Journal of Convex Analysis
Volume 11 (2004), No. 1, 1–16

On Total Convexity, Bregman Projections and Stability in Banach Spaces

Elena Resmerita*
Department of Mathematics, University of Haifa,
31905 Haifa, Israel

Received May 14, 2002
Revised manuscript received March 13, 2003

Totally convex functions and Bregman projections associated to them are of special interest for building optimization and feasibility algorithms. This motivates one to investigate existence of totally convex functions in Banach spaces. Also, this raises the question whether and under which conditions the corresponding Bregman projections have the properties needed for guaranteeing convergence and stability of the algorithms based on them. We show that a reflexive Banach space in which some power $r \in (1, +\infty)$ of the norm is totally convex is an E-space and conversely. Also we prove that totally convex functions in reflexive Banach spaces are necessarily essentially strictly convex in the sense of [6]. We use these facts in order to establish continuity and stability properties of Bregman projections.

Keywords: Bregman distance, Bregman projection, total convexity, essential strict convexity, E-space, Mosco convergence

1. Introduction

Let X be a Banach space and $g : X \to (-\infty, +\infty]$ be a proper convex function. According to [14, Section 1.2], the *modulus of total convexity of g at the point $x \in dom\, g$* is the function $\nu_g(x, \cdot) : [0, +\infty) \to [0, +\infty]$ defined by

$$\nu_g(x, t) = \inf \{D_g(y, x) : y \in dom\, g, \|y - x\| = t\}, \tag{1}$$

where $D_g : dom\, g \times dom\, g \to [0, +\infty]$ is the *Bregman distance with respect to g* given by

$$D_g(y, x) := g(y) - g(x) - g^\circ(x, y - x).$$

For any $x \in dom\, g$ and $z \in X$ we denote by $g^\circ(x, z)$ the *right-hand sided derivative of g at x in the direction z*, that is,

$$g^\circ(x, z) = \lim_{t \searrow 0} \frac{g(x + tz) - g(x)}{t}.$$

The function g is called *totally convex at $x \in dom\, g$* if $\nu_g(x, t) > 0$, for all $t \in (0, \infty)$. The function g is called *totally convex on the convex set $C \subseteq dom\, g$* if it is totally convex at any point $x \in C$. It was shown in [14, Section 1.2.4] that (locally) uniformly convex functions are totally convex functions, while the latter are strictly convex (cf. [14, Proposition 1.2.6]). Total convexity is a property of interest in the convergence analysis of the proximal

*The author gratefully acknowledges the support of the Israel Science Foundation founded by the Israel Academy of Sciences and Humanities.

ISSN 0944-6532 / $ 2.50

point algorithm with generalized distances (see, e.g., [11], [12], [16], [27], [30], [31], [32], [38]) and of some projection type methods for solving convex feasibility problems (see, for instance, [1], [2], [5], [8], [10], [14], [21], [23], [36]). This concept was introduced in [13] and further studied in detail in [15], [17]. A natural question is how this notion relates with other concepts studied for similar purposes, as for instance, essential strict convexity (as defined in [6], by generalizing the one proposed in [37]) and Kadec-Klee property (see, e.g., [25]). Using results from [19] makes it possible to prove that totally convex functions in reflexive Banach spaces are essentially strictly convex (see Proposition 2.1). In practice, the most convenient totally convex functions to work with in infinite dimensional Banach spaces are powers greater than one of the norm (provided that they have this property). For example, this can be seen in the analysis of the augmented Lagrangian methods ([14], [30]), in algorithms for solving the stochastic convex feasibility problem ([14], [21]) or linear operator equations ([18]). Therefore, it is interesting to know in which Banach spaces the powers greater than one of the norm are totally convex. It has been shown that the function $g_r = \|\cdot\|^r$ with $r > 1$ is totally convex in $\mathcal{L}^p$ and ℓ^p with $p > 1$ (see [14]) or, more general, in any uniformly convex space ([18]). This holds even in any locally uniformly convex space where the function g_r is uniformly convex at any point (cf. [40]), and, hence, totally convex (cf. [14], [19]). We prove below that a Banach space on which the functions $g_r = \|\cdot\|^r$ with $r > 1$ are totally convex is necessarily strictly convex and has the Kadec-Klee property. Moreover, we give a characterization of the reflexive Banach spaces in which, for some $r > 1$, g_r is totally convex. Precisely, we prove that those spaces are exactly the E-spaces (see Theorem 3.3). Another problem which occurs in the same algorithmic context is whether or not the Bregman projections with respect to totally convex functions in reflexive Banach spaces have continuity and stability properties which will ensure enough stability of the computational procedures when errors are involved. In Section 4, we show that Bregman projections with respect to Fréchet differentiable totally convex functions in reflexive Banach spaces are continuous and have good stability properties with respect to Mosco convergence. This is relevant because it opens a way for deeper studying the effect of computational errors on the behavior of various iterative algorithms involving Bregman projections like, for instance, those in [30].

2. Totally convex functions

We start by discussing the relationship between essential strict convexity and total convexity. Recall that, according to [6], the function g is *essentially strictly convex* if $(\partial g)^{-1}$ is locally bounded on its domain and g is strictly convex on every convex subset of $dom(\partial g)$.

Proposition 2.1. *Let X be a reflexive Banach space and $g : X \to (-\infty, +\infty]$ be a proper convex lower semicontinuous function such that $dom(\partial g)$ is a convex set. If g is totally convex on $dom(\partial g)$, then it is essentially strictly convex.*

Proof. According to [6, Theorem 5.4], the function g is essentially strictly convex if and only if its Fenchel conjugate g^* is essentially smooth. Showing essential smoothness of g^* is equivalent, by [6, Theorem 5.6], to proving that $int\,dom\,g^* \neq \emptyset$ and ∂g^* is single valued on its domain. In order to establish these properties, take $x \in dom(\partial g)$ (such an x exists since $dom\,g \subset \overline{dom(\partial g)}$ - see, e.g., [4, Corollary 2.2, p. 110]). By [19, Proposition 3.4], total convexity of g at x implies $x^* \in int\,dom\,g^*$, whenever $x^* \in \partial g(x)$. If y^* belongs to $dom(\partial g^*)$, which is necessarily nonempty because it contains $int\, dom\, g^*$, then there exists $y \in X$ such that $y^* \in \partial g(y)$. Applying again [19, Proposition 3.4], it follows that $y^* \in int\, dom\, g^*$ and g^* is Fréchet differentiable at y^*. Hence, $\partial g^*(y^*)$ is singleton. $\square$

The converse implication in Proposition 2.1 is not true. In order to see that, one can take into account Theorem 3.3 in Section 3 which states that a reflexive and strictly convex Banach space has the Kadec-Klee property if and only if the function $\frac{1}{2}\|\cdot\|^2$ is totally convex. In this respect, the reference [7] provides an example of a reflexive, strictly convex Banach space which fails to have the Kadec-Klee property. Therefore, the function $\frac{1}{2}\|\cdot\|^2$ on that space cannot be totally convex, although it is essentially strictly convex.

In finite dimensional Banach spaces, essential strict convexity seems to be slightly different from total convexity. In this respect, it is interesting to note that the most useful essentially strictly convex functions in $\mathbb{R}^n$ (see, for instance, [5]) are also totally convex. Moreover, any function which has closed domain and which is strictly convex and continuous on its domain as well as any strictly convex function whose domain is the entire space is essentially strictly convex and totally convex at the same time (see [14, Proposition 1.2.6]). On one hand, we do not have any example of a function which simultaneously satisfies the assumptions of Proposition 2.1 and is essentially strictly convex without being totally convex. On the other hand, we do not have a proof for the equivalence of the two notions in finite dimensional Banach spaces.

The following result, which gives a sequential characterization of the total convexity of a function, will be repeatedly used later:

Proposition 2.2. *Let X be a Banach space, let $g : X \to (-\infty, +\infty]$ be a convex function and take $x \in dom\, g$. The following statements are equivalent:*

(*i*) *The function g is totally convex at x;*

(*ii*) *For any sequence $\{y_n\}_{n\in\mathbb{N}} \subseteq dom\, g$,*

$$\lim_{n\to\infty} D_g(y_n, x) = 0 \Rightarrow \lim_{n\to\infty} \|y_n - x\| = 0; \tag{2}$$

(*iii*) *For any sequence $\{y_n\}_{n\in\mathbb{N}} \subseteq dom\, g$,*

$$\liminf_{n\to\infty} D_g(y_n, x) = 0 \Rightarrow \liminf_{n\to\infty} \|y_n - x\| = 0;$$

(*iv*) *For any sequence $\{y_n\}_{n\in\mathbb{N}} \subseteq dom\, g$,*

$$\lim_{n\to\infty} D_g(y_n, x) = 0 \Rightarrow \liminf_{n\to\infty} \|y_n - x\| = 0.$$

Proof. $(i) \Leftrightarrow (ii)$: Suppose that g is totally convex at x. Take $\{y_n\}_{n\in\mathbb{N}} \subseteq dom\, g$ such that $\lim_{n\to\infty} D_g(y_n, x) = 0$. Since $\nu_g(x, \|y_n - x\|) \le D_g(y_n, x)$ for all $n \in \mathbb{N}$, it follows that

$$\lim_{n\to\infty} \nu_g(x, \|y_n - x\|) = 0. \tag{3}$$

Suppose, by contradiction, that there exist a positive number ε and a subsequence $\{y_{n_k}\}_{k\in\mathbb{N}}$ of $\{y_n\}_{n\in\mathbb{N}}$ such that $\|y_{n_k} - x\| \ge \varepsilon$, for all $k \in \mathbb{N}$. It was shown in [14, Proposition 1.2.2] that the function $\nu_g(x, \cdot)$ is strictly increasing, whenever x is in the algebraic interior of *dom* g. It is easy to see that this result is still valid when $x \in \mathit{dom}\, g$. Consequently, we get

$$\lim_{k\to\infty} \nu_g(x, \|y_{n_k} - x\|) \geq \nu_g(x, \varepsilon) > \nu_g(x, 0) = 0,$$

contradicting (3). Conversely, suppose that there exists $t_0 > 0$ such that $\nu_g(x, t_0) = 0$, that is, there exists $\{y_n\}_{n\in\mathbb{N}} \subseteq \mathit{dom}\, g$ with $\|y_n - x\| = t_0$ and $\lim_{n\to\infty} D_g(y_n, x) = 0$. Then (2) yields $t_0 = 0$, a contradiction. $(i) \Leftrightarrow (iii)$ can be shown in the same way as $(i) \Leftrightarrow (ii)$, using the strict monotonicity of $\nu_g(x, \cdot)$. Obviously, $(iii) \Rightarrow (iv)$. The proof for the converse implication is the same as for $(ii) \Rightarrow (i)$, the contradiction being reached as follows: $0 = \liminf_{n\to\infty} \|y_n - x\| = t_0 > 0$. □

In the sequel, we emphasize some special properties of Bregman distances. This result improves in some respects Theorem 3.5(v) in [5] and Lemma 7.3(vii) in [6].

Proposition 2.3. *Let X be a reflexive Banach space and let $g : X \to (-\infty, +\infty]$ be a convex, lower semicontinuous function which is totally convex at some $x \in \mathit{dom}\, g$. Then,*

(i) *for all $y \in \mathit{dom}\, g$,*

$$x^* \in \partial g(x), y^* \in \partial g(y) \Rightarrow D_g(y, x) + D_{g^*}(y^*, x^*) \leq \langle y^* - x^*, y - x\rangle;$$

If, in addition, g is Gâteax differentiable at x, then the inequality becomes equality.

(ii) *For all $y \in \mathit{dom}\, g$,*

$$x^* \in \partial g(x), y^* \in \partial g(y) \Rightarrow D_g(x, y) \leq D_{g^*}(y^*, x^*);$$

If, in addition, g is Gâteax differentiable at y, then the inequality becomes equality.

(iii) *If g satisfies*

$$\lim_{\|z\|\to\infty} g(z)/\|z\| = +\infty,$$

then, for any $x^*, y^* \in X^*$ there exist $x_0 \in (\partial g)^{-1}(x^*)$ and $y_0 \in (\partial g)^{-1}(y^*)$, such that

$$D_{g^*}(y^*, x^*) \geq D_g(x_0, y_0).$$

If, in addition, g is Gâteax differentiable on dom(∂g), then the inequality becomes equality.

Proof. (i) Note that $x^*, y^* \in \mathit{dom}\, \partial g^* \subseteq \mathit{dom}\, g^*$ and, thus, it makes sense to discuss about $D_{g^*}(y^*, x^*)$. Recall that whenever $z^* \in \partial g(z)$, we have

$$g(z) + g^*(z^*) = \langle z^*, z\rangle.$$

According to Proposition 3.4 in [19], if $x^* \in \partial g(x)$, then $x^* \in \mathit{int}\, \mathit{dom}(g^*))$, the function g^* is Fréchet differentiable at x^* and $(g^*)'(x^*) = x$. Consequently, if $y^* \in \partial g(y)$, then

$$\begin{aligned} D_g(y, x) + D_{g^*}(y^*, x^*) &= g(y) - g(x) - g^\circ(x, y - x) \\ &+ g^*(y^*) - g^*(x^*) - \langle y^* - x^*, (g^*)'(x^*)\rangle \\ &\leq \langle y^*, y\rangle - \langle x^*, x\rangle - \langle x^*, y - x\rangle - \langle y^* - x^*, x\rangle \\ &= \langle y^* - x^*, y - x\rangle. \end{aligned}$$

If g is differentiable at x, then $\partial g(x)$ has a unique element x^* and, hence, $g^\circ(x, y-x)$ equals to $\langle x^*, y-x\rangle$.

(ii) Suppose that $x^* \in \partial g(x)$ and $y^* \in \partial g(y)$ with $y \in dom\, g$. Then

$$\begin{aligned} D_{g^*}(y^*, x^*) &= g^*(y^*) - g^*(x^*) - \langle y^* - x^*, (g^*)'(x^*)\rangle \\ &= g^*(y^*) - g^*(x^*) - \langle y^* - x^*, x\rangle \\ &= \langle y^*, y\rangle - g(y) - \langle x^*, x\rangle + g(x) - \langle y^* - x^*, x\rangle \\ &= g(x) - g(y) - \langle y^*, x-y\rangle \\ &\geq g(x) - g(y) - g^\circ(y, x-y). \end{aligned}$$

(iii) Take $x^*, y^* \in X^*$. By [4, Proposition 2.5, p.112], we deduce that there exist $x_0, y_0 \in X$ such that $x^* \in \partial g(x_0)$ and $y^* \in \partial g(y_0)$. Then we apply (ii). □

3. Locally totally convex spaces

Let X be a Banach space. Recall that the function $\varphi : \mathbb{R}_+ \to \mathbb{R}_+$ is *a weight function* if it is increasing, continuous, $\varphi(0) = 0$ and $\lim_{t\to+\infty} \varphi(t) = +\infty$. We denote by $J_\varphi : X \to \mathcal{P}(X^*)$ and $J : X \to \mathcal{P}(X^*)$ the *duality mapping of weight* φ and the *normalized duality mapping*, respectively (see, e.g., [24]). We follow [39] and associate to the weight function φ the functions $\psi : \mathbb{R}_+ \to \mathbb{R}_+$ and $\Phi : X \to \mathbb{R}_+$ given by

$$\psi(t) = \int_0^t \varphi(s)ds \text{ and } \Phi(x) = \psi(\|x\|). \tag{4}$$

The function ψ is strictly convex, increasing, differentiable, $\psi' = \varphi$, it has

$$\lim_{t\to+\infty} \frac{\psi(t)}{t} = +\infty, \tag{5}$$

and Φ is convex and continuous (cf. [40], Lemma 3.7.1 and Theorem 3.7.2). We use h to denote the function $\frac{1}{2}\|\cdot\|^2$ all over this section. It is known that the space X is uniformly convex if and only if h is uniformly convex on bounded sets. Also, X is locally uniformly convex if and only if h is uniformly convex at any point (see, e.g., [40]). We call a Banach space *locally totally convex* if the function $h = \frac{1}{2}\|\cdot\|^2$ is totally convex at each $x \in X$. Locally uniformly convex spaces are locally totally convex (cf. [14, Section 1.2.4]); separable or reflexive Banach spaces can be equivalently renormed for becoming locally totally convex spaces (see [9] and [24], respectively). In this section, we establish some properties of locally totally convex spaces and show that there are strong connections between these spaces and the E-spaces. We start with a useful characterization of the locally totally convex spaces.

Theorem 3.1. *The following conditions are equivalent:*

(i) *X is locally totally convex;*

(ii) *There exists a weight function φ such that the function Φ is totally convex;*

(iii) *For any weight function φ, the function Φ is totally convex.*

Proof. $(i) \Rightarrow (iii)$: Let φ be a weight function. Consider $x \in X$ and take $\{y_n\}_{n\in\mathbb{N}} \subseteq X$ such that

$$\lim_{n\to\infty} D_\Phi(y_n, x) = 0. \tag{6}$$

We show that $\lim_{n\to\infty} \|y_n - x\| = 0$. If $x = 0$, then

$$0 = \lim_{n\to\infty} D_\Phi(y_n, x) = \lim_{n\to\infty} \psi(\|y_n\|).$$

Taking into account that the function ψ is increasing, we necessarily obtain that $\lim_{n\to\infty} \|y_n\| = 0$. Now, suppose that $x \neq 0$. Since Φ is continuous at x and convex, for any $n \in \mathbb{N}$, there exists $x_n^* \in \partial\Phi(x) = J_\varphi(x)$ such that

$$\Phi^\circ(x, y_n - x) = \max\{\langle z^*, y_n - x\rangle : z^* \in \partial\Phi(x)\} = \langle x_n^*, y_n - x\rangle$$

First, note that $\{y_n\}_{n\in\mathbb{N}}$ is bounded. Indeed, if we assume that $\{y_n\}_{n\in\mathbb{N}}$ is unbounded, i.e., $\lim_{n\to\infty} \|y_n\| = +\infty$ (or a subsequence of it), then we have

$$\begin{aligned}
D_\Phi(y_n, x) &= \psi(\|y_n\|) - \psi(\|x\|) - \langle x_n^*, y_n - x\rangle \\
&= \psi(\|y_n\|) - \psi(\|x\|) + \|x\|\varphi(\|x\|) - \langle x_n^*, y_n\rangle \\
&\geq \psi(\|y_n\|) - \psi(\|x\|) + \|x\|\varphi(\|x\|) - \|y_n\|\varphi(\|x\|) \\
&= \|y_n\| \left[\psi(\|y_n\|)/\|y_n\| - \psi(\|x\|)/\|y_n\| + \|x\|\varphi(\|x\|)/\|y_n\| - \varphi(\|x\|)\right].
\end{aligned}$$

Letting $n \to \infty$ and taking into account (5), we get a contradiction. Since $\{y_n\}_{n\in\mathbb{N}}$ is bounded, there exists a subsequence $\{\|y_{n_k}\|\}_{k\in\mathbb{N}}$ of $\{\|y_n\|\}_{n\in\mathbb{N}}$ which converges to some $\alpha \in \mathbb{R}_+$. For any $k \in \mathbb{N}$,

$$\delta_k = \frac{1}{2}\psi(\|y_{n_k}\|) + \frac{1}{2}\psi(\|x\|) - \psi\left(\frac{\|y_{n_k} + x\|}{2}\right).$$

From the convexity of Φ we obtain

$$\Phi^\circ(x, y_{n_k} - x) \leq 2\left[\Phi\left(\frac{y_{n_k} + x}{2}\right) - \Phi(x)\right]$$

and, thus, we deduce that

$$\frac{1}{2} D_\Phi(y_{n_k}, x) \geq \delta_k \geq 0,$$

for all $k \in \mathbb{N}$. Consequently, $\lim_{k\to\infty} \delta_k = 0$ and, hence,

$$\lim_{k\to\infty}\left[\frac{1}{2}\psi(\|y_{n_k}\|) + \frac{1}{2}\psi(\|x\|) - \psi\left(\frac{\|y_{n_k}\| + \|x\|}{2}\right)\right] \leq \lim_{k\to\infty} \delta_k = 0.$$

This implies

$$\frac{1}{2}\psi(\alpha) + \frac{1}{2}\psi(\|x\|) \leq \psi\left(\frac{\alpha + \|x\|}{2}\right), \tag{7}$$

because ψ is continuous. Since ψ is also strictly convex, (7) cannot hold unless $\alpha = \|x\|$. This, combined with (6) yields

$$0 = \lim_{k\to\infty} \Phi^\circ(x, y_{n_k} - x) = \lim_{k\to\infty} \varphi(\|x\|) g^\circ(x, y_{n_k} - x),$$

i.e., $\lim_{k\to\infty} g^\circ(x, y_{n_k} - x) = 0$, where $g = \|\cdot\|$. Thus,

$$\lim_{k\to\infty} D_h(y_{n_k}, x) = \lim_{k\to\infty}\left[\frac{1}{2}\|y_{n_k}\|^2 - \frac{1}{2}\|x\|^2 - \|x\|g^\circ(x, y_{n_k} - x)\right] = 0.$$

Consequently, by Proposition 2.2, we obtain

$$\lim_{k\to\infty} \|y_{n_k} - x\| = 0.$$

Observing that the entire sequence $\{y_n\}_{n\in\mathbb{N}}$ converges to x completes the proof. $(ii) \Rightarrow (i)$: Suppose that for some weight function φ, Φ is totally convex. Let $\{y_n\}_{n\in\mathbb{N}}$ be such that $\lim_{n\to\infty} D_h(y_n, x) = 0$. We would like to establish that $\lim_{n\to\infty} \|y_n - x\|$. If $x = 0$, then

$$0 = \lim_{n\to\infty} D_h(y_n, x) = \lim_{n\to\infty} (1/2)\|y_n\|^2.$$

Now, suppose that $x \neq 0$. Note that, for any $n \in \mathbb{N}$, there exists $x_n^* \in \partial h(x) = J(x)$ such that

$$\begin{aligned} D_h(y_n, x) &= \frac{1}{2}\|y_n\|^2 - \frac{1}{2}\|x\|^2 - \langle x_n^*, y_n - x\rangle \\ &\geq \frac{1}{2}\|y_n\|^2 + \frac{1}{2}\|x\|^2 - \|x\|\|y_n\| \\ &= \frac{1}{2}\left(\|y_n\| - \|x\|\right)^2, \end{aligned}$$

and, consequently, $\lim_{n\to\infty} \|y_n\| = \|x\|$. We have

$$0 = \lim_{n\to\infty} D_h(y_n, x) = \lim_{n\to\infty} (-\|x\|g^\circ(x, y_n - x)),$$

that is, $\lim_{n\to\infty} g^\circ(x, y_n - x) = 0$. By a simple calculation, we get $\lim_{n\to\infty} D_\Phi(y_n, x) = 0$. Therefore, Proposition 2.2 and the hypothesis imply $\lim_{n\to\infty} \|y_n - x\| = 0$. □

The next lemma provides a sufficient condition for the locally total convexity of a Banach space.

Lemma 3.2. *If $h = \frac{1}{2}\|\cdot\|^2$ is totally convex at each point of the unit sphere of X, then X is locally totally convex.*

Proof. Let $x \in X$ and denote $z := \frac{1}{\|x\|}x$. Take $\{y_n\}_{n\in\mathbb{N}} \subseteq X$ such that $\lim_{n\to\infty} D_h(y_n, x) = 0$. It is sufficient to show that $\lim_{n\to\infty} \left\|\frac{1}{\|x\|}y_n - z\right\| = 0$. Note that

$$\begin{aligned} D_h\left(\frac{1}{\|x\|}y_n, z\right) &= \frac{\|y_n\|^2}{2\|x\|^2} - \frac{1}{2} - \frac{1}{\|x\|}\max\{\langle z^*, y_n - x\rangle : z^* \in \partial h(z)\} \\ &= \frac{\|y_n\|^2}{2\|x\|^2} - \frac{1}{2} - \frac{1}{\|x\|^2}\max\left\{\langle \|x\|z^*, y_n - x\rangle : \|x\|z^* \in \partial h(x)\right\} \\ &= \frac{1}{\|x\|^2}D_h(y_n, x). \end{aligned}$$

Since h is totally convex at z with $\|z\| = 1$ and $\lim_{n\to\infty} D_h\left(\frac{1}{\|x\|}y_n, z\right) = 0$, we get $\lim_{n\to\infty}\left\|\frac{1}{\|x\|}y_n - z\right\| = 0$. □

The E-spaces, introduced by Fan and Glicksberg [28], were studied over the years because they are the natural setting for the research of strong Tykhonov and Hadamard well-posedness of convex best approximation problems (see, e.g., [26], [29]), and variational inequalities in Banach spaces (see, e.g., [3]). Recall that the Banach space X is an *E-space* if it is reflexive, strictly convex and has the Kadec-Klee property, that is,

$$\left(w - \lim_{n\to\infty} x_n = x \text{ and } \lim_{n\to\infty} \|x_n\| = \|x\|\right) \Rightarrow \lim_{n\to\infty} x_n = x.$$

Note that the notion of E-space is equivalent to that of weakly uniformly convex space as defined in [24, p. 47]. A Banach space X is called *weakly uniformly convex* if for any two sequences $\{x_n\}_{n\in\mathbb{N}}$, $\{y_n\}_{n\in\mathbb{N}} \subseteq X$ with $\|x_n\| = \|y_n\| = 1, n \in \mathbb{N}$, we have

$$\lim_{n\to\infty} \langle x^*, x_n + y_n\rangle = 2 \quad \Rightarrow \quad \lim_{n\to\infty} \|x_n - y_n\| = 0,$$

whenever $x^* \in X^*, \|x^*\| = 1$. We show next that the E-spaces are exactly those Banach spaces which are reflexive and locally totally convex.

Theorem 3.3. *The Banach space X is an E-space if and only if it is reflexive and locally totally convex.*

Proof. "$\Rightarrow$": Let $x \in X$. In order to show that h is totally convex at x, it is sufficient to prove that

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ D_h(y, x) < \delta \Rightarrow \|y - x\| < \varepsilon. \tag{8}$$

Since h is convex and continuous on X, there exists $x^* \in \partial h(x)$ such that $h^\circ(x, y - x) = \langle x^*, y - x\rangle$. Note that (cf. [26, Theorem 3, p. 41]), X is an E-space if and only if the dual norm $\|\cdot\|_*$ is Fréchet differentiable on $X^* \setminus \{0\}$ and, consequently, the function $h^* = \frac{1}{2}\|\cdot\|_*^2$ is Fréchet differentiable on X^*. According to [26, Proposition 26 and Theorem 27, p. 15], the Fréchet differentiability of h^* at the point x^* is equivalent to the strong rotundity of h at $x = (h^*)'(x^*)$ with slope x^*, i.e.,

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ h(y) - h(x) - \langle x^*, y - x\rangle < \delta \Rightarrow \|y - x\| < \varepsilon.$$

This is, in fact, (8). "$\Leftarrow$": X is strictly convex because, (cf. [14, Proposition 1.2.6]), any totally convex function is strictly convex. Thus, it remains to show that X has the Kadec-Klee property. Indeed, let $\{x_n\}_{n\in\mathbb{N}}$ such that w–$\lim_{n\to\infty} x_n = x$ and $\lim_{n\to\infty} \|x_n\| = \|x\|$. Since $h^\circ(x, \cdot)$ is convex and continuous on X, it is also weakly lower semicontinuous. Hence, we have

$$\begin{aligned}
\limsup_{n\to\infty} D_h(x_n, x) &= \limsup_{n\to\infty} \left[\frac{1}{2}\|x_n\|^2 - \frac{1}{2}\|x\|^2 - h^\circ(x, x_n - x)\right] \\
&= -\liminf_{n\to\infty} h^\circ(x, x_n - x) \\
&\leq -h^\circ(x, 0) = 0.
\end{aligned}$$

Consequently, $\lim_{n\to\infty} D_h(x_n, x) = 0$ and, by the total convexity of h at x, we obtain $\lim_{n\to\infty} x_n = x$. □

Now, one can summarize the connections between locally totally convex spaces and various classes of Banach spaces as follows:

Corollary 3.4. *The following implications hold among the statements below:* $(i)\Rightarrow(ii)\Rightarrow(iii)\Rightarrow(iv)$*:*

(*i*) *X is uniformly convex;*
(*ii*) *X is an E-space/locally uniformly convex space;*
(*iii*) *X is locally totally convex;*
(*iv*) *X is strictly convex.*

Note that locally totally convex spaces with additional properties are locally uniformly convex, as one can see below.

Proposition 3.5. *If X is locally totally convex and if $\|\cdot\|^2$ is Fréchet differentiable on X, then X is locally uniformly convex.*

Proof. This follows from [19, Proposition 2.3] and [40, Theorem 4.1 (*i*)]. □

4. Continuity and stability properties of the Bregman projections

The concept of Bregman projection was first used by Bregman [10], while the terminology is due to Censor and Lent [22]. It has been shown that this generalized projection is a good replacement for the metric projection in optimization methods and in algorithms for solving convex feasibility problems. Let $K \subseteq int\,dom\,g$ be a nonempty closed convex set and take $x \in int\,dom\,g$. Recall that *the Bregman projection of x onto K with respect to g* is defined by

$$\Pi_K^g(x) = argmin\{D_g(y,x) : y \in K\}.$$

If X is reflexive, g is totally convex and Gâteaux differentiable on $int\,dom\,g$, and lower semicontinuous, then there exists a unique minimizer of the function $D_g(\cdot,x)$ in K (see Proposition 2.1.5 in [14] for a proof). In other words, $\Pi_K^g(x)$ is the only point in which the Bregman distance from x to K, i.e.,

$$D_g(K,x) = \inf_{y\in K}\{D_g(y,x)\},$$

is attained. Moreover, the Bregman projection $\Pi_K^g(x)$ is characterized by the following inequality

$$D_g(z,\Pi_K^g(x)) + D_g(\Pi_K^g(x),x) \le D_g(z,x), \tag{9}$$

for every $z \in K$. Note that, if X is a Hilbert space and $g = \|\cdot\|^2$, then $\Pi_K^g(x)$ is the metric projection of x onto K. All over this section, unless otherwise stated, we assume that X is a reflexive Banach space. We prove next some results which are useful in the sequel, but may also be of more general interest.

Proposition 4.1. *Let X be a Banach space, $g : X \to (-\infty,+\infty]$ be a convex function which is Gâteaux differentiable on an open convex set $U \subseteq dom\,g$. Then, the following statements are true:*

(*i*) *If g is continuous on U, then the function $D_g(\cdot,\cdot)$ is continuous on $U \times U$;*
(*ii*) *If g is Fréchet differentiable on U and lower semicontinuous, then the function $D_g(\cdot,\cdot)$ is sequentially weak-to-norm lower semicontinuous on $U \times U$;*

(*iii*) *If $\{y_n\}_{n\in\mathbb{N}}$ is a sequence contained in U, $x, y \in U$ and g is totally convex at y, then*

$$\left(w\text{-}\lim_{n\to\infty} y_n = y \text{ and } \lim_{n\to\infty} D_g(y_n, x) = D_g(y, x)\right) \Rightarrow \lim_{n\to\infty} y_n = y.$$

Proof. (*i*) Consider $x, y \in U$ and let two sequences $\{x_n\}_{n\in\mathbb{N}}$, $\{y_n\}_{n\in\mathbb{N}} \subseteq U$ such that $\lim_{n\to\infty} x_n = x$ and $\lim_{n\to\infty} y_n = y$. According to [34, Proposition 2.1], the function $\langle g'(\cdot), \cdot\rangle$ is continuous on $U \times X$. Then, letting $n \to \infty$ in

$$D_g(y_n, x_n) = g(y_n) - g(x_n) - \langle g'(x_n), y_n - x_n\rangle,$$

we obtain $\lim_{n\to\infty} D_g(y_n, x_n) = D_g(y, x)$. (*ii*) Take $x, y, \{x_n\}_{n\in\mathbb{N}}$ as in (*i*) and let $\{y_n\}_{n\in\mathbb{N}} \subseteq U$ with $w - \lim_{n\to\infty} y_n = y$. Using the weakly lower semicontinuity of g and the continuity of the differential g', we obtain

$$\begin{aligned}
\liminf_{n\to\infty} D_g(y_n, x_n) &= \liminf_{n\to\infty} [g(y_n) - g(x_n) - \langle g'(x_n) - g'(x), y_n - x_n\rangle] \\
&- \lim_{n\to\infty} \langle g'(x), y_n - x_n\rangle \\
&= \liminf_{n\to\infty} g(y_n) - g(x) - \langle g'(x), y - x\rangle \\
&\geq D_g(y, x).
\end{aligned}$$

(*iii*) Note that

$$D_g(y_n, x) - D_g(y, x) - D_g(y_n, y) = \langle g'(x) - g'(y), y - y_n\rangle.$$

Letting $n \to \infty$ and using the hypothesis, we obtain $\lim_{n\to\infty} D_g(y_n, y) = 0$. Since g is totally convex at y, we get $\lim_{n\to\infty} y_n = y$. □

Continuity of the Bregman projection operator $\Pi_K^g : int\, dom\, g \to K$ and of the function $D_g(K, \cdot)$, when happens, is an important property in applications (see, for instance, [20] and Theorem 2.3.6 in [14]). As far as we know, this fact is already established for the special case when $X = \mathbb{R}^n$ and g is the negentropy (see [13]). Also, the continuity of $D_g(K, \cdot)$ when g is a totally convex function in reflexive Banach spaces was shown in [14, Section 2.3.6] under quite restrictive conditions. We can prove now that this property as well as the continuity of the operator Π_K^g hold in more general situations. First, let us denote, for any $\alpha \in (0, +\infty)$ and $y \in dom\, g$,

$$R_\alpha^g(y; K) := \{x \in K : D_g(y, x) \leq \alpha\}. \tag{10}$$

Proposition 4.2. *Let $g : X \to (-\infty, +\infty]$ be a convex lower semicontinuous function which is totally convex and Fréchet differentiable on $int\, dom\, g$. If the set $R_\alpha^g(y; K)$ is bounded, whenever $\alpha \in (0, +\infty)$ and $y \in K$, then the function $D_g(K, \cdot)$ is continuous on $int\, dom\, g$.*

Proof. Since $D_g(K, \cdot)$ is upper semicontinuous on $int\, dom\, g$ (as a consequence of Proposition 4.1(*i*)), it remains to prove only lower semicontinuity. To this end, consider $z \in int\, dom\, g$ and let $\{z_n\}_{n\in\mathbb{N}}$ be a sequence in $int\, dom\, g$ such that $\lim_{n\to\infty} z_n = z$. Denote $\hat{z} := \Pi_K^g(z)$ and $\hat{z}_n := \Pi_K^g(z_n)$, for all $n \in \mathbb{N}$. Then,

$$\begin{aligned}
D_g(\hat{z}_n, z_n) - D_g(\hat{z}, z) &\geq D_g(\hat{z}_n, z_n) - D_g(\hat{z}_n, z) \\
&= [g(z) - g(z_n)] + \langle g'(z_n) - g'(z), z_n - \hat{z}_n\rangle \\
&+ \langle g'(z), z_n - z\rangle.
\end{aligned}$$

If we prove that the sequence $\{\hat{z}_n\}_{n\in\mathbb{N}}$ is bounded, then using the continuity of the function g and of the differential g', we obtain

$$\liminf_{n\to\infty} D_g(\hat{z}_n, z_n) \geq D_g(\hat{z}, z).$$

For $y \in K$, inequality (9) applied to $\hat{z}_n = \Pi_K^g(z_n)$ yields, for any $n \in \mathbb{N}$,

$$D_g(y, \hat{z}_n) + D_g(\hat{z}_n, z_n) \leq D_g(y, z_n).$$

Clearly, the sequence $\{D_g(y, z_n)\}_{n\in\mathbb{N}}$ is convergent and, thus, bounded. Since the set $R_\alpha^g(y; K)$ is bounded for some upper bound α of the sequence $\{D_g(y, z_n)\}_{n\in\mathbb{N}}$, the boundedness of $\{\hat{z}_n\}_{n\in\mathbb{N}}$ immediately follows. □

Using the preceding result, we show the continuity of Π_K^g : *int dom* $g \to K$ under the same requirements.

Proposition 4.3. *Let* $g : X \to (-\infty, +\infty]$ *be a convex lower semicontinuous function which is totally convex and Fréchet differentiable on* int dom g. *If the set* $R_\alpha^g(y; K)$ *is bounded, whenever* $\alpha \in (0, +\infty)$ *and* $y \in K$, *then the Bregman projection operator* Π_K^g : *int dom* $g \to K$ *is norm-to-norm continuous on int dom* g.

Proof. Consider $\hat{z}$ and $\{\hat{z}_n\}_{n\in\mathbb{N}}$ as above. From (9), we have, for any $n \in \mathbb{N}$,

$$D_g(\hat{z}, \hat{z}_n) + D_g(\hat{z}_n, z_n) \leq D_g(\hat{z}, z_n).$$

Taking into account Proposition 4.2 and the continuity of the function $D_g(\hat{z}, \cdot)$, we deduce that $\lim_{n\to\infty} D_g(\hat{z}, \hat{z}_n) = 0$. Now Proposition 2.3($ii$) applies and yields

$$\lim_{n\to\infty} D_{g^*}(g'(\hat{z}_n), g'(\hat{z})) = \lim_{n\to\infty} D_g(\hat{z}, \hat{z}_n) = 0.$$

Our further argument is based on the following fact.

Claim. The function g^* is totally convex at $g'(\hat{z}) \in$ *int dom* g^*.

For proving the claim, it is sufficient to show, cf. [19, Proposition 3.5], that $g'(\hat{z}) \in$ *int dom* g^*, the conjugate function g^* is continuous at $g'(\hat{z})$, $(g^*)'(g'(\hat{z})) \in$ *int dom*(g^{**}) and g^{**} is Fréchet differentiable at $(g^*)'(g'(\hat{z}))$. Indeed, by [19, Proposition 3.4], it follows that $g'(\hat{z}) \in$ *int dom* g^* and g^* is Fréchet differentiable at $g'(\hat{z})$. Hence, g^* is also continuous at $g'(\hat{z})$. The fact that $g = g^{**}$ is Fréchet differentiable on *int dom* g and, thus, at $(g^*)'(g'(\hat{z})) = \hat{z}$ completes the proof of the claim. Consequently, by Proposition 2.2, we get $\lim_{n\to\infty} \|g'(\hat{z}) - g'(\hat{z}_n)\|_* = 0$. As shown in the previous proposition, the sequence $\{\hat{z}_n\}_{n\in\mathbb{N}}$ is bounded. Therefore, taking limit in

$$D_g(\hat{z}, \hat{z}_n) + D_g(\hat{z}_n, \hat{z}) = \langle g'(\hat{z}_n) - g'(\hat{z}), \hat{z} - \hat{z}_n \rangle$$

implies $\lim_{n\to\infty} D_g(\hat{z}_n, \hat{z}) = 0$. Finally, from total convexity of g we obtain that $\lim_{n\to\infty} \hat{z}_n = \hat{z}$. □

In the significant particular case of $g = \|\cdot\|^r$, $r > 1$, Proposition 4.2 and Proposition 4.3 become the following:

Corollary 4.4. *If* X *and* X^* *are E-spaces and* $g = \|\cdot\|^r$, $r > 1$, *then* (*i*) *The function* $D_g(K, \cdot)$ *is continuous on* X*;*

(*ii*) *The Bregman projection operator* Π_K^g *is norm-to-norm continuous on* X*.*

Proof. According to Theorems 3.1 and 3.3, the functions g and g^* are totally convex. In order to apply Propositions 4.2 and 4.3 and to get the conclusion, it is sufficient to take into account [18, Corollary 2.4(i)] and to prove that g is Fréchet differentiable on X. Indeed, by [19, Proposition 3.4], it follows that $g^{**} = g$ is Fréchet differentiable on the image of $\partial(g^*)$. It is known (see, e.g., [39, Theorem 3.7.2]) that $g^* = c\|\cdot\|_*^q$ where $1/r + 1/q = 1$ and c is some positive constant. Applying Asplund's Theorem, we deduce that $\partial(g^*)$ is a duality mapping of some weight function. Moreover, this mapping acting from X^* into X^{**} is surjective because X is reflexive (see [24, Theorem 3.4, p. 62]). By consequence, the function g is Fréchet differentiable on the entire space X. □

As observed in the Introduction, Bregman distances and Bregman projections are useful in optimization and feasibility algorithms. Therefore, knowing how they vary when the set that one projects onto is perturbed is of interest for investigating stability of such methods. We study below the behavior of the Bregman distances and Bregman projections when X is reflexive and the set K is "approximated" in Mosco's sense [33], generalizing in some respects several results from [26, p. 49]. A sequence $\{K_n\}_{n\in\mathbb{N}}$ of subsets of X is said *to converge in the sense of Mosco to* $K \subseteq X$, and we write $M - \lim_{n\to\infty} K_n = K$, if the following two conditions hold:

($M1$) *for every* $u \in K$ *there exists* $u_n \in K_n$ *for every* n *sufficiently large such that* $\lim_{n\to\infty} u_n = u$; ($M2$) *for every subsequence* n_j *of natural numbers, if* $x_j \in K_{n_j}$ *for every* j *and* $w - \lim_{j\to\infty} x_j = x$, *then* $x \in K$.

The next result shows that if $\{K_n\}_{n\in\mathbb{N}}$ converges to K in Mosco's sense, then $\{\Pi_{K_n}^g\}_{n\in\mathbb{N}}$ also approaches Π_K^g. A careful analysis of the Theorem 4.5 below shows that the implications (i)⇒ (ii)⇒(iii) remain true even if g is not Fréchet, but only Gâteaux differentiable.

Theorem 4.5. *Let* $g : X \to (-\infty, +\infty]$ *be a convex lower semicontinuous function which is totally convex and Fréchet differentiable on* $\operatorname{int}\operatorname{dom} g$, *and* $\{K_n\}_{n\in\mathbb{N}}$, K *be closed convex subsets of* $\operatorname{int}\operatorname{dom} g$. *Among the statements below the implications* (i)⇔(ii)⇒(iii) *hold*:

(*i*) *The sequence* $\{K_n\}_{n\in\mathbb{N}}$ *converges in the sense of Mosco to* K;

(*ii*)

$$\lim_{n\to\infty} \Pi_{K_n}^g(x) = \Pi_K^g(x), \text{ for every } x \in \operatorname{int}\operatorname{dom} g;$$

(*iii*)

$$\lim_{n\to\infty} D_g(K_n, x) = D_g(K, x), \text{ for every } x \in \operatorname{int}\operatorname{dom} g.$$

Proof. (i)⇒ (iii): Fix $x \in \operatorname{int}\operatorname{dom} g$ and denote $x_0 := \Pi_K^g(x)$ and $x_n := \Pi_{K_n}^g(x)$. Let $u \in K$ and $u_n \in K_n$ such that $\lim_{n\to\infty} u_n = u$. Then, for any $n \in \mathbb{N}$,

$$D_g(u_n, x_n) + D_g(x_n, x) \le D_g(u_n, x).$$

Since the sequence $\{D_g(u_n, x)\}_{n\in\mathbb{N}}$ is convergent (to $D_g(u, x)$), it is bounded and so is the sequence $\{D_g(x_n, x)\}_{n\in\mathbb{N}}$. Note that $D_g(x_n, x) \ge \nu_g(x, \|x_n - x\|)$ for all $n \in \mathbb{N}$; this, combined with the strict monotonicity of $\nu_g(x, \cdot)$, yields the boundedness of the sequence $\{x_n\}_{n\in\mathbb{N}}$. Hence, there exists some subsequence $\{x_{n_j}\}_{j\in\mathbb{N}}$ which converges weakly to some

$y \in X$. By $(M2)$, it follows that $y \in K$. Since g is lower semicontinuous and convex, it is weakly lower semicontinuous. By consequence,

$$D_g(y, x) \leq \liminf_{j \to \infty} D_g(x_{n_j}, x) \leq \lim_{j \to \infty} D_g(u_n, x) = D_g(u, x).$$

Since u was arbitrarily chosen in K, we deduce that $y = \Pi_K^g(x)$. As this weak cluster point is unique, we obtain that the entire sequence $\{x_n\}_{n \in \mathbb{N}}$ converges weakly to x_0. Hence,

$$D_g(x_0, x) \leq \liminf_{n \to \infty} D_g(x_n, x) \leq \limsup_{n \to \infty} D_g(x_n, x) \leq D_g(u, x).$$

for all $u \in K$. In particular, this holds for x_0. Therefore, the following limit exists and

$$\lim_{n \to \infty} D_g(x_n, x) = D_g(x_0, x),$$

which means that (iii) is proved. $(i) \Rightarrow$ (ii): Since (iii) holds, too, and $\{x_n\}_{n \in \mathbb{N}}$ converges weakly to x_0, we apply Proposition 4.1 (iii) and obtain $\lim_{n \to \infty} x_n = x_0$. $(ii) \Rightarrow$ (i): Clearly, $(M1)$ holds: if $x \in K$, there exists $\{\Pi_{K_n}^g(x)\}_{n \in \mathbb{N}}$ such that $\Pi_{K_n}^g(x) \in K_n$, for all $n \in \mathbb{N}$, and $\lim_{n \to \infty} \Pi_{K_n}^g(x) = \Pi_K^g(x) = x$. It remains to prove $(M2)$. Let $\{x_i\}_{i \in \mathbb{N}}$ with $x_i \in K_i$ such that it converges weakly to some $x \in X$. If $y_0 := \Pi_K^g(x)$ and $y_i := \Pi_{K_i}^g(x)$, then the hypothesis yields $\lim_{i \to \infty} y_i = y_0$. By (9) we have, for any $i \in \mathbb{N}$,

$$\langle g'(y_i) - g'(x), x_i - y_i \rangle \geq 0.$$

Letting $i \to \infty$, we obtain $\langle g'(y_0) - g'(x), x - y_0 \rangle \geq 0$. As g' is strictly monotone (because g is strictly convex), we deduce that $x = y_0 \in K$. $\square$

In the case of special interest when g is a power greater than one of the norm, Theorem 4.5 leads to the following result:

Corollary 4.6. *Suppose that X and X^* are E-spaces, $\{K_n\}_{n \in \mathbb{N}}$, K is a sequence of closed convex subsets of X and $g = \|\cdot\|^r$, $r > 1$. Among the statements below the implications $(i) \Leftrightarrow$ $(ii) \Rightarrow (iii)$ hold:*

(i) *The sequence $\{K_n\}_{n \in \mathbb{N}}$ converges in the sense of Mosco to K;*

(ii)

$$\lim_{n \to \infty} \Pi_{K_n}^g(x) = \Pi_K^g(x), \quad \text{for every } x \in X;$$

(iii)

$$\lim_{n \to \infty} D_g(K_n, x) = D_g(K, x), \quad \text{for every } x \in X.$$

Theorem 4.5 can be improved by using a technique occurring in the proof of [26, Theorem 10, p. 49]. We show next that the Bregman distances $D_g(K, x)$ and the Bregman projections $\Pi_K^g(x)$ with respect to certain totally convex functions are stable under simultaneous variation of the set K and of the point x.

Theorem 4.7. *Let $g : X \to (-\infty, +\infty]$ be a convex lower semicontinuous function which is totally convex and Fréchet differentiable on $\operatorname{int}\operatorname{dom} g$, and it satisfies the following condition: If $\{x_n\}_{n \in \mathbb{N}}$ and $\{y_n\}_{n \in \mathbb{N}}$ are sequences in $\operatorname{int}\operatorname{dom} g$ such that $\{x_n\}_{n \in \mathbb{N}}$ and $\{D_g(y_n, x_n)\}_{n \in \mathbb{N}}$ are bounded, then $\{y_n\}_{n \in \mathbb{N}}$ is bounded. Then the following statement is true:*

For every sequence x, $\{x_n\}_{n\in\mathbb{N}}$ in int dom g and for every sequence K, $\{K_n\}_{n\in\mathbb{N}}$ of closed convex subsets of int dom g such that $\lim_{n\to\infty} x_n = x$ and $M - \lim_{n\to\infty} K_n = K$, we have

$$\lim_{n\to\infty} D_g(K_n, x_n) = D_g(K, x) \quad \text{and} \quad w - \lim_{n\to\infty} \Pi^g_{K_n}(x_n) = \Pi^g_K(x). \tag{11}$$

Moreover, if the differential g' is bounded on bounded subsets of int dom g, then the convergence of Bregman projections in (11) is strong.

Proof. Denote $\hat{x}_n := \Pi^g_{K_n}(x_n)$ and $\hat{x} := \Pi^g_K(x)$. Let $z \in K$ and $z_n \in K_n$, $n \in \mathbb{N}$ such that $\lim_{n\to\infty} z_n = z$. Applying (9), it follows that, for any $n \in \mathbb{N}$,

$$D_g(z_n, \hat{x}_n) + D_g(\hat{x}_n, x_n) \le D_g(z_n, x_n).$$

Since $\{D_g(z_n, x_n)\}_{n\in\mathbb{N}}$ is convergent (as a consequence of Proposition 4.1(i)), it is also bounded. Then the sequence $\{D_g(\hat{x}_n, x_n)\}_{n\in\mathbb{N}}$ is bounded and, by hypothesis, so is the sequence $\{\hat{x}_n\}_{n\in\mathbb{N}}$. Consequently, $\{\hat{x}_n\}_{n\in\mathbb{N}}$ has some weakly convergent subsequence $\{\hat{x}_{n_j}\}_{j\in\mathbb{N}}$ to some point y which necessarily belongs to K. We apply Proposition 4.1(ii) and obtain

$$D_g(\hat{x}, x) \le D_g(y, x) \le \liminf_{j\to\infty} D_g(\hat{x}_{n_j}, x_{n_j}).$$

Now let $v_n \in K_n$, $n \in \mathbb{N}$ such that $\lim_{n\to\infty} v_n = \hat{x}$. From the previous inequality combined with the chain of inequalities below

$$\liminf_{j\to\infty} D_g(\hat{x}_{n_j}, x_{n_j}) \le \limsup_{j\to\infty} D_g(\hat{x}_{n_j}, x_{n_j}) \le \lim_{n\to\infty} D_g(v_n, x_n) = D_g(\hat{x}, x),$$

we deduce that $\lim_{n\to\infty} D_g(\hat{x}_n, x_n) = D_g(\hat{x}, x)$, the weak cluster point y coincides with $\hat{x}$ and, therefore, the entire sequence $\{\hat{x}_n\}_{n\in\mathbb{N}}$ converges weakly to $\hat{x}$. Thus, the first part of the theorem is proved. Suppose now that g' is bounded on bounded subsets of *int dom g*. Note that

$$D_g(\hat{x}_n, x) - D_g(\hat{x}_n, x_n) - D_g(x_n, x) = \langle g'(x) - g'(x_n), x_n - \hat{x}_n \rangle.$$

Taking limit above, we obtain that $\lim_{n\to\infty} D_g(\hat{x}_n, x) = D_g(\hat{x}, x)$. Now Proposition 4.1($iii$) applies and yields $\lim_{n\to\infty} \hat{x}_n = \hat{x}$. □

One can easily see that the function $g = \|\cdot\|^r$, $r > 1$ satisfies the hypothesis of the preceding theorem whenever X and X^* are E-spaces.

Acknowledgements. The author thanks Professor Dan Butnariu for many useful discussions which helped in the preparation of this paper, as well as the anonymous referee whose comments led to an improvement of an earlier version.

References

[1] Y. Alber: Metric and generalized projection operators in Banach spaces: Properties and applications, in: Theory and Applications of Nonlinear Operators of Accretive and Monotone Type, Dekker, New York (1996) 15–50.

[2] Y. Alber, D. Butnariu: Convergence of Bregman projection methods for solving consistent convex feasibility problems in reflexive Banach spaces, J. Optimization Theory Appl. 92(1) (1997) 33–61.

[3] Y. Alber, D. Butnariu, I. Ryazantseva: Regularization methods for ill-posed inclusions and variational inequalities with domain perturbations, J. Nonlinear Convex Anal. 1(2) (2001) 53–79.

[4] V. Barbu, Th. Precupanu: Convexity and Optimization in Banach Spaces, D. Reidel Publishing Co., Dordrecht, 2nd edition (1986).

[5] H. H. Bauschke, J. M. Borwein: Legendre functions and the method of random Bregman projections, J. Convex Anal. 4(1) (1997) 27–67.

[6] H. H. Bauschke, J. M. Borwein, P. L. Combettes: Essential smoothness, essential strict convexity, and Legendre functions in Banach spaces, Commun. Contemp. Math. 3(4) (2001) 615–647.

[7] H. H. Bauschke, P. L. Combettes: Construction of best Bregman approximations in reflexive Banach spaces, Proc. Am. Math. Soc. 31(12) (2003) 3757–3766.

[8] H. H. Bauschke, A. S. Lewis: Dykstra's algorithm with Bregman projections: A convergence proof, Optimization 48 (2000) 409–427.

[9] C. Bessaga, A. Pelczyński: Selected Topics in Infinite-Dimensional Topology, PWN-Polish Scientific Publishers, Warszawa (1975).

[10] L. M. Bregman: The relaxation method for finding common points of convex sets and its application to the solution of problems in convex programming, USSR Computational Mathematics and Mathematical Physics 7 (1967) 200–217.

[11] R. S. Burachik: Generalized Proximal Point Methods for the Variational Inequality Problem, PhD Thesis, Instituto de Mathematica Pura e Aplicada (IMPA), Rio de Janeiro, RJ, Brazil (1995).

[12] R. S. Burachik, S. Scheimberg: A proximal point method for the variational inequality problem in Banach spaces, SIAM J. Control Optim. 39(5) (2001) 1633–1649.

[13] D. Butnariu, Y. Censor, S. Reich: Iterative averaging of entropic projections for solving stochastic convex feasibility problems, Comput. Optim. Appl. 8 (1997) 21–39.

[14] D. Butnariu, A. N. Iusem: Totally Convex Functions for Fixed Points Computation and Infinite Dimensional Optimization, Kluwer Academic Publishers, Dordrecht (2000).

[15] D. Butnariu, A. N. Iusem: Local moduli of convexity and their application to finding almost common fixed points of measurable families of operators, in: Recent Developments in Optimization Theory and Nonlinear Analysis, Y. Censor, S. Reich (eds.), Contemp. Math. 204 (1997) 61–92.

[16] D. Butnariu, A. N. Iusem: On a proximal point method for convex optimization in Banach spaces, Numer. Funct. Anal. Optimization 18 (1998) 723–744.

[17] D. Butnariu, A. N. Iusem, R. S. Burachik: Iterative methods for solving stochastic convex feasibility problems and applications, Comput. Optim. Appl. 15 (2000) 269–307.

[18] D. Butnariu, A. N. Iusem, E. Resmerita: Total convexity for powers of the norm in uniformly convex Banach spaces, J. Convex Anal. 7 (2000) 319–334.

[19] D. Butnariu, A. N. Iusem, C. Zălinescu: On uniform convexity, total convexity and convergence of the proximal point and outer Bregman projection algorithms in Banach spaces, J. Convex Anal. 10(1) (2003) 35–61.

[20] D. Butnariu, S. Reich, A. Zaslavski: Weak convergence of orbits of nonlinear operators in reflexive Banach spaces, Numer. Funct. Anal. Optimization 24 (2003) 489–508.

[21] D. Butnariu, E. Resmerita: The outer Bregman projection method for stochastic feasibility problems in Banach spaces, in: Inherently Parallel Algorithms for Feasibility and Optimization and their Applications, D. Butnariu, Y. Censor, S. Reich (eds.), Elsevier, Amsterdam, Stud. Comput. Math. 8 (2001) 69–86.

[22] Y. Censor, A. Lent: An iterative row action method for interval convex programming, J. Optimization Theory Appl. 34 (1981) 321–353.

[23] Y. Censor, S. Zenios: Parallel Optimization: Theory, Algorithms and Applications, Oxford University Press, New York (1997).

[24] I. Cioărănescu: Geometry of Banach Spaces, Duality mappings, and Nonlinear Problems, Kluwer Academic Publishers, Dordrecht (1990).

[25] R. Deville, G. Godefroy, V. Zizler: Smoothness and Renormings in Banach Spaces, Longman Scientific & Technical, Harlow (1993).

[26] A. L. Dontchev, T. Zolezzi: Well-Posed Optimization Problems, Springer-Verlag, Berlin (1993).

[27] J. Eckstein: Nonlinear proximal point algorithms using Bregman functions, with applications to convex programming, Math. Oper. Res. 18(1) (1993) 202–226.

[28] K. Fan, I. Glicksberg: Some geometric properties of the spheres in a normed linear space, Duke Math. J. 25 (1958) 553–568.

[29] R. Holmes: A Course on Optimization and Best Approximation, Lecture Notes in Mathematics 257, Springer (1972).

[30] A. N. Iusem, R. Gárciga-Otero: Inexact versions of proximal point and augmented Lagrangian algorithms in Banach spaces, Numer. Funct. Anal. Optimization 22 (2001) 609–640.

[31] A. N. Iusem, R. Gárciga-Otero: Augmented Lagrangian methods for cone-constrained convex optimization in Banach spaces, J. Nonlinear Convex Analysis 3 (2002) 155–176.

[32] K. C. Kiwiel: Proximal minimization methods with generalized Bregman functions, SIAM J. Control Optimization 35(4) (1997) 1142–1168.

[33] U. Mosco: Approximation of the solutions of some variational inequalities, Ann. Sc. Norm. Sup. Pisa 21 (1967) 373–394.

[34] J.-P. Penot, R. Ratsimahalo: Characterizations of metric projections in Banach spaces and applications, Abstr. Appl. Anal. 3 (1998) 85–103.

[35] R. R. Phelps: Convex Functions, Monotone Operators and Differentiability, Springer Verlag, Berlin (1993).

[36] S. Reich: A weak convergence theorem for the alternating method with Bregman distances, in: Theory and Applications of Nonlinear Operators of Accretive and Monotone Type, Dekker, New York (1996) 313–318.

[37] R. T. Rockafellar: Convex Analysis, Princeton University Press, Princeton (1970).

[38] M. V. Solodov, B. F. Svaiter: An inexact hybrid generalized proximal point algorithm and some new results on the theory of Bregman functions, Math. Oper. Res. 25(2) (2000) 214–230.

[39] C. Zălinescu: Convex Analysis in General Vector Spaces, World Scientific, River Edge, N.J. (2002).

[40] C. Zălinescu: On uniformly convex functions, J. Math. Anal. Appl. 95 (1983) 344–374.