Journal of Convex Analysis
Volume 17 (2010), No. 3&4, 991–1006

A Strong Convergence Theorem for Resolvents of Monotone Operators

Dan Butnariu
*Department of Mathematics,
University of Haifa, 31905 Haifa, Israel
dbutnaru@math.haifa.ac.il*

Simeon Reich*
*Department of Mathematics, The Technion,
Israel Institute of Technology, 32000 Haifa, Israel
sreich@tx.technion.ac.il*

Shoham Sabach
*Department of Mathematics, The Technion,
Israel Institute of Technology, 32000 Haifa, Israel
ssabach@tx.technion.ac.il*

Dedicated to Hedy Attouch on the occasion of his 60th birthday.

Received: December 23, 2008
Revised manuscript received: December 04, 2009

We prove a strong convergence theorem for resolvents of monotone operators in Banach spaces. These resolvents are associated with totally convex Legendre functions.

Keywords: Banach space, Bregman projection, Legendre function, monotone operator, resolvent, totally convex function

2000 Mathematics Subject Classification: Primary 47H05; Secondary 52A41

1. Introduction

In this paper X denotes a real reflexive Banach space with norm $\|\cdot\|$ and X^* stands for the (topological) dual of X. Let $A : X \to 2^{X^*}$ be a monotone operator, that is, for any $x, y \in \operatorname{dom} A$, we have

$$\xi \in Ax \text{ and } \eta \in Ay \quad \Longrightarrow \quad \langle \xi - \eta, x - y \rangle \geq 0.$$

(Recall that the set $\operatorname{dom} A = \{x \in X \mid Ax \neq \varnothing\}$ is called the effective domain of such an operator A.) A monotone operator A is said to be maximal if the graph of A is not a proper subset of the graph of any other monotone operator. The operator A is said

*The second author was partially supported by the Israel Science Foundation (Grant 647/07), by the Fund for the Promotion of Research at the Technion, and by the Technion President's Research Fund.

ISSN 0944-6532 / $ 2.50

to be demiclosed at $x \in \operatorname{dom} A$ if for any sequence $\{(x_n, \xi_n)\}_{n\in\mathbb{N}}$ in $X \times X^*$,

$$\left.\begin{array}{c} x_n \rightharpoonup x \\ \xi_n \in Ax_n, \quad n \in \mathbb{N} \\ \xi_n \to \xi \end{array}\right\} \quad \Longrightarrow \quad \xi \in Ax.$$

Let $f : X \to (-\infty, +\infty]$ be a proper, lower semicontinuous and convex function, and let $f^* : X^* \to (-\infty, +\infty]$ be the Fenchel conjugate of f.

A problem of practical interest with which we are concerned in this paper is that of strong convergence of the resolvents $\operatorname{Res}^f_{\lambda A}(x)$ of monotone operators A relative to convex functions f. More precisely, given a monotone operator $A : X \to 2^{X^*}$ and a Legendre function $f : X \to (-\infty, +\infty]$, the question is whether and under which conditions $\lim_{\lambda\to+\infty} \operatorname{Res}^f_{\lambda A}(x)$ exists and is a zero of the operator A (see Section 2 for the definitions of all relevant concepts). The fact that this indeed happens in certain circumstances has been known for quite a long time. See [8] for the case of a maximal monotone operator A in Hilbert space ([24] and [26] for the case of m-accretive operators in Banach spaces) and, subsequently, [20] for the case of a maximal monotone operator A in smooth and uniformly convex Banach spaces. In these papers it is shown that if $f = \frac{1}{2}\|\cdot\|^2$, then $\operatorname{Res}^f_{\lambda A}(x)$ converges strongly to a point in $A^{-1}(0^*)$ as $\lambda \to +\infty$, provided such a point exists. More recently, in [21] it is claimed that for a maximal monotone operator A and a well-chosen function f $\left(\text{which is not necessarily } \frac{1}{2}\|\cdot\|^2\right)$, the strong limit $\lim_{\lambda\to+\infty} \operatorname{Res}^f_{\lambda A}(x)$ exists and is a point in $A^{-1}(0^*)$. We show below (see Theorem 3.1) that $\lim_{\lambda\to+\infty} \operatorname{Res}^f_{\lambda A}(x)$ exists and is exactly the Bregman projection of x relative to f onto $A^{-1}(0^*)$ for those monotone operators A which are not necessarily maximal monotone, but instead satisfy a certain range condition which is compatible with the Legendre function f. Theorem 5.1 extends this result to the case where A is approximated in some sense by more regular operators. Our results regarding strong convergence of resolvents of monotone operators show ways of strongly approximating zeroes of monotone operators in Banach spaces. Finding, even by approximation, zeroes of monotone operators is of interest in many fields. For instance, the minimization of lower semicontinuous convex functions reduces to finding zeroes of their subgradients which are monotone operators. More generally, as Kimura has already pointed out in [21], finding strong approximations of zeroes of monotone operators can be used in the process of solving variational inequalities. The literature contains several other methods for finding zeroes of monotone operators. See for example [1, 5, 6, 9, 11, 12, 13, 14, 18, 22, 30, 31] and the references therein. Many of them are fixed point methods which calculate fixed points of Res^f_A. Obviously, each fixed point of Res^f_A is, necessarily, a zero of A. An example of such a method is presented in [28]. Typically, the successful application of these fixed point methods is guaranteed under conditions on f and A which are not required when one approximates zeroes of A by $\operatorname{Res}^f_{\lambda A}(x)$ with large λ. Our paper is organized as follows. The next section is devoted to several preliminary definitions and results. Our main result (Theorem 3.1) is formulated and proved in Section 3. The fourth section contains three corollaries of our main result. In the fifth and last section we present an extension of Theorem 3.1 (Theorem 5.1) and two related propositions.

2. Preliminaries

2.1. Some facts about Legendre functions.

Legendre functions mapping a general Banach space X into $(-\infty, +\infty]$ are defined in [3]. According to [3, Theorems 5.4 and 5.6], since X is reflexive, the function f is Legendre if and only if it satisfies the following conditions:

(L1) The interior of the domain of f, $\operatorname{int}\operatorname{dom} f$, is nonempty, f is Gâteaux differentiable on $\operatorname{int}\operatorname{dom} f$ and

$$\operatorname{dom} \nabla f = \operatorname{int}\operatorname{dom} f; \tag{1}$$

(L2) The interior of the domain of f^*, $\operatorname{int}\operatorname{dom} f^*$, is nonempty, f^* is Gâteaux differentiable on $\operatorname{int}\operatorname{dom} f^*$ and

$$\operatorname{dom} \nabla f^* = \operatorname{int}\operatorname{dom} f^*. \tag{2}$$

Since X is reflexive, we always have $(\partial f)^{-1} = \partial f^*$ (see [7, p. 83]). This fact, when combined with conditions (L1) and (L2), implies the following equalities which we are going to use in the sequel:

$$\nabla f = (\nabla f^*)^{-1}, \tag{3}$$

$$\operatorname{ran} \nabla f = \operatorname{dom} \nabla f^* = \operatorname{int}\operatorname{dom} f^*, \tag{4}$$

$$\operatorname{ran} \nabla f^* = \operatorname{dom} \nabla f = \operatorname{int}\operatorname{dom} f. \tag{5}$$

Also, conditions (L1) and (L2) in conjunction with [3, Theorem 5.4] imply that the functions f and f^* are strictly convex on the interior of their respective domains.

Several interesting examples of Legendre functions are presented in [2] and [3]. Among them are the functions $\frac{1}{s}\|\cdot\|^s$ with $s \in (1, \infty)$, where the Banach space X is smooth and strictly convex and, in particular, a Hilbert space. From now on we assume that the convex function $f : X \to (-\infty, +\infty]$ is Legendre.

2.2. Some facts about totally convex functions.

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

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

The function $D_f : \operatorname{dom} f \times \operatorname{int}\operatorname{dom} f \to [0, +\infty)$, defined by

$$D_f(y, x) := f(y) - f(x) - f^\circ(x, y - x),$$

is called the *Bregman distance with respect to f* (cf. [16]). If f is a Gâteaux differentiable function, then the Bregman distance has the following important property, called the *three point identity*: for any $x \in \operatorname{dom} f$ and $y, z \in \operatorname{int}\operatorname{dom} f$,

$$D_f(x, y) + D_f(y, z) - D_f(x, z) = \langle \nabla f(z) - \nabla f(y), x - y \rangle. \tag{6}$$

Recall that, according to [10, Section 1.2, p. 17], the function f is called *totally convex at a point* $x \in \operatorname{int}\operatorname{dom} f$ if its *modulus of total convexity at* x, that is, the function $\upsilon_f : \operatorname{int}\operatorname{dom} f \times [0, +\infty) \to [0, +\infty]$ defined by

$$\upsilon_f(x, t) := \inf \{D_f(y, x) \mid y \in \operatorname{dom} f,\ \|y - x\| = t\}, \tag{7}$$

is positive whenever $t > 0$. The function f is called *totally convex* when it is totally convex at every point $x \in \operatorname{int}\operatorname{dom} f$. The following proposition summarizes some properties of the modulus of total convexity.

Proposition 2.1 (cf. [10, Propostion 1.2.2, p. 18]). *Let f be a proper, convex and lower semicontinuous function. If $x \in \operatorname{int}\operatorname{dom} f$, then*

(*i*) *The domain of $\upsilon_f(x, \cdot)$ is an interval of the form $[0, \tau_f(x))$ or $[0, \tau_f(x)]$ with $\tau_f(x) \in (0, +\infty]$.*
(*ii*) *If $c \in [1, +\infty)$ and $t \geq 0$, then $\upsilon_f(x, ct) \geq c\upsilon_f(x, t)$.*
(*iii*) *The function $\upsilon_f(x, \cdot)$ is superadditive, that is, for any $s, t \in [0, +\infty)$, we have $\upsilon_f(x, s + t) \geq \upsilon_f(x, s) + \upsilon_f(x, t)$.*
(*iv*) *The function $\upsilon_f(x, \cdot)$ is increasing; it is strictly increasing if and only if f is totally convex at x.*

Another proposition which is very useful in the proof of our main result is the following one.

Proposition 2.2 (cf. [29, Proposition 2.2, p. 3]). *Let $f : X \to (-\infty, +\infty]$ be a convex function and take $x \in \operatorname{int}\operatorname{dom} f$. Then f is totally convex at x if and only if* $\lim_{n\to\infty} D_f(y_n, x) = 0$ *implies that* $\lim_{n\to\infty} \|y_n - x\| = 0$ *for every sequence* $\{y_n\}_{n\in\mathbb{N}} \subset \operatorname{dom} f$.

2.3. The Resolvent of A relative to f.

Let $A : X \to 2^{X^*}$ be an operator such that

$$(\operatorname{int}\operatorname{dom} f) \bigcap (\operatorname{dom} A) \neq \varnothing. \tag{8}$$

The operator

$$\operatorname{Prt}_A^f := (\nabla f + A)^{-1} : X^* \to 2^X$$

is called the *protoresolvent* of A, or, more precisely, the *protoresolvent of A relative to* f. This allows us to define the *resolvent* of A, or, more precisely, the *resolvent of A relative to* f, introduced and studied in [4], as the operator $\operatorname{Res}_A^f : X \to 2^X$ given by

$$\operatorname{Res}_A^f := \operatorname{Prt}_A^f \circ \nabla f.$$

This operator is single-valued when A is monotone and f is strictly convex on $\operatorname{int}\operatorname{dom} f$. If $A = \partial\varphi$, where φ is a proper, lower semicontinuous and convex function, then we denote

$$\operatorname{prox}_\varphi^f := \operatorname{Prt}_{\partial\varphi}^f \quad \text{and} \quad \operatorname{prox}_\varphi^f := \operatorname{Res}_{\partial\varphi}^f.$$

If C is a nonempty, closed and convex subset of X, then the indicator function ι_C of C, that is, the function

$$\iota_C(x) := \begin{cases} 0 & \text{if } x \in C \\ +\infty & \text{if } x \notin C, \end{cases}$$

is proper, convex and lower semicontinuous, and therefore $\partial\iota_C$ exists and is a maximal monotone operator with domain C. The operator $\mathrm{prox}^f_{\iota_C}$ is called the *Bregman projection* onto C with respect to f (cf. [15]) and we denote it by proj^f_C.

We denote the closure of a subset K of X by $\mathrm{cl}(K)$. For each x and u in $\mathrm{int}\,\mathrm{dom}\, f$, set

$$H(x,u) = \{y \in \mathrm{cl}(\mathrm{dom}\, A) \mid \langle \nabla f(x) - \nabla f(u), y - u\rangle \le 0\}$$

and let

$$H = \bigcap_{u \in \mathrm{Res}^f_A(x)} H(x,u).$$

Proposition 2.3. *Let $A : X \to 2^{X^*}$ be a monotone mapping which satisfies the range condition*

$$\nabla f(\mathrm{cl}(\mathrm{dom}\, A)) \subset \mathrm{ran}(\nabla f + A). \tag{9}$$

Then

$$\mathrm{Fix}\,\mathrm{Res}^f_A = H \bigcap (\mathrm{int}\,\mathrm{dom}\, f).$$

If, in addition, $\mathrm{cl}(\mathrm{dom}\, A)$ *is convex, then* $\mathrm{Fix}\,\mathrm{Res}^f_A$ *is convex too.*

Proof. If $y \in \mathrm{Fix}\,\mathrm{Res}^f_A$, then $y = \mathrm{Res}^f_A(y)$ and therefore $0^* \in A(y)$, since

$$y = \mathrm{Res}^f_A(y) \Leftrightarrow y = (\nabla f + A)^{-1}\nabla f(y) \Leftrightarrow \nabla f(y) \in \nabla f(y) + A(y) \Longleftrightarrow 0^* \in A(y).$$

Hence $y \in \mathrm{cl}(\mathrm{dom}\, A)$. Take $(x,u) \in \mathrm{Graph}\,\mathrm{Res}^f_A$. Then $u \in (\nabla f + A)^{-1}\nabla f(x)$ and therefore $\nabla f(x) - \nabla f(u) \in A(u)$. Denote $\nabla f(x) - \nabla f(u)$ by $\xi \in A(u)$. The monotonicity of A implies that

$$\langle \nabla f(x) - \nabla f(u), y - u\rangle = \langle \xi - 0^*, y - u\rangle \le 0.$$

Thus $y \in H(x,u)$ for every $(x,u) \in \mathrm{Graph}\,\mathrm{Res}^f_A$. This means that $y \in H$. Conversely, take $y \in H \cap (\mathrm{int}\,\mathrm{dom}\, f)$. Then $y \in \mathrm{cl}(\mathrm{dom}\, A) \subset \mathrm{dom}\,\mathrm{Res}^f_A$ and $y \in \cap_{u \in \mathrm{Res}^f_A(y)} H(y,u)$. Hence

$$\langle \nabla f(y) - \nabla f(u), y - u\rangle \le 0 \quad \text{for } u \in \mathrm{Res}^f_A(y).$$

The operator ∇f is strictly monotone on $\mathrm{int}\,\mathrm{dom}\, f$ because f is strictly convex on $\mathrm{int}\,\mathrm{dom}\, f$ and therefore $u = y$. That is $y = \mathrm{Res}^f_A(y)$ and $y \in \mathrm{Fix}\,\mathrm{Res}^f_A$.

If, in addition, $\mathrm{cl}(\mathrm{dom}\, A)$ is a convex set, then the convexity of $\mathrm{Fix}\,\mathrm{Res}^f_A$ follows from the fact that $\mathrm{Fix}\,\mathrm{Res}^f_A = H \cap (\mathrm{int}\,\mathrm{dom}\, f)$ because $H(x,u)$ is convex for any $(x,u) \in \mathrm{Graph}\,\mathrm{Res}^f_A$. □

Our new range condition (9) is analogous to and weaker than other range conditions which appear in the literature. This is illustrated by the following examples.

If the Banach space X is a Hilbert space H and the function f is $(1/2)\|\cdot\|^2$, then our range condition (9) becomes the range condition

$$\mathrm{cl}(\mathrm{dom}\, A)) \subset \mathrm{ran}(I + A),$$

which is well known in semigroup theory (see, for example, [25, 27]).

In [4, Prop. 3.8(iv)(c), p. 604] Bauschke, Borwein and Combettes use the range condition

$$\operatorname{ran}(\nabla f) \subset \operatorname{ran}(\nabla f + A), \tag{10}$$

which is a stronger than our range condition (9).

If A is a maximal monotone operator, $\operatorname{dom} A \subset \operatorname{int} \operatorname{dom} f$ and $A^{-1}(0^*) \neq \varnothing$, then $\operatorname{dom} \operatorname{Res}_A^f = X$ (see [4, Proposition 3.14(ii), p. 606]). Hence $\operatorname{ran}(\nabla f) \subset \operatorname{ran}(\nabla f + A)$. This means that (10) holds and therefore (9) holds too.

If A is a maximal monotone operator, and ∇f is bounded on bounded subsets of X and coercive, then $\nabla f + A$ is surjective, that is, $\operatorname{ran}(\nabla f + A) = X^*$ (see [17, Theorem 3.4, p. 163]). Thus our range condition (9) certainly holds in this case.

2.4. A special set of functions.

By $\mathcal{F}_f$ we denote the set of proper, lower semicontinuous and convex functions $\varphi : X \to (-\infty, +\infty]$ which satisfy the following two conditions:

$$\operatorname{dom} \varphi \bigcap \operatorname{int} \operatorname{dom} f \neq \varnothing$$

and

$$\varphi_f := \inf \{\varphi(x) : x \in \operatorname{dom} \varphi \cap \operatorname{dom} f\} > -\infty.$$

With any Legendre function f we associate the function $W^f : X^* \times X \to [0, +\infty]$ defined by

$$W^f(\xi, x) = f(x) - \langle \xi, x \rangle + f^*(\xi).$$

Proposition 2.4 (cf. [12, Lemma 2.1, p. 2101]). *Suppose that $\varphi \in \mathcal{F}_f$. Then for any $\xi \in \operatorname{int} \operatorname{dom} f^*$, there exists a unique global minimizer of the function $\varphi(\cdot) + W^f(\xi, \cdot)$ which is exactly $\operatorname{Prox}_\varphi^f(\xi)$. The vector $\operatorname{prox}_\varphi^f(\xi)$ belongs to $\operatorname{dom} \partial\varphi \cap \operatorname{int} \operatorname{dom} f$ and we have*

$$\operatorname{prox}_\varphi^f(\xi) = (\partial\varphi + \nabla f)^{-1}(\xi) = [\partial(\varphi + f)]^{-1}(\xi).$$

3. A Strong Convergence Theorem for the Resolvent

In this section we give sufficient conditions for strong convergence of resolvents which require neither the maximal monotonicity of A nor any smoothness properties of the space X.

Theorem 3.1. *Let $A : X \to 2^{X^*}$ be a demiclosed monotone operator with nonempty zero set $A^{-1}(0^*)$. Assume that $\operatorname{cl}(\operatorname{dom} A)$, the closure of $\operatorname{dom} A$, is convex. If $f : X \to \mathbb{R}$ is a totally convex and lower semicontinuous Legendre function, which is bounded on bounded subsets of X, and satisfies the range condition*

$$\nabla f(\operatorname{cl}(\operatorname{dom} A)) \subset \operatorname{ran}(\nabla f + \lambda A), \quad \forall \lambda > 0, \tag{11}$$

then, for each $x \in \operatorname{cl}(\operatorname{dom} A)$, the Bregman projection $\operatorname{proj}_{A^{-1}(0^*)}^f(x)$ exists, the net

$$x_\lambda = \operatorname{Res}_{\lambda A}^f(x), \quad \lambda > 0, \tag{12}$$

is well defined, and converges strongly to $\operatorname{proj}_{A^{-1}(0^*)}^f(x)$ as $\lambda \to +\infty$.

Proof. First we note that the net $\{x_\lambda\}_{\lambda>0}$ given by (12) is well defined for any $x \in \mathrm{cl}\,(\mathrm{dom}\,A)$ because

$$\begin{aligned} x \in \mathrm{cl}\,(\mathrm{dom}\,A) &\Longrightarrow \nabla f(x) \in \nabla f(\mathrm{cl}\,(\mathrm{dom}\,A)) \\ &\Longrightarrow \nabla f(x) \in \mathrm{ran}(\nabla f + \lambda A) \\ &\Longrightarrow x \in \mathrm{dom}\,\mathrm{Res}^f_{\lambda A}\,. \end{aligned}$$

Next, by Proposition 2.3 and the fact that $\mathrm{Fix}\,\mathrm{Res}^f_A = A^{-1}(0^*) \cap (\mathrm{int}\,\mathrm{dom}\,f) = A^{-1}(0^*)$ (since in our case $\mathrm{int}\,\mathrm{dom}\,f = X$), the zero set $A^{-1}(0^*)$ of A is convex. Since the operator A is demiclosed, the set $A^{-1}(0^*)$ is also closed, because, if $\{u_n\}_{n\in\mathbb{N}}$ is a sequence in $A^{-1}(0^*)$ with $u_0 = \lim_{n\to+\infty} u_n$, then $(u_n, 0^*) \in \mathrm{graph}\,A$ for all $n \in \mathbb{N}$ and, therefore, since u_0 is also the weak limit of $\{u_n\}_{n\in\mathbb{N}}$, it follows that it belongs to $A^{-1}(0^*)$. Finally, since $A^{-1}(0^*)$ is nonempty, closed and convex, the Bregman projection $\mathrm{proj}^f_{A^{-1}(0^*)}(x)$ exists. We are now going to establish our theorem by successively proving the following four claims.

Claim 1: The net $\{x_\lambda\}_{\lambda>0}$ is bounded. In order to prove this claim, observe that, by (12), for each positive number λ, $\nabla f(x) \in \nabla f(x_\lambda) + \lambda A(x_\lambda)$. Hence, for each $\lambda > 0$, there exists $\xi_\lambda \in Ax_\lambda$ such that

$$\nabla f(x) = \nabla f(x_\lambda) + \lambda \xi_\lambda. \tag{13}$$

From the three point identity (6), for any $y \in X$, we have

$$D_f(x_\lambda, x) = D_f(y, x) - D_f(y, x_\lambda) + \langle \nabla f(x) - \nabla f(x_\lambda), y - x_\lambda \rangle\,.$$

Since A is a monotone operator, for every $y \in A^{-1}(0^*)$, it follows from (13) that

$$\begin{aligned} D_f(x_\lambda, x) &\le D_f(y, x) + \langle \lambda \xi_\lambda, y - x_\lambda \rangle \\ &= D_f(y, x) - \lambda \langle 0^* - \xi_\lambda, y - x_\lambda \rangle \\ &\le D_f(y, x). \end{aligned} \tag{14}$$

Hence the net $\{D_f(x_\lambda, x)\}_{\lambda>0}$ is bounded by $D_f(y, x)$ for any $y \in A^{-1}(0^*)$. Therefore the net $\{\nu_f(x, \|x_\lambda - x\|)\}_{\lambda>0}$ is bounded by $D_f(y, x)$, since from the definition of the modulus of total convexity (see (6)) and from (14) we get

$$\nu_f(x, \|x_\lambda - x\|) \le D_f(x_\lambda, x) \le D_f(y, x). \tag{15}$$

Since the function f is totally convex, the function $\nu_f(x, \cdot)$ is strictly increasing and positive on $(0, \infty)$ (cf. Proposition 2.1(iv)). It is not difficult to see that this implies that the net $\{x_\lambda\}_{\lambda>0}$ is indeed bounded, as claimed.

Observe that, since X is reflexive and $\{x_\lambda\}_{\lambda>0}$ is bounded, it follows that the net $\{x_\lambda\}_{\lambda>0}$ has weak sequential limit points. Observe also that ∇f is bounded on bounded subsets of X (see [10, Proposition 1.1.11, p. 16]).

Claim 2: Every weak sequential limit point of $\{x_\lambda\}_{\lambda>0}$ as $\lambda \to +\infty$ belongs to $A^{-1}(0^*)$. Let x_0 be a weak sequential limit point of $\{x_\lambda\}_{\lambda>0}$ as $\lambda \to +\infty$. Then there is a sequence $\{x_{\lambda_n}\}_{n\in\mathbb{N}} \subset \{x_\lambda\}_{\lambda>0}$ such that $\lambda_n \to +\infty$ and $x_{\lambda_n} \rightharpoonup x_0$. From (13) we have that

$$\xi_\lambda = \frac{\nabla f(x) - \nabla f(x_\lambda)}{\lambda}, \quad \forall \lambda > 0.$$

The net $\{\nabla f(x_\lambda)\}_{\lambda>0}$ is bounded because ∇f is bounded on bounded subsets of X and $\{x_\lambda\}_{\lambda>0}$ is bounded by *Claim 1*. Therefore $\xi_\lambda \to 0^*$ as $\lambda \to +\infty$. Hence, from the demiclosededness of A we have that $x_0 \in A^{-1}(0^*)$ because $(x_{\lambda_n}, \xi_{\lambda_n}) \in \operatorname{graph} A$ for each $n \in \mathbb{N}$, $x_{\lambda_n} \rightharpoonup x_0$ and $\xi_{\lambda_n} \to 0^*$. This proves *Claim 2*.

Now we are going to use this fact in order to establish our next claim.

Claim 3: The net $\{x_\lambda\}_{\lambda>0}$ has a weak limit which is exactly $\operatorname{proj}^f_{A^{-1}(0^*)}(x)$ as $\lambda \to +\infty$. Let $\{x_{\lambda_n}\}_{n\in\mathbb{N}}$ converge weakly to x_0, where $\lambda_n \to +\infty$. In order to prove this claim, we use the weak lower semicontinuity of the function $D_f(\cdot, x)$ which follows from the weak lower semicontinuity of f. This and (14) imply that

$$D_f(x_0, x) \le \liminf_{n\to+\infty} D_f(x_{\lambda_n}, x) \le D_f(y, x), \quad \forall y \in A^{-1}(0^*).$$

Since y is an arbitrary element of $A^{-1}(0^*)$, this shows that x_0 is the minimizer of $D_f(\cdot, x)$ on the closed and convex set $A^{-1}(0^*)$. Note that

$$D_f(y, x) + \iota_{A^{-1}(0^*)}(y) = W^f(\nabla f(x), y) + \iota_{A^{-1}(0^*)}(y),$$

and by Proposition 2.4 we know that the minimum of $W^f(\nabla f(x), y) + \iota_{A^{-1}(0^*)}(y)$ is exactly $\operatorname{proj}^f_{A^{-1}(0^*)}(x)$. Hence, any weak sequential limit point x_0 of $\{x_\lambda\}_{\lambda>0}$ coincides with $\operatorname{proj}^f_{A^{-1}(0^*)}(x)$. This implies that $\{x_\lambda\}_{\lambda>0}$ itself converges weakly to $\operatorname{proj}^f_{A^{-1}(0^*)}(x)$ as $\lambda \to +\infty$, as claimed.

Finally, we are going to establish strong convergence of the net $\{x_\lambda\}_{\lambda>0}$.

Claim 4: The net $\{x_\lambda\}_{\lambda>0}$ converges strongly to $\operatorname{proj}^f_{A^{-1}(0^*)}(x)$ as $\lambda \to +\infty$. In order to prove this claim, we take $y = \operatorname{proj}^f_{A^{-1}(0^*)}(x)$ in (14). In view of *Claim 3* and the weak lower semicontinuity of $D_f(\cdot, x)$, we obtain

$$\begin{aligned} D_f(\operatorname{proj}^f_{A^{-1}(0^*)}(x), x) \le\ & \liminf_{\lambda\to+\infty} D_f(x_\lambda, x) \\ \le\ & \limsup_{\lambda\to+\infty} D_f(x_\lambda, x) \le D_f(\operatorname{proj}^f_{A^{-1}(0^*)}(x), x). \end{aligned}$$

Hence

$$\lim_{\lambda\to+\infty} D_f(x_\lambda, x) = D_f(\operatorname{proj}^f_{A^{-1}(0^*)}(x), x). \tag{16}$$

By using again the three point identity (6), we deduce that

$$\begin{aligned} & D_f(x_\lambda, \operatorname{proj}^f_{A^{-1}(0^*)}(x)) \\ =\ & \left[D_f(x_\lambda, x) - D_f(\operatorname{proj}^f_{A^{-1}(0^*)}(x), x)\right] \\ & + \left\langle \nabla f(x) - \nabla f(\operatorname{proj}^f_{A^{-1}(0^*)}(x)), x_\lambda - \operatorname{proj}^f_{A^{-1}(0^*)}(x) \right\rangle. \end{aligned}$$

Observe that the quantity between square brackets converges to zero as $\lambda \to +\infty$ by (16). Also, the inner product on the right-hand side of this equality converges to zero by *Claim 3*. This implies that the net $\left\{D_f(x_\lambda, \operatorname{proj}^f_{A^{-1}(0^*)}(x))\right\}_{\lambda>0}$ converges to zero as $\lambda \to +\infty$. Since f is a totally convex function, this fact and Proposition 2.2 show that the net $\{x_\lambda\}_{\lambda>0}$ converges strongly to $\operatorname{proj}^f_{A^{-1}(0^*)}(x)$ as $\lambda \to +\infty$. This proves *Claim 4* and completes the proof of the theorem itself. □

4. Consequences of the Strong Convergence Theorem

An interesting particular instance of Theorem 3.1 occurs when A is a maximal monotone operator. In this case λA is also maximal monotone for any $\lambda > 0$ and, consequently $\operatorname{dom} \operatorname{Res}^f_{\lambda A} = X$ (see [4, Proposition 3.14(ii), p. 606]). Therefore it follows that $\operatorname{ran}(\nabla f) \subset \operatorname{ran}(\nabla f + \lambda A)$ for any $\lambda > 0$. This means that (11) holds. Also, from the maximal monotonicity of A it follows that $\operatorname{cl}(\operatorname{dom} A)$ is convex (see [19, Proposition 2.3.1, p. 327]). Therefore Theorem 3.1 yields the following corollary.

Corollary 4.1. *Let $A : X \to 2^{X^*}$ be a maximal monotone operator such that $A^{-1}(0^*) \neq \varnothing$. If $f : X \to \mathbb{R}$ is a totally convex and lower semicontinuous Legendre function which is bounded on bounded subsets of X, then, for any $x \in X$, the Bregman projection $\operatorname{proj}^f_{A^{-1}(0^*)}(x)$ exists, the net*

$$x_\lambda = \operatorname{Res}^f_{\lambda A}(x), \quad \lambda > 0,$$

is well defined, and converges strongly to $\operatorname{proj}^f_{A^{-1}(0^*)}(x)$ as $\lambda \to +\infty$.

Two other interesting corollaries of Theorem 3.1 occur when the space X is smooth and has the Kadec-Klee property, that is,

$$(x_n \rightharpoonup x \text{ and } \|x_n\| \to \|x\|) \Longrightarrow x_n \to x. \tag{17}$$

In this case the function $f(x) = (1/2)\|x\|^2$ is Legendre (cf. [3, Lemma 6.2, p. 24]) and ∇f is exactly the duality mapping J_X of the space X. If A is also a maximal monotone mapping, then λA is maximal monotone for any $\lambda > 0$ and, consequently, the mapping $\nabla f + \lambda A$ is surjective (cf. [17, Theorem 3.11, p. 166]). Thus the range condition (11) holds. Also, as in Corollary 4.1, $\operatorname{cl}(\operatorname{dom} A)$ is a convex set. According to ([13, Proposition 3.2, p. 17]), since X is reflexive, f is totally convex whenever X has the Kadec-Klee property or, equivalently, X is an E-space. (Recall that a Banach space X is called an *E-space* if it is reflexive, strictly convex and has the Kadec-Klee property.) Therefore Theorem 3.1 applies in this context and leads us to the following two results which, in some sense, complement Theorem 1 in [26] (see also [20]).

Corollary 4.2. *Let X be a smooth Banach space with the Kadec-Klee property (or equivalently, a smooth E-space) and let $A : X \to 2^{X^*}$ be a maximal monotone operator such that $A^{-1}(0^*) \neq \varnothing$. Then, for any $x \in X$, the net*

$$x_\lambda = \operatorname{Res}^{(1/2)\|\cdot\|^2}_{\lambda A}(x), \quad \lambda > 0,$$

is well defined, the Bregman projection $\operatorname{proj}^{(1/2)\|\cdot\|^2}_{A^{-1}(0^*)}(x)$ exists, and $\{x_\lambda\}_{\lambda>0}$ converges strongly to $\operatorname{proj}^{(1/2)\|\cdot\|^2}_{A^{-1}(0^*)}(x)$ as $\lambda \to +\infty$.

Corollary 4.3. *Let X be a smooth Banach space with the Kadec-Klee property (or equivalently, a smooth E-space) and let $A : X \to 2^{X^*}$ be a maximal monotone operator such that $A^{-1}(0^*) \neq \varnothing$. Then, for any $\eta \in X^*$, the net*

$$x_\lambda = \operatorname{Prt}^{(1/2)\|\cdot\|^2}_{\lambda A}(\eta), \quad \lambda > 0,$$

is well defined, the Bregman projection $\mathrm{proj}_{A^{-1}(0^*)}^{(1/2)\|\cdot\|^2}(J_{X^*}(\eta))$ *exists, and* $\{x_\lambda\}_{\lambda>0}$ *converges strongly to* $\mathrm{proj}_{A^{-1}(0^*)}^{(1/2)\|\cdot\|^2}(J_{X^*}(\eta))$ *as* $\lambda \to +\infty$.

Proof. The operator J_X is surjective because X is reflexive. Therefore for any $\eta \in X^*$, there exists $x \in X$ such that $J_X(x) = \eta$. It follows from Corollary 4.2 that the net $\{x_\lambda\}_{\lambda>0}$ is well defined and converges strongly as $\lambda \to +\infty$ to

$$\mathrm{proj}_{A^{-1}(0^*)}^{(1/2)\|\cdot\|^2}(x) = \mathrm{proj}_{A^{-1}(0^*)}^{(1/2)\|\cdot\|^2}(J_{X^*}(\eta)).$$

5. An Extension of the Strong Convergence Theorem

The next result extends Theorem 3.1 to the case where A is approximated in some sense by more regular operators $\{A_n\}_{n\in\mathbb{N}}$. More precisely, we say that a sequence $\{A_n\}_{n\in\mathbb{N}}$ of operators from X into 2^{X^*} approximates A regularly at 0^* if the following two conditions hold:

(*S1*) *If* $\{(x_k, \xi_k)\}_{k\in\mathbb{N}}$ *is a sequence in* $X \times X^*$, *then*

$$\left.\begin{array}{c} x_k \rightharpoonup x \\ \xi_k \in A_k x_k, \quad k \in \mathbb{N} \\ \xi_k \to 0^* \end{array}\right\} \Longrightarrow 0^* \in Ax.$$

(*S2*) *For any* y *in* $A^{-1}(0^*)$, *there exists a sequence* $\{y_n\}_{n\in\mathbb{N}}$ *with* $y_n \in A_n^{-1}(0^*)$ *for each* $n \in \mathbb{N}$, *such that* $\lim_{n\to+\infty} y_n = y$.

The notion of regular approximation at 0^* is related to the well-known notion of graph convergence. In fact, in [1, Prop. 3.59, p. 361] Attouch proves that if A and $\{A_n\}_{n\in\mathbb{N}}$ are maximal monotone operators and the sequence $\{A_n\}_{n\in\mathbb{N}}$ is graph convergent to A, then condition (*S1*) holds. In addition, it is clear that condition (*S2*) is equivalent to the inclusion $A^{-1}(0^*) \subset \mathrm{Li}_n A_n^{-1}(0^*)$, where $y \in \mathrm{Li}_n C_n$ if there exists a sequence $\{y_n\}_{n\in\mathbb{N}}$ in X such that $\lim_{n\to+\infty} y_n = y$ and $y_n \in C_n$ for all $n \in \mathbb{N}$.

Theorem 5.1. *Let* $A, A_n : X \to 2^{X^*}$, $n \in \mathbb{N}$, *be monotone operators such that the sequence* $\{A_n\}_{n\in\mathbb{N}}$ *approximates* A *regularly at* 0^*. *Let* $f : X \to \mathbb{R}$ *be a totally convex and Legendre function which is bounded on bounded subsets of* X. *Assume that* (11) *holds and the following conditions are satisfied:*

(*A*) $\{\lambda_n\}_{n\in\mathbb{N}}$ *is a sequence of positive real numbers with* $\lim_{n\to+\infty} \lambda_n = +\infty$ *such that*

$$\nabla f(\mathrm{cl}(\mathrm{dom}\, A_n)) \subset \mathrm{ran}(\nabla f + \lambda_n A_n), \quad n \in \mathbb{N}. \tag{18}$$

(*B*) *The sets* $A^{-1}(0^*)$ *and* $A_n^{-1}(0^*)$ *are nonempty,* $A^{-1}(0^*)$ *is closed,* $\mathrm{cl}(\mathrm{dom}\, A)$ *is convex, and* $clx(\mathrm{dom}\, A) \subset \cap_{n\in\mathbb{N}} \mathrm{cl}(\mathrm{dom}\, A_n)$.

Then, for any $x \in \mathrm{cl}(\mathrm{dom}\, A)$, *the Bregman projection* $\mathrm{proj}^f_{A^{-1}(0^*)}(x)$ *exists, the sequence*

$$x_n = \mathrm{Res}^f_{\lambda_n A_n}(x), \quad n \in \mathbb{N}, \tag{19}$$

is well defined and converges strongly to $\mathrm{proj}^f_{A^{-1}(0^*)}(x)$ *as* $n \to +\infty$.

Proof. The sequence $\{x_n\}_{n\in\mathbb{N}}$ given by (19) is well defined for any $x \in \mathrm{cl}(\mathrm{dom}\, A) \subset \cap_{n\in\mathbb{N}} \mathrm{cl}(\mathrm{dom}\, A_n)$ because

$$\begin{aligned} & x \in \cap_{n\in\mathbb{N}} \mathrm{cl}(\mathrm{dom}\, A_n) \Rightarrow x \in \mathrm{cl}(\mathrm{dom}\, A_n) \\ \Longrightarrow\; & \nabla f(x) \in \nabla f(\mathrm{cl}(\mathrm{dom}\, A_n)) \\ \Longrightarrow\; & \nabla f(x) \in \mathrm{ran}(\nabla f + \lambda_n A_n) \Longrightarrow x \in \mathrm{dom}\, \mathrm{Res}^f_{\lambda_n A_n}. \end{aligned}$$

The set $A^{-1}(0^*)$ is closed and nonempty, and in view of Proposition 2.3 and the fact that $\mathrm{Fix}\,\mathrm{Res}^f_A = A^{-1}(0^*) \cap (\mathrm{int}\,\mathrm{dom}\, f) = A^{-1}(0^*)$ (since in our case $\mathrm{int}\ \mathrm{dom}\, f = X$), the zero set $A^{-1}(0^*)$ of A is convex. Therefore the Bregman projection $\mathrm{proj}^f_{A^{-1}(0^*)}(x)$ exists.

Claim 1: The sequence $\{x_n\}_{n\in\mathbb{N}}$ is bounded. In order to prove this claim, observe that, by (19), for each positive integer n, $\nabla f(x) \in \nabla f(x_n) + \lambda_n A(x_n)$. Hence, for each $n \in \mathbb{N}$, there exists $\xi_n \in A_n x_n$ such that

$$\nabla f(x) = \nabla f(x_n) + \lambda_n \xi_n. \tag{20}$$

By Condition (*S2*), given any $y \in A^{-1}(0^*)$, there exists a sequence $\{y_n\}_{n\in\mathbb{N}}$ such that $y_n \in A_n^{-1}(0^*)$ for each $n \in \mathbb{N}$ and $\lim_{n\to+\infty} y_n = y$. Therefore, from the three point identity (6), it follows that for each $n \in \mathbb{N}$, we have

$$D_f(x_n, x) = D_f(y_n, x) - D_f(y_n, x_n) + \langle \nabla f(x) - \nabla f(x_n), y_n - x_n \rangle . \tag{21}$$

Since A_n is a monotone operator, taking into account (20) and (21) for any $n \in \mathbb{N}$, we obtain

$$\begin{aligned} D_f(x_n, x) &= D_f(y_n, x) - D_f(y_n, x_n) + \langle \nabla f(x) - \nabla f(x_n), y_n - x_n \rangle \\ &\le D_f(y_n, x) + \langle \lambda_n \xi_n, y_n - x_n \rangle \\ &= D_f(y_n, x) - \lambda_n \langle 0^* - \xi_n, y_n - x_n \rangle \\ &\le D_f(y_n, x). \end{aligned} \tag{22}$$

The function $D_f(\cdot, x)$ is continuous on X because it is lower semicontinuous and convex with domain X. Thus the sequence $\{D_f(y_n, x)\}_{n\in\mathbb{N}}$ converges to $D_f(y, x)$. Hence the sequence $\{D_f(x_n, x)\}_{n\in\mathbb{N}}$ is also bounded (see (22)). It follows that the sequence $\{\nu_f(x, \|x_n - x\|)\}_{n\in\mathbb{N}}$ is bounded too. Indeed, using the definition of the modulus of total convexity (see (7)) and (22), we obtain that

$$\nu_f(x, \|x_n - x\|) \le D_f(x_n, x) \le D_f(y_n, x), \quad \forall n \in \mathbb{N}. \tag{23}$$

The function f is totally convex and therefore the function $\nu_f(x, \cdot)$ is strictly increasing and positive on $(0, \infty)$ (cf. Proposition 2.1(*iv*)). As in the case of *Claim 1* in the proof of Theorem 3.1, it is not difficult to see that this implies that the sequence $\{x_n\}_{n\in\mathbb{N}}$ is indeed bounded, as claimed.

Observe that, since X is reflexive and $\{x_n\}_{n\in\mathbb{N}}$ is bounded, the sequence $\{x_n\}_{n\in\mathbb{N}}$ has weak subsequential limit points. Recall also that ∇f is bounded on bounded subsets of X (see [10, Proposition 1.1.11, p. 16]).

Claim 2: Every weak subsequential limit point of $\{x_n\}_{n\in\mathbb{N}}$ *belongs to* $A^{-1}(0^*)$. In order to prove this claim, we take a subsequence $\{x_{n_k}\}_{k\in\mathbb{N}}$ of $\{x_n\}_{n\in\mathbb{N}}$ which converges weakly to some $\bar{x} \in X$. From (20) it follows that

$$\xi_n = \frac{\nabla f(x) - \nabla f(x_n)}{\lambda_n}.$$

The sequence $\{\nabla f(x_n)\}_{n\in\mathbb{N}}$ is bounded because ∇f is bounded on bounded subsets of X and $\{x_n\}_{n\in\mathbb{N}}$ is bounded by *Claim 1*. Therefore, $\xi_n \to 0^*$. Since $\xi_{n_k} \in A_{n_k}x_{n_k}$ and $x_{n_k} \rightharpoonup \bar{x}$, an application of Condition (*S1*) implies that $\bar{x} \in A^{-1}(0^*)$. This proves *Claim 2*.

Claim 3: The sequence $\{x_n\}_{n\in\mathbb{N}}$ *has a weak limit which is exactly* $\mathrm{proj}^f_{A^{-1}(0^*)}(x)$ *as* $n \to +\infty$. Let $\{x_{n_k}\}_{k\in\mathbb{N}}$ be a subsequence of $\{x_n\}_{n\in\mathbb{N}}$ which converges weakly to $\bar{x}$. In order to prove this claim, we invoke the weak lower semicontinuity of the function $D_f(\cdot, x)$ which is a consequence of the weak lower semicontinuity of f. This and (22) imply that

$$D_f(\bar{x}, x) \le \liminf_{k\to+\infty} D_f(x_{n_k}, x) \le \liminf_{k\to+\infty} D_f(y_{n_k}, x) = D_f(y, x), \quad y \in A^{-1}(0^*).$$

Since y is an arbitrary element of $A^{-1}(0^*)$, this shows that $\bar{x}$ is the minimizer of $D_f(\cdot, x)$ on the closed and convex set $A^{-1}(0^*)$. Note that

$$D_f(y, x) + \iota_{A^{-1}(0^*)}(y) = W^f(\nabla f(x), y) + \iota_{A^{-1}(0^*)}(y),$$

and by Proposition 2.3 we know that the minimum of $W^f(\nabla f(x), \cdot) + \iota_{A^{-1}(0^*)}(\cdot)$ is exactly $\mathrm{proj}^f_{A^{-1}(0^*)}(x)$. Hence, any weak subsequential limit point $\bar{x}$ of $\{x_n\}_{n\in\mathbb{N}}$ coincides with $\mathrm{proj}^f_{A^{-1}(0^*)}(x)$. This implies that the sequence $\{x_n\}_{n\in\mathbb{N}}$ itself converges weakly to $\mathrm{proj}^f_{A^{-1}(0^*)}(x)$ as $n \to +\infty$, as claimed.

Claim 4: The sequence $\{x_n\}_{n\in\mathbb{N}}$ *converges strongly to* $\mathrm{proj}^f_{A^{-1}(0^*)}(x)$ *as* $n \to +\infty$. In order to prove this claim, take $\{y_n\}_{n\in\mathbb{N}}$ to be a sequence with $y_n \in A_n^{-1}(0^*)$ for each $n \in \mathbb{N}$ such that $\lim_{n\to+\infty} y_n = \mathrm{proj}^f_{A^{-1}(0^*)}(x)$. Such a sequence exists by Condition (*S2*). From *Claim 3*, (22) and the weak lower semicontinuity of $D_f(\cdot, x)$, we get

$$\begin{aligned} D_f(\mathrm{proj}^f_{A^{-1}(0^*)}(x), x) &\le \liminf_{n\to+\infty} D_f(x_n, x) \\ &\le \limsup_{n\to+\infty} D_f(x_n, x) \le \limsup_{n\to+\infty} D_f(y_n, x) \\ &= D_f(\mathrm{proj}^f_{A^{-1}(0^*)}(x), x). \end{aligned}$$

Hence

$$\lim_{n\to+\infty} D_f(x_n, x) = D_f(\mathrm{proj}^f_{A^{-1}(0^*)}(x), x). \tag{24}$$

Using again the three point identity (6), we infer that

$$\begin{aligned} & D_f(x_n, \mathrm{proj}^f_{A^{-1}(0^*)}(x)) \\ = & \left[D_f(x_n, x) - D_f(\mathrm{proj}^f_{A^{-1}(0^*)}(x), x)\right] \\ & \quad + \left\langle \nabla f(x) - \nabla f(\mathrm{proj}^f_{A^{-1}(0^*)}(x)), x_n - \mathrm{proj}^f_{A^{-1}(0^*)}(x) \right\rangle. \end{aligned}$$

Observe that the quantity between square brackets converges to zero by (24). Also, the inner product on the right-hand side of this equality converges to zero by *Claim 3*. This implies that the sequence $\left\{D_f(x_n, \operatorname{proj}^f_{A^{-1}(0^*)}(x))\right\}_{n\in\mathbb{N}}$ also converges to zero as $n \to +\infty$. Since f is totally convex, an application of Proposition 2.2 shows that the sequence $\{x_n\}_{n\in\mathbb{N}}$ itself converges strongly to $\operatorname{proj}^f_{A^{-1}(0^*)}(x)$ as $n \to +\infty$. This proves *Claim 4* and completes the proof of the theorem itself. □

The following two propositions show that Conditions (*S1*) and (*S2*) are in some sense necessary (if f is Fréchet differentiable) for the conclusions of Theorem 5.1 to hold (cf. [21]).

Proposition 5.2. *Let $A, A_n : X \to 2^{X^*}$, $n \in \mathbb{N}$, be monotone operators. Let $f : X \to \mathbb{R}$ be a Fréchet differentiable and totally convex Legendre function which is bounded on bounded subsets of X. Assume that* (11) *holds and the following two conditions are satisfied:*

(*A*) *For each $\lambda > 0$ and $n \in \mathbb{N}$,*

$$\operatorname{ran}(\nabla f) \subset \operatorname{ran}(\nabla f + \lambda A_n).$$

(*B*) *The sets $A^{-1}(0^*)$ and $A_n^{-1}(0^*)$ are nonempty, $A^{-1}(0^*)$ is closed and* $\operatorname{cl}(\operatorname{dom} A)$ *is convex.*

Suppose that, for any $x \in X$ and any sequence $\{\lambda_n\}_{n\in\mathbb{N}}$ with $\lim_{n\to+\infty} \lambda_n = +\infty$, the sequence $\left\{\operatorname{Res}^f_{\lambda_n A_n}(x)\right\}_{n\in\mathbb{N}}$ converges strongly to $\operatorname{proj}^f_{A^{-1}(0^*)}(x)$ as $n \to +\infty$. Then Condition (*S1*) *holds.*

Proof. Note first that $\operatorname{proj}^f_{A^{-1}(0^*)}(x)$ is well defined for any $x \in X$ because $A^{-1}(0^*)$ is convex by Proposition 2.3. Suppose that $\{\xi_k\}_{k\in\mathbb{N}}$ is a sequence in X^* which converges strongly to 0^* and $\{x_k\}_{k\in\mathbb{N}}$ is a sequence in X which converges weakly to some x such that $\xi_k \in A_k x_k$ for all $k \in \mathbb{N}$. Let $\{\lambda_k\}_{k\in\mathbb{N}}$ be a sequence of positive real numbers such that $\lambda_k \to +\infty$ as $k \to +\infty$ and $\{\lambda_k \xi_k\}_{k\in\mathbb{N}}$ converges strongly to 0^*. Set $y_k = \operatorname{Res}^f_{\lambda_k A_k}(x)$ for $k \in \mathbb{N}$. Since

$$\frac{\nabla f(x) - \nabla f(y_k)}{\lambda_k} \in A_k y_k$$

for each $k \in \mathbb{N}$, it follows from the monotonicity of A_k that

$$\left\langle \frac{\nabla f(x) - \nabla f(y_k)}{\lambda_k} - \xi_k, y_k - x_k \right\rangle \geq 0, \quad k \in \mathbb{N},$$

which is equivalent to

$$\langle \nabla f(x) - \nabla f(y_k) - \lambda_k \xi_k, y_k - x_k \rangle \geq 0, \quad k \in \mathbb{N}.$$

Letting $k \to +\infty$, we obtain that

$$\left\langle \nabla f(x) - \nabla f(\operatorname{proj}^f_{A^{-1}(0^*)}(x)), \operatorname{proj}^f_{A^{-1}(0^*)}(x) - x \right\rangle \geq 0$$

because ∇f is norn-to-norm continuous by [23, Propostion 2.8, p. 19]. Using the strict monotonicity of ∇f, we see that $x = \mathrm{proj}^f_{A^{-1}(0^*)}(x)$ and hence $x \in A^{-1}(0^*)$. This shows that Condition (*S1*) holds, as asserted. □

Proposition 5.3. *Let $A, A_n : X \to 2^{X^*}$, $n \in \mathbb{N}$, be monotone operators. Let $f : X \to \mathbb{R}$ be a totally convex Legendre function which is bounded on bounded subsets of X. Assume that* (11) *holds and the following two conditions are satisfied:*

(*A*) *For each $\lambda > 0$ and $n \in \mathbb{N}$,*

$$\nabla f(\mathrm{cl}(\mathrm{dom}\, A_n)) \subset \mathrm{ran}(\nabla f + \lambda A_n).$$

(*B*) *The sets $A^{-1}(0^*)$ and $A_n^{-1}(0^*)$ are nonempty and closed,* $\mathrm{cl}(\mathrm{dom}\, A)$ *is convex, and* $\mathrm{cl}(\mathrm{dom}\, A) \subset \bigcap_{n\in\mathbb{N}} \mathrm{cl}(\mathrm{dom}\, A_n)$.

Suppose that, for any $x \in \mathrm{cl}(\mathrm{dom}\, A)$ and any sequence $\{\lambda_n\}_{n\in\mathbb{N}}$ with $\lim_{n\to+\infty} \lambda_n = +\infty$, the sequence $\left\{\mathrm{Res}^f_{\lambda_n A_n}(x)\right\}_{n\in\mathbb{N}}$ converges strongly to $\mathrm{proj}^f_{A^{-1}(0^*)}(x)$ as $n \to +\infty$. Then Condition (S2) holds.

Proof. Note first that $\mathrm{proj}^f_{A^{-1}(0^*)}(x)$ and $\mathrm{proj}^f_{A_n^{-1}(0^*)}(x)$ are well defined for any $x \in X$ and $n \in \mathbb{N}$ because $A^{-1}(0^*)$ and $A_n^{-1}(0^*)$ are convex by Proposition 2.3. Let $y \in A^{-1}(0^*)$ and $n \in \mathbb{N}$. By Theorem 3.1, there exists $k_n \in \mathbb{N}$ such that $\lambda_{k_n} > n$ and

$$\left\| \mathrm{Res}^f_{\lambda_{k_n} A_n}(y) - \mathrm{proj}^f_{A_n^{-1}(0^*)}(y) \right\| < \frac{1}{n}.$$

Let $z_n = \mathrm{Res}^f_{\lambda_{k_n} A_n}(y)$ for each $n \in \mathbb{N}$. Since

$$d\left(z_n, A_n^{-1}(0^*)\right) \le \left\| z_n - \mathrm{proj}^f_{A_n^{-1}(0^*)}(y) \right\| < \frac{1}{n},$$

it follows that the sequence $\{d(z_n, A_n^{-1}(0^*))\}_{n\in\mathbb{N}}$ converges to 0 as $n \to +\infty$. On the other hand, from the assumption of the proposition, $\{z_n\}_{n\in\mathbb{N}}$ converges strongly to $\mathrm{proj}^f_{A^{-1}(0^*)}(y) = y$. Therefore we have

$$\lim_{n\to+\infty} d\left(y, A_n^{-1}(0^*)\right) \le \lim_{n\to+\infty} \|y - z_n\| + \lim_{n\to+\infty} d\left(z_n, A_n^{-1}(0^*)\right) = 0.$$

So for each $n \in \mathbb{N}$, there exist $y_n \in A_n^{-1}(0^*)$ such that the sequence $\{y_n\}_{n\in\mathbb{N}}$ converges strongly to y. Hence Condition (*S2*) holds, as asserted. □

Acknowledgements. The very detailed and useful comments of the referee are gratefully acknowledged.

References

[1] H. Attouch: Variational Convergence for Functions and Operators, Pitman, Boston (1984).

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

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

[4] H. H. Bauschke, J. M. Borwein, P. L. Combettes: Bregman monotone optimization algorithms, SIAM J. Control Optimization 42 (2003) 596–636.

[5] H. H. Bauschke, P. L. Combettes, S. Reich: The asymptotic behavior of the composition of two resolvents, Nonlinear Anal. 60 (2005) 283–301.

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

[7] J. F. Bonnans, A. Shapiro: Perturbation Analysis of Optimization Problems, Springer, New York (2000).

[8] R. E. Bruck: A strongly convergent iterative solution of $0 \in U(x)$ for a maximal monotone operator U in Hilbert space, J. Math. Anal. Appl. 48 (1974) 114–126.

[9] R. E. Bruck, S. Reich: Nonexpansive projections and resolvents of accretive operators in Banach spaces, Houston J. Math. 3 (1977) 459–470.

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

[11] 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 Analysis 10 (2003) 35–61.

[12] D. Butnariu, G. Kassay: A proximal-projection method for finding zeros of set-valued operators, SIAM J. Control Optimization 47 (2008) 2096–2136.

[13] D. Butnariu, E. Resmerita: Bregman distances, totally convex functions and a method for solving operator equations in Banach spaces, Abstr. Appl. Anal. 2006 (2006) Art. ID 84919, 39 pp.

[14] D. Butnariu, E. Resmerita: Mosco stability of proximal mappings in reflexive Banach spaces, J. Nonlinear Convex Anal. 8 (2007) 1–10.

[15] L. M. Bregman: The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming, U.S.S.R. Comput. Math. Math. Phys. 7 (1967) 200–217.

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

[17] I. Cioranescu: Geometry of Banach Spaces, Duality Mappings and Nonlinear Problems, Kluwer, Dordrecht (1990).

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

[19] S. Hu, N. S. Papageorgiou: Handbook of Multivalued Analysis. Volume I: Theory, Kluwer, Dordrecht (1997).

[20] K. Kido: Strong convergence of resolvents of monotone operators in Banach spaces, Proc. Amer. Math. Soc. 103 (1988) 755–758.

[21] Y. Kimura: A characterization of strong convergence for a sequence of resolvents of maximal monotone operators, in: Fixed Point Theory and its Applications (Guanajuato, 2005), H. Fetter Nathansky et al. (ed.), Yokohama Publishers, Yokohama (2006) 149–159.

[22] O. Nevanlinna, S. Reich: Strong convergence of contraction semigroups and of iterative methods for accretive operators in Banach spaces, Israel J. Math. 32 (1979) 44–58.

[23] R. R. Phelps: Convex Functions, Monotone Operators, and Differentiability, 2nd Ed., Springer, Berlin (1993).

[24] S. Reich: Approximating zeros of accretive operators, Proc. Amer. Math. Soc. 51 (1975) 381–384.

[25] S. Reich: Product formulas, nonlinear semigroups, and accretive operators, J. Funct. Anal. 36 (1980) 147–168.

[26] S. Reich: Strong convergence theorems for resolvents of accretive operators in Banach spaces, J. Math. Anal. Appl. 75 (1980) 287–292.

[27] S. Reich: Convergence and approximation of nonlinear semigroups, J. Math. Anal. Appl. 76 (1980) 77–83.

[28] 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, A. G. Kartsatos (ed.), Marcel Dekker, New York (1996) 313–318.

[29] E. Resmerita: On total convexity, Bregman projections and stability in Banach spaces, J. Convex Analysis 11 (2004) 1–16.

[30] R. T. Rockafellar: Monotone operators and the proximal point algorithm, SIAM J. Control Optimization 14 (1976) 877–898.

[31] R. T. Rockafellar: Augmented Lagrangians and applications of the proximal point algorithm in convex programming, Math. Oper. Res. 1 (1976) 97–116.

[32] R. T. Rockafellar, R. J.-B. Wets: Variational Analysis, Springer, Berlin (1998).