, G. Godefroy, V. Zizler: A smooth variational principle with applications to Hamilton–Jacobi equations in infinite dimensions J. of Func. Anal. Vol. 111, 1993, 197–212

[7] R. Deville, G. Godefroy, V. Zizler: Smoothness and Renormings in Banach Spaces, Longman Sc. & Tec., Harlow–New York, 1993

[8] I. Ekeland, G. Lebourg: Generic Fréchet differentiability and perturbed optimization problems in Banach spaces Trans. Am. Math. Soc. Vol. 224, 1976, 193–216

[9] R. Haydon: A counterexample to several questions about scattered compact spaces Bull. London Math. Soc. Vol. 22 , 1990, 261–268

[10] R. Haydon: Trees and renorming theory, to appear

[11] J.C. Oxtoby: The Banach–Mazur game and Banach category theorem, in: Contributions to the Theory of Games, vol.3, Annals of Math. Studies 39, Princeton, 1957, 159–163

[12] Y. Li, S. Shi: A generalization of Ekeland's ε-variational principle and of its Borwein–Preiss' smooth variant, to appear

[13] R.R. Phelps: Support cones in Banach space and their applications Advances in Math. Vol. 13, 1974, 1–19

[14] R.R Phelps: Convex functions, monotone operators and differentiability, Lecture Notes in Math., No. 1364, Springer–Verlag , 1989

[15] D. Preiss, R.R. Phelps, I. Namioka: Smooth Banach spaces and monotone or usco mappings, Israel J. of Math. Vol. 72, 1990, 257–279