A Characterization of Bregman Firmly Nonexpansive Operators Using a New Monotonicity Concept

Jonathan M. Borwein[*], Simeon Reich[†] and Shoham Sabach[‡]

February 23, 2011

Abstract

The property of nonexpansivity (1-Lipschitz) is very important in the analysis of many optimization problems. In this paper we study a more general notion of nonexpansivity – Bregman nonexpansivity. We present a characterization of Bregman firmly nonexpansive operators in general reflexive Banach spaces. This characterization allows us to construct Bregman firmly nonexpansive operators explicitly. We provide several examples of such operators with respect to the Boltzmann-Shannon entropy and the Fermi-Dirac entropy in Euclidean spaces. We also compute resolvents with respect to these functions.

AMS 2010 Subject Classification: 26B25, 47H05, 47H09, 52A41.
Key words and phrases. Boltzmann-Shannon entropy, Bregman distance, Bregman firmly nonexpansive operator, Fermi-Dirac entropy, nonexpansive operator, monotone mapping, resolvent, T-monotone mapping, totally convex function.

1 Introduction

Nonexpansive operator theory and monotone mapping theory have transpired to be crucial in both the algorithmic design and analysis of optimization problems. Over

[*]CARMA, University of Newcastle, Australia. Email: jonathan.borwein@newcastle.edu.au.
[†]Department of Mathematics, The Technion – Israel Institute of Technology, 32000 Haifa, Israel. Email: sreich@tx.technion.ac.il.
[‡]Department of Mathematics, The Technion – Israel Institute of Technology, 32000 Haifa, Israel. Email: ssabach@tx.technion.ac.il.

the past few decades many results were published which make the connection between these two notions. In this paper we establish another such result which brings out a connection between the concepts of Bregman firmly nonexpansive operators and T-monotone mappings. This result both improves and generalizes previous results.

One of the earliest results in this direction is due to Rockafellar [30]. Rockafellar was interested in the problem of finding zeroes of a maximal monotone set-valued mapping A in Hilbert spaces. In order to achieve this he used the operator

$$R_A := (I + A)^{-1}.$$

This operator is called the *(classical) resolvent*. The importance of the resolvent stems from its good properties: it is, for instance, single-valued and firmly nonexpansive (hence nonexpansive), and has full domain. Since the fixed point set of R_A is exactly the zero set of A, Rockafellar used a variant of the Picard method in order to approximate fixed points of R_A which are zeroes of A. This method is called the *proximal point algorithm*. It is also known that any firmly nonexpansive operator is the resolvent of a monotone mapping (see [13, 24]).

When we try to extend the theory of resolvents of monotone mappings and firmly nonexpansive operators to general reflexive Banach spaces we encounter several problems. It is well-known that in general the classical resolvent does not enjoy the good properties that we mentioned earlier. In order to overcome these difficulties and extend this theory one uses different types of monotonicity, nonexpansivity and resolvents. In this paper we use Bregman firmly nonexpansive operators and the generalized resolvent. Both of these notions are based on the concept of a Bregman function f. These notions were investigated intensively in the last ten years by Bauschke, Borwein and Combettes [3, 4, 5, 6, 7] who established many properties of the generalized resolvent. Among these is the fact that the resolvent of a monotone mapping is a Bregman firmly nonexpansive operator. The converse implication was only proven two years ago when Bauschke, Wang and Yao showed that any Bregman firmly nonexpansive operator is a generalized resolvent of a monotone mapping. For a recent study of the existence and approximation of fixed points of Bregman firmly nonexpansive operators see [28].

The problem of finding zeroes of monotone mappings is important in the theory of optimization because any minimization problem can be written as a problem of finding zeroes of the subdifferential mapping. In the case of Banach spaces, the application of Bregman distances instead of the norm gives us alternative ways to find zeroes of monotone mappings and fixed points of nonexpansive operators. In the literature we find various papers that used the Bregman distances to generalize the theory from Hilbert to Banach space. See, for instance, [3, 4, 5, 6, 7, 15, 14, 8, 16, 17, 19, 25, 26, 27, 28, 29] and the references therein. The application of Bregman distances to the solution of this problem is also very useful in the case of finite dimensional spaces. Many papers considered the problem of minimizing functions in finite dimensional spaces using Bregman distances. See, for example, [9, 20, 21, 31].

Motivated by all these facts, our aim in this paper is to present a characterization of Bregman firmly nonexpansive operators using our new monotonicity concept. This characterization leads us to many examples of Bregman firmly nonexpansive operators in finite dimensional spaces.

Our paper is organized as follows. The next section (Section 2) is devoted to the preliminaries that are needed in our work. In the third section we prove the main results regarding the characterization of Bregman firmly nonexpansive operators. In this work we deal with two main notions: Bregman firmly nonexpansive operators and resolvents. Therefore we present examples of these notions with respect to different choices of Bregman functions. In the fourth section we present several examples of Bregman firmly nonexpansive operators and in the fifth section we present several examples of resolvents.

2 Preliminaries

Let X be a real reflexive Banach space with dual space X^*. The norms in X and X^* are denoted by $\|\cdot\|$ and $\|\cdot\|_*$, respectively. The pairing $\langle \xi, x\rangle$ is defined by the action of $\xi \in X^*$ at $x \in X$, that is, $\langle \xi, x\rangle = \xi(x)$. The set of all real numbers is denoted by $\mathbb{R}$ while $\mathbb{N}$ denotes the set of nonnegative integers. The closure of a subset C of X is denoted by $\overline{C}$. We refer to [10, 11] for notation and facts not proven within.

In the following three subsections (Sections 2.1-2.3) we give the definitions, notation and basic results that are needed in the sequel.

2.1 Admissible functions

Let $f : X \to (-\infty, +\infty]$ be a function. The *domain* of f is defined to be

$$\operatorname{dom} f := \{x \in X : f(x) < +\infty\}.$$

When $\operatorname{dom} f \neq \varnothing$ we say that f is *proper*. We denote by $\operatorname{int}\operatorname{dom} f$ the *interior* of the domain of f.

The *Fenchel conjugate* of f is the function $f^* : X^* \to (-\infty, +\infty]$ which is defined by

$$f^*(\xi) = \sup\{\langle \xi, x\rangle - f(x) : x \in X\}.$$

The function f is called *cofinite* if $\mathrm{dom}\ f^* = X^*$.

Let $x \in \operatorname{int}\operatorname{dom} f$. For any $y \in X$, we define the *right-hand derivative* of f at x by

$$f^\circ(x, y) := \lim_{t \to 0^+} \frac{f(x + ty) - f(x)}{t}. \tag{1}$$

If the limit as $t \to 0$ in (1) exists for any y, then the function f is said to be *Gâteaux differentiable at x* (see, for instance, [23, Definition 1.3, p. 3]). In this case, the *gradient* of f at x is the function $\nabla f(x)$ which is defined by $\langle \nabla f(x), y\rangle = f^\circ(x, y)$ for any $y \in X$. The function f is called *Gâteaux differentiable* if it is Gâteaux differentiable for any $x \in \operatorname{int}\operatorname{dom} f$. Throughout this paper, the function $f : X \to (0, +\infty]$ is a proper, convex and lower semicontinuous function which is also Gâteaux differentiable on $\operatorname{int}\operatorname{dom} f$. We will call such functions *admissible*.

The *Bregman distance* determined by a function f is the function $D_f : \mathrm{dom}\ f \times \operatorname{int}\operatorname{dom}\ f \to [0, +\infty]$ given by

$$D_f(y, x) := f(y) - f(x) - \langle \nabla f(x), y - x\rangle. \tag{2}$$

It is clear that the Bregman distance $D_f(y, x)$ with respect to the function $f = \|\cdot\|^2$ in Hilbert space is just $\|x - y\|^2$.

The function f is *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

$$\upsilon_f(x, t) := \inf\{D_f(y, x) : y \in \mathrm{dom}\ f,\ \ \|y - x\| = t\}, \tag{3}$$

is positive whenever $t > 0$. The function f is said to be *totally convex* when it is totally convex at any point $x \in \operatorname{int}\operatorname{dom} f$.

The following proposition summarizes known properties of the modulus of total convexity (*cf.* [15, Proposition 1.2.2, p. 18] and [14, Proposition 2.4, p. 26]).

Proposition 1 (Modulus properties)**.** *Let $f : X \to (0, +\infty]$ 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 $[0, \tau_f(x))$ 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.

Moreover, if $X=\mathbb{R}^n$ and $f:C\to\mathbb{R}$, where C is an open, convex and unbounded subset of $\mathbb{R}^n$, then the following statements also hold:

(v) $\upsilon_f(x,\cdot)$ is continuous from the right on $(0,+\infty)$;

(vi) If $\bar{f}:\overline{C}\to\mathbb{R}$ is a convex and continuous extension of f to $\overline{C}$ and if $\upsilon_f(x,\cdot)$ is continuous, then, for each $t\in[0,+\infty)$,

$$\upsilon_f(x,t)=\inf\left\{D_{\bar{f}}(y,x):y\in\overline{C},\ \ \|y-x\|=t\right\}.$$

Similarly, the function f is *uniformly convex* if the function $\delta_f:[0,+\infty)\to[0,+\infty]$, defined by

$$\delta_f(t):=\inf\left\{\frac{1}{2}f(x)+\frac{1}{2}f(y)-f\left(\frac{x+y}{2}\right):\|y-x\|=t,\ \ x,y\in\operatorname{dom}f\right\},\tag{4}$$

is positive whenever $t>0$. The function $\delta_f(\cdot)$ is called the *modulus of convexity* of f. For more details see, for instance, [11, 12].

According to [15, Proposition 1.2.5, p. 25], if $x\in\operatorname{int}\operatorname{dom}f$ and $t\in[0,+\infty)$, then $\upsilon_f(x,t)\geq\delta_f(t)$ and, therefore, if f is uniformly convex, then it is totally convex. The converse implication is not generally valid, that is, a function f may be totally convex without being uniformly convex (for such an example see [15, Section 1.3, p. 30]).

Definition 1 (Uniform smoothness)**.** The function f is called *uniformly smooth* if the function $\rho_f:[0,+\infty)\to\mathbb{R}$, defined by

$$\rho_f(t):=\sup\left\{\frac{1}{2}f(x)+\frac{1}{2}f(y)-f\left(\frac{x+y}{2}\right):\|y-x\|=t\right\},\tag{5}$$

satisfies $\lim_{t\to0^+}\rho_f(t)/t=0$.

The connection between uniform convexity and uniform smoothness is brought out by the following result (see [33, Theorem 3.5.5(i), p. 158]).

Proposition 2. *Let $f:X\to(-\infty,+\infty]$ be proper, lower semicontinuous and convex. Then f is uniformly convex if and only if f^* is uniformly smooth.*

The following result provides a characterization of uniform smoothness (see [33, Theorem 3.5.6(i)(xi), p. 159]).

Proposition 3. *Let $f : X \to (-\infty, +\infty]$ be proper, lower semicontinuous and convex. Then f is uniformly smooth if and only if* $\operatorname{dom} f = \mathbb{R}$*, f is Gâteaux differentiable and ∇f is uniformly continuous.*

2.2 Bregman operators

We fix a function f as above, and let K be a nonempty subset of $\operatorname{int}\operatorname{dom} f$. The *fixed point set* of an operator $T : K \to \operatorname{int}\operatorname{dom} f$ is the set $\{x \in K : Tx = x\}$ and is denoted by $\operatorname{Fix}(T)$.

We next list significant types of nonexpansivity with respect to the Bregman distance.

Definition 2 (Bregman nonexpansivity)**.** We say:

(i) the operator $T : K \to \operatorname{int}\operatorname{dom} f$ is *Bregman nonexpansive* (BNE) if

$$D_f(Tx, Ty) \leq D_f(x, y), \ \forall\, x, y \in K; \tag{6}$$

(ii) the operator $T : K \to \operatorname{int}\operatorname{dom} f$ is *quasi-Bregman nonexpansive* (QBNE) if

$$D_f(p, Tx) \leq D_f(p, x), \ \forall\, x \in K,\ p \in \operatorname{Fix}(T); \tag{7}$$

(iii) the operator $T : K \to \operatorname{int}\operatorname{dom} f$ is *Bregman firmly nonexpansive* (BFNE) if

$$\langle \nabla f(Tx) - \nabla f(Ty), Tx - Ty \rangle \leq \langle \nabla f(x) - \nabla f(y), Tx - Ty \rangle \tag{8}$$

for any $x, y \in K$, or equivalently,

$$\begin{aligned} D_f(Tx, Ty) + D_f(Ty, Tx) + D_f(Tx, x) + D_f(Ty, y) \\ \leq D_f(Tx, y) + D_f(Ty, x); \end{aligned} \tag{9}$$

(iv) the operator $T : K \to \operatorname{int}\operatorname{dom} f$ is *quasi-Bregman firmly nonexpansive* (QBFNE) if

$$0 \leq \langle \nabla f(x) - \nabla f(Tx), Tx - p \rangle \ \forall\, x \in K,\ p \in \operatorname{Fix}(T), \tag{10}$$

or equivalently,

$$D_f(p, Tx) + D_f(Tx, x) \leq D_f(p, x). \tag{11}$$

Assume now that $f = \|\cdot\|^2$ and the space X is a Hilbert space H, so that $\nabla f = I$ (the identity operator) and $D_f(y,x) = \|x-y\|^2$. Thence, Definition 2(i)-(iv) takes the following form.

(i') The operator $T : K \to H$ is *nonexpansive* (NE) if

$$\|Tx - Ty\| \le \|x - y\|\,,\ \forall\ x,y \in K; \tag{12}$$

(ii') the operator $T : K \to H$ is *quasi-nonexpansive* (QNE) if

$$\|Tx - p\| \le \|x - p\|\,,\ \forall\ x \in K,\ p \in \mathrm{Fix}\,(T)\,; \tag{13}$$

(iii') the operator $T : K \to H$ is *firmly nonexpansive* (FNE) if

$$\|Tx - Ty\|^2 \le \langle x - y, Tx - Ty\rangle\,,\ \forall\ x,y \in K; \tag{14}$$

(iv') the operator $T : K \to H$ is *quasi-firmly nonexpansive* (QFNE) if

$$\|Tx - p\|^2 + \|Tx - x\|^2 \le \|x - p\|^2\,,\ \forall\ x \in K,\ p \in \mathrm{Fix}\,(T)\,, \tag{15}$$

or equivalently, $0 \le \langle x - Tx, Tx - p\rangle$.

2.3 Resolvent mappings

Let $A : X \to 2^{X^*}$ be an arbitrary mapping (multi-function). Recall that the set $\mathrm{dom}\ A = \{x \in X : Ax \neq \varnothing\}$ is called the (effective) *domain* of the mapping A. We say that A is a *monotone* mapping if for any $x, y \in \mathrm{dom}\ A$ we have

$$\xi \in Ax \text{ and } \eta \in Ay \quad \Longrightarrow \quad 0 \le \langle \xi - \eta, x - y\rangle\,. \tag{16}$$

A monotone mapping A is said to be *maximal* if the graph of A is not a proper subset of the graph of any other monotone mapping.

Definition 3 (T-monotonicity)**.** Let $A : X \to 2^{X^*}$, $K \subset \mathrm{dom}\ A$ and $T : K \to X$. We say that the mapping A is monotone with respect to the operator T, or T*-monotone*, if

$$0 \le \langle \xi - \eta, Tx - Ty\rangle \tag{17}$$

for any $x, y \in K$, where $\xi \in Ax$ and $\eta \in Ay$.

Clearly, when $T = I$ the classes of monotone and I-monotone operators coincide.

Definition 4 (Set-valued indicator)**.** The set-valued indicator of a subset S of X is defined by

$$\mathbb{I}_S : x \mapsto \begin{cases} \{0\}, & x \in S; \\ \emptyset, & \text{otherwise.} \end{cases}$$

The concept of T-monotonicity can also be defined by using this set-valued indicator (as kindly suggested by Heinz H. Bauschke [2]).

Remark 1. A mapping A is T-monotone if and only if $T \circ \left(A^{-1} + \mathbb{I}_{A(K)}\right)$ is monotone. $\Diamond$

Remark 2. An unrelated concept of a T-monotone operator can be found in several papers of Calvert (see, for example, [18]). $\Diamond$

Remark 3. Let $F : X \to X$ be an operator which satisfies

$$0 \leq \langle Jx - Jy, Fx - Fy \rangle \tag{18}$$

for any $x, y \in \operatorname{dom} f$, where J is the normalized duality mapping of the space X. An operator F which satisfies inequality (18) is called *d-accretive* (see [1]). Clearly in our terms J is F-monotone whenever F is d-accretive. $\Diamond$

Definition 5 (f-Resolvents)**.** Let $A : X \to 2^{X^*}$ be a monotone mapping. The *resolvent* of A with respect to f is the operator $\operatorname{Res}_A^f : X \to 2^X$ which is given by

$$\operatorname{Res}_A^f := (\nabla f + A)^{-1} \circ \nabla f. \tag{19}$$

Remark 4. In the case of a Hilbert space and when $f = (1/2) \|\cdot\|^2$, the resolvent Res_A^f is the *classical resolvent*

$$\mathrm{R}_A := (I + A)^{-1},$$

and when A is maximal, the Minty surjectivity theorem [11, p. 435] assures us that R_A will be everywhere defined. $\Diamond$

Finally, we record various basic properties of f-resolvents as established in [5, Proposition 3.8, p. 604].

Proposition 4 (Properties of f-resolvents)**.** *Let $f : X \to (-\infty, +\infty]$ be an admissible function and let $A : X \to 2^{X^*}$ be a mapping such that* $\operatorname{int} \operatorname{dom} f \cap \operatorname{dom} A \neq \varnothing$. *The following statements hold:*

(i) dom $\operatorname{Res}_A^f \subset \operatorname{int} \operatorname{dom} f$*;*

(ii) ran $\operatorname{Res}_A^f \subset \operatorname{int} \operatorname{dom} f$*;*

(iii) Fix $\left(\operatorname{Res}_A^f\right) = \operatorname{int} \operatorname{dom} f \cap A^{-1}(0^*)$.

(iv) Suppose additionally that A is a monotone mapping. Then also the following assertions hold:

(a) The operator Res_A^f is BFNE.

(b) If, in addition, $f|_{\operatorname{int} \operatorname{dom} f}$ is strictly convex, then the operator Res_A^f is single-valued on its domain.

(c) If $f : X \to \mathbb{R}$ is such that $\operatorname{ran} \nabla f \subset \operatorname{ran} (\nabla f + A)$, *then* $\operatorname{dom} \operatorname{Res}_A^f = X$.

Thus Proposition 4(iv)(a) recaptures Rockafellar's result [30] that the classical resolvent in Hilbert space is firmly nonexpansive.

The following result gives us a sufficient condition for the f-resolvent to have full domain.

Proposition 5. *[8, Corollary 2.3, p. 59] Assume that $A : X \to 2^{X^*}$ is a monotone mapping and that $f : X \to \mathbb{R}$ is a Gâteaux differentiable, strictly convex and cofinite function. Then A is maximal monotone if and only if* $\operatorname{ran} (A + \nabla f) = X^*$.

Combining Propositions 4(iv)(c) and 5, we obtain the following result.

Remark 5. If $A : X \to 2^{X^*}$ is a maximal monotone mapping and $f : X \to \mathbb{R}$ is a Gâteaux differentiable, strictly convex and cofinite function, then $\operatorname{dom} \operatorname{Res}_A^f = X$. $\Diamond$

3 Characterization of BFNE operators

In this section we establish a characterization of BFNE operators. This characterization emphasizes the strong connection between the nonexpansivity of T and the monotonicity of S_T, where

$$S_T := \nabla f - (\nabla f) \circ T. \tag{20}$$

Results in this direction have been known for a long time. We cite two of them.

Proposition 6 (Rockafellar, 1976)**.** *Let X be a Hilbert space. Then T is firmly nonexpansive if and only if $I - T$ is T-monotone.*

Proposition 7 (Bauschke, Wang and Yao, 2008)**.** *Let X be a reflexive Banach space. Let K be a subset of X and let $T : K \to X$. Fix an admissible function $f : X \to \mathbb{R}$ and set*

$$A_T := \nabla f \circ T^{-1} - \nabla f.$$

If T is BFNE, then A_T is monotone (this operator is not necessarily single-valued).

Motivated by these results, we offer the following generalization.

Theorem 1 (Characterization of BFNE operators)**.** *Let $K \subset \operatorname{int}\operatorname{dom} f$ and suppose that $T : K \to \operatorname{int}\operatorname{dom} f$ for an admissible function f. Then T is BFNE if and only if $S_T = \nabla f - (\nabla f) \circ T$ is T-monotone.*

Proof. Suppose that T is BFNE. Take x, y in K and denote $\xi = S_T(x)$ and $\eta = S_T(y)$.

Then by the definition of S_T (see (20)) we obtain

$$\nabla f\left(T(x)\right) = \nabla f(x) - \xi \quad \text{and} \quad \nabla f\left(T(y)\right) = \nabla f(y) - \eta. \tag{21}$$

Since T is BFNE, we have

$$\left\langle \nabla f\left(Tx\right) - \nabla f\left(Ty\right), Tx - Ty\right\rangle \leq \left\langle \nabla f\left(x\right) - \nabla f\left(y\right), Tx - Ty\right\rangle. \tag{22}$$

Now, substituting (21) on the left-hand side of (22), we obtain

$$\left\langle \left(\nabla f(x) - \xi\right) - \left(\nabla f(y) - \eta\right), Tx - Ty\right\rangle \leq \left\langle \nabla f\left(x\right) - \nabla f\left(y\right), Tx - Ty\right\rangle,$$

which means that

$$0 \leq \left\langle S_T\left(x\right) - S_T\left(y\right), Tx - Ty\right\rangle.$$

Thus S_T is T-monotone. Conversely, if S_T is T-monotone, then

$$0 \leq \left\langle S_T\left(x\right) - S_T\left(y\right), Tx - Ty\right\rangle$$

for any $x, y \in K$ and therefore

$$0 \leq \left\langle \left(\nabla f(x) - \nabla f(Tx)\right) - \left(\nabla f(y) - \nabla f(Ty)\right), Tx - Ty\right\rangle,$$

which means that

$$\left\langle \nabla f\left(Tx\right) - \nabla f\left(Ty\right), Tx - Ty\right\rangle \leq \left\langle \nabla f\left(x\right) - \nabla f\left(y\right), Tx - Ty\right\rangle$$

for any $x, y \in K$. In other words, T is indeed a BFNE operator. □

Remark 6. It is clear that when X is a Hilbert space and $f = \|\cdot\|^2$, then BFNE operators are firmly nonexpansive operators and $S_T = I - T$. Therefore Proposition 6 is an immediate consequence of Theorem 1. $\Diamond$

Remark 7. If T is a BFNE operator, then S_T is T-monotone by Theorem 1. Take $\xi \in A_T(x)$ and $\eta \in A_T(y)$. From the definition of A_T we get $\xi = \nabla f(z) - \nabla f(x)$, where $Tz = x$, and $\eta = \nabla f(w) - \nabla f(y)$, where $Tw = y$. Hence

$$\begin{aligned} \langle \xi - \eta, x - y \rangle &= \langle (\nabla f(z) - \nabla f(Tz)) - (\nabla f(w) - \nabla f(Tw)), Tz - Tw \rangle \\ &= \langle S_T(z) - S_T(w), Tz - Tw \rangle \\ &\geq 0 \end{aligned}$$

for all $x, y \in \mathrm{dom} A_T$, and so A_T is monotone. Hence Proposition 7 follows from Theorem 1. $\Diamond$

Motivated by our characterization (Theorem 1), we now show that the converse implication of Proposition 7 is also true.

Proposition 8. *Let $K \subset \operatorname{int}\operatorname{dom} f$ and suppose that $T : K \to \operatorname{int}\operatorname{dom} f$ for an admissible function f. If A_T is monotone, then T is BFNE.*

Proof. Suppose that A_T is monotone. Then for any $x, y \in \mathrm{dom}\ A_T$, we have

$$0 \leq \langle \xi - \eta, x - y \rangle$$

for any $\xi \in A_T(x)$ and $\eta \in A_T(y)$. Let $w, z \in K$. Set $\xi = \nabla f(z) - \nabla f(x)$, where $Tz = x$, and $\eta = \nabla f(w) - \nabla f(y)$, where $Tw = y$. We have

$$0 \leq \langle (\nabla f(z) - \nabla f(x)) - (\nabla f(w) - \nabla f(y)), x - y \rangle,$$

which means that

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle \leq \langle \nabla f(z) - \nabla f(w), x - y \rangle.$$

Thus

$$\langle \nabla f(Tz) - \nabla f(Tw), Tz - Tw \rangle \leq \langle \nabla f(z) - \nabla f(w), Tz - Tw \rangle$$

and so T is a BFNE operator, as asserted. $\square$

Remark 8. Combining Propositions 7 and 8, we obtain another characterization of BFNE operators. $\Diamond$

Remark 9. Our characterization of BFNE operators is based on a new type of monotonicity which seems to be harder to check than the regular one. On the other hand, our mapping S_T is defined without any inverse operation, and hence is easier to compute (see Subsections 4.1.1 and 4.1.2). In the case of the mapping A_T, similar computations seem to be much harder because of the presence of the inverse operator T^{-1}. $\Diamond$

We underscore the importance of the new operator S_T. This operator is a generalization of the *Yosida approximation* operator, because, when in Hilbert space we take the function f to be $(1/2)\left\|\cdot\right\|^2$ and $T = R_A = (I + A)^{-1}$, where A is a monotone mapping, then $S_T = I - T$ is exactly the classical Yosida approximation operator with $\lambda = 1$. In the following result we present two properties of S_T when the operator T is taken to be the resolvent Res_A^f of a monotone mapping A. To facilitate matters, for any $\lambda > 0$ we shall denote

$$T_{\lambda A}^f := \operatorname{Res}_{\lambda A}^f. \tag{23}$$

We may now prove the following result.

Theorem 2. *Let $A : X \to 2^{X^*}$ be a monotone mapping and let $f : X \to (0, +\infty]$ be an admissible strictly convex function. Then with $T_{\lambda A}^f$ given by (23) we have the following implications:*

(i) $\left(T_{\lambda A}^f(x), \lambda^{-1} S_{T_{\lambda A}^f}(x)\right) \in \operatorname{graph} A$*; and*

(ii) $0 \in Ax$ *if and only if* $0 \in S_{T_{\lambda A}^f}(x)$.

Proof. (i) Indeed,

$$\begin{aligned} T_{\lambda A}^f(x) &= (\nabla f + \lambda A)^{-1} \circ \nabla f(x) \Leftrightarrow \nabla f(x) \in (\nabla f + \lambda A) \circ T_{\lambda A}^f(x) \\ &\Leftrightarrow \lambda^{-1}\left(\nabla f - \nabla f \circ T_{\lambda A}^f\right)(x) \in A\left(T_{\lambda A}^f(x)\right) \\ &\Leftrightarrow \lambda^{-1} S_{T_{\lambda A}^f}(x) \in A\left(T_{\lambda A}^f(x)\right). \end{aligned}$$

(ii) Likewise, since ∇f is injective (see [4, Theorem 5.10, p. 636]) it follows that

$$\begin{aligned} 0 \in Ax &\Leftrightarrow 0 \in \lambda Ax \Leftrightarrow \nabla f(x) \in (\nabla f + \lambda A)(x) \Leftrightarrow x \in (\nabla f + \lambda A)^{-1} \circ \nabla f(x) \\ &\Leftrightarrow \nabla f(x) \in \nabla f\left(T_{\lambda A}^f(x)\right) \Leftrightarrow 0 \in \left(\nabla f - \nabla f \circ T_{\lambda A}^f\right)(x) \Leftrightarrow 0 \in S_{T_{\lambda A}^f}(x), \end{aligned}$$

as required. $\square$

4 Examples of BFNE operators in Euclidean spaces

In this section we shall use Theorem 1 to present various examples of BFNE operators in Euclidean spaces. Indeed, we have already seen that BFNE operators can be generated from T-monotone mappings. Moreover, the notion of T-monotonicity can be simplified in the case of the real line.

Remark 10. If $X = \mathbb{R}$ and both T and S_T are increasing (decreasing), then S_T is T-monotone. $\Diamond$

We begin this section with critical one-dimensional examples for various choices of the function f.

4.1 The case of the real line

Assume that $X = \mathbb{R}$. In this section we study in detail the classes of BFNE operators with respect to the *Boltzmann-Shannon entropy*

$$\mathcal{BS}\left(x\right) := x\log\left(x\right) - x, \;\; 0 < x < +\infty \tag{24}$$

and the *Fermi-Dirac entropy*

$$\mathcal{FD}\left(x\right) := x\log\left(x\right) + \left(1-x\right)\log\left(1-x\right), \;\; 0 < x < 1. \tag{25}$$

Each of these functions can be defined to be zero, by its limits, at the endpoints of their domains. At the end of this section we present a table of sufficient conditions for an operator to be BFNE with respect to various other admissible functions.

We study the entropies $\mathcal{BS}$ and $\mathcal{FD}$ in detail because of their importance in applications. These two functions, which form a large part of the basis for classical information theory, arguably provide the only consistent measures of the average uncertainty in predicting outcomes of a random experiment (see [22]).

Moreover, both $D_{\mathcal{BS}}$ and $D_{\mathcal{FD}}$ are jointly convex [10, 11], an uncommon property which they share with $(x,y) \mapsto \|x-y\|^2$. The utility of both the Boltzmann-Shannon and the Fermi-Dirac entropies is enhanced because they are totally convex. In the following two results (Propositions 9 and 10) we calculate the modulus of total convexity of the $\mathcal{BS}$ entropy and show that $\mathcal{BS}$ is totally convex (*cf.* [14, 15]). See Propositions 11 and 12 for analogous results concerning the $\mathcal{FD}$ entropy.

Proposition 9 (Modulus of total convexity of $\mathcal{BS}$)**.** *The modulus of total convexity of $\mathcal{BS}$ on $(0,+\infty)$ is given by*

$$\upsilon_{\mathcal{BS}}\left(x,t\right) = x\left[\left(1+\frac{t}{x}\right)\log\left(1+\frac{t}{x}\right) - \frac{t}{x}\right], \; x \in (0,+\infty), \; t \geq 0. \tag{26}$$

Proof. Let $x_0 \in (0, +\infty)$ and $0 < t < x_0$. It is clear from the definition of the modulus of total convexity that

$$\begin{aligned}\upsilon_{\mathcal{BS}}(x_0, t) &= \min\{D_{\mathcal{BS}}(x_0 + t, x_0), D_{\mathcal{BS}}(x_0 - t, x_0)\}\\ &= \min\left\{(x_0 + t)\log\left(\frac{x_0 + t}{x_0}\right) - t, (x_0 - t)\log\left(\frac{x_0 - t}{x_0}\right) + t\right\}\\ &= \min\left\{x_0\left[\left(1 + \frac{t}{x_0}\right)\log\left(1 + \frac{t}{x_0}\right) - \frac{t}{x_0}\right],\right.\\ &\left. x_0\left[\left(1 - \frac{t}{x_0}\right)\log\left(1 - \frac{t}{x_0}\right) + \frac{t}{x_0}\right]\right\}.\end{aligned}$$

In order to find this minimum we define a function $\varphi : [0, x_0) \to \mathbb{R}$ by

$$\varphi(t) := x_0\left[\left(1 - \frac{t}{x_0}\right)\log\left(1 - \frac{t}{x_0}\right) + \frac{t}{x_0} - \left(1 + \frac{t}{x_0}\right)\log\left(1 + \frac{t}{x_0}\right) + \frac{t}{x_0}\right].$$

It is clear that $\varphi(0) = 0$ and

$$\varphi'(t) = -\log\left(1 - \left(\frac{t}{x_0}\right)^2\right),$$

and so φ is increasing for all $t < x_0$. Thus $\varphi(t) > 0$ for any $t < x_0$, which means that

$$\upsilon_{\mathcal{BS}}(x_0, t) = x_0\left[\left(1 + \frac{t}{x_0}\right)\log\left(1 + \frac{t}{x_0}\right) - \frac{t}{x_0}\right]$$

for any $t < x_0$. If $t \geq x_0$ then the point $t - x_0$ does not belong to the domain of $\mathcal{BS}$ and therefore

$$\upsilon_{\mathcal{BS}}(x_0, t) = D_{\mathcal{BS}}(x_0 + t, x_0) = x_0\left[\left(1 + \frac{t}{x_0}\right)\log\left(1 + \frac{t}{x_0}\right) - \frac{t}{x_0}\right].$$

Hence (26) holds for any $t \geq 0$. $\square$

Now we show that $\mathcal{BS}$ is totally convex but not uniformly convex.

Proposition 10 (Total convexity of $\mathcal{BS}$)**.** *The function $\mathcal{BS}$ is totally convex but not uniformly convex.*

Proof. We need to show that $\upsilon_{\mathcal{BS}}(x_0, t) > 0$ for any $t > 0$. We know that $\upsilon_{\mathcal{BS}}(x_0, 0) = 0$ and from Proposition 9 we obtain that

$$\frac{\partial}{\partial t}\left(\upsilon_{\mathcal{BS}}(x_0, t)\right) = \log\left(1 + \frac{t}{x_0}\right) > 0,\ t > 0.$$

This means that $\upsilon_{\mathcal{BS}}(x_0, t)$ is a strictly increasing function for all $t > 0$. Thence, $\upsilon_{\mathcal{BS}}(x_0, t) > 0$ for any $t > 0$ and so $\mathcal{BS}$ is totally convex on $(0, +\infty)$, as asserted. Since for any $t > 0$, we have

$$0 \le \delta_{\mathcal{BS}}(t) \le \lim_{x \to +\infty} \upsilon_{\mathcal{BS}}(x, t) = 0.$$

It follows that $\delta_{\mathcal{BS}}(t) = 0$ and thus $\mathcal{BS}$ is not uniformly convex. □

Remark 11. In [14] it is mentioned that the modulus of total convexity of $f(x) = x\log(x)$ is also given by (26) and that f is totally convex. Note that $D_f = D_{\mathcal{BS}}$. ♢

The following results show that $\mathcal{FD}$ is both totally convex and uniformly convex.

Proposition 11 (Modulus of total convexity of $\mathcal{FD}$)**.** *The modulus of total convexity of $\mathcal{FD}$ on $(0,1)$ is given by*

$$\upsilon_{\mathcal{FD}}(x, t) = x\left[\left(1 + \frac{t}{x}\right)\log\left(1 + \frac{t}{x}\right) + \left(\frac{1-t}{x} - 1\right)\log\left(1 - \frac{t}{1-x}\right)\right], \tag{27}$$

when $0 < x \le 1/2$ and $0 < t < 1 - x$, and by

$$\upsilon_{\mathcal{FD}}(x, t) = x\left[\left(1 - \frac{t}{x}\right)\log\left(1 - \frac{t}{x}\right) + \left(\frac{1+t}{x} - 1\right)\log\left(1 + \frac{t}{1-x}\right)\right] \tag{28}$$

when $1/2 \le x < 1$ and $0 < t < x$.

Proof. Let $x_0 \in (0,1)$. Denote $M = \max\{x_0, 1 - x_0\}$ and $m = \min\{x_0, 1 - x_0\}$. If $0 < t < m$, then it is clear from the definition of the modulus of total convexity that

$$\begin{aligned}
\upsilon_{\mathcal{FD}}(x_0, t) &= \min\left\{D_{\mathcal{FD}}(x_0 + t, x_0), D_{\mathcal{FD}}(x_0 - t, x_0)\right\} \\
&= \min\left\{x_0\left[\left(1 + \frac{t}{x_0}\right)\log\left(1 + \frac{t}{x_0}\right) + \left(\frac{1-t}{x_0} - 1\right)\log\left(1 - \frac{t}{1-x_0}\right)\right],\right. \\
&\left. x_0\left[\left(1 - \frac{t}{x_0}\right)\log\left(1 - \frac{t}{x_0}\right) + \left(\frac{1+t}{x_0} - 1\right)\log\left(1 + \frac{t}{1-x_0}\right)\right]\right\}.
\end{aligned}$$

In order to find this minimum we define a function $\psi : [0, m) \to \mathbb{R}$ by

$$\psi(t) := x_0\left[\left(1-\frac{t}{x_0}\right)\log\left(1-\frac{t}{x_0}\right)+\left(\frac{1+t}{x_0}-1\right)\log\left(1+\frac{t}{1-x_0}\right)\right] \\ -x_0\left[\left(1+\frac{t}{x_0}\right)\log\left(1+\frac{t}{x_0}\right)+\left(\frac{1-t}{x_0}-1\right)\log\left(1-\frac{t}{1-x_0}\right)\right].$$

It is clear that $\psi(0) = 0$ and

$$\psi'(t) = \log\left(1-\left(\frac{t}{1-x_0}\right)^2\right)-\log\left(1-\left(\frac{t}{x_0}\right)^2\right),$$

Therefore, for any $0 < t < m$, the function ψ is increasing when $0 < x \leq 1/2$ and decreasing when $1/2 \leq x < 1$. Hence, for any $0 < t < m$, the function $\psi(t) > 0$ when $0 < x \leq 1/2$ and $\psi(t) < 0$ when $1/2 \leq x < 1$. If $m \leq t < M$, then one of the points $x_0 - t$ or $x_0 + t$ belongs to the domain of $\mathcal{FD}$ and the second does not. Therefore the modulus of total convexity of $\mathcal{FD}$ is given by (27) and (28) in all cases. □

Proposition 12 (Total convexity of $\mathcal{FD}$)**.** *The function $\mathcal{FD}$ is totally convex.*

Proof. We need to show that $\upsilon_{\mathcal{FD}}(x_0, t) > 0$ for any $t > 0$. We know that $\upsilon_{\mathcal{FD}}(x_0, 0) = 0$ and from Proposition 11 we obtain that

$$\frac{\partial}{\partial t}\left(\upsilon_{\mathcal{FD}}(x_0,t)\right)=\begin{cases}\log\left(1+\frac{t}{x(1-x-t)}\right), & 0<x\leq 1/2,\quad 0<t<1-x,\\ \log\left(1+\frac{t}{(1-x)(x-t)}\right), & 1/2\leq x<1,\quad\ 0<t<x.\end{cases}$$

This means that $\upsilon_{\mathcal{FD}}(x_0, t)$ is a strictly increasing function for all $t > 0$. Thence, $\upsilon_{\mathcal{FD}}(x_0, t) > 0$ for any $t > 0$ and so $\mathcal{FD}$ is totally convex on $(0, 1)$, as asserted. □

Lemma 1. Let $f : (a, b) \to \mathbb{R}$ be twice differentiable. If $f''(x) \geq m > 0$ on (a, b), then f is uniformly convex there.

Proof. Let $x, y \in (a, b)$ with $\|y - x\| = t > 0$. Then

$$f(x)=f\left(\frac{x+y}{2}\right)+f'\left(\frac{x+y}{2}\right)\left(\frac{x-y}{2}\right)+\frac{f''(\xi)}{2}\left(\frac{x-y}{2}\right)^2$$

and

$$f(y)=f\left(\frac{x+y}{2}\right)+f'\left(\frac{x+y}{2}\right)\left(\frac{y-x}{2}\right)+\frac{f''(\eta)}{2}\left(\frac{y-x}{2}\right)^2$$

for some $\xi, \eta \in (a, b)$. Therefore

$$\frac{f(x)}{2} + \frac{f(y)}{2} - f\left(\frac{x+y}{2}\right) = \frac{f''(\xi)}{4}\left(\frac{x-y}{2}\right)^2 + \frac{f''(\eta)}{4}\left(\frac{y-x}{2}\right)^2 \geq \frac{mt^2}{8} > 0,$$

as asserted. □

Proposition 13 (Uniform convexity of $\mathcal{FD}$)**.** *The function $\mathcal{FD}$ is uniformly convex.*

Proof. This result follows immediately from Lemma 1 because $\mathcal{FD}''(x) \geq 4$ for all $x \in (0, 1)$. □

The following remark provides another proof of Proposition 13. It was kindly suggested by Heinz H. Bauschke [2].

Remark 12. Define $f : \mathbb{R} \to (-\infty, +\infty]$ by

$$f(x) = \begin{cases} 0 = \lim_{x \to 0^+} \mathcal{FD}(x), & \text{if } x = 0 \\ \mathcal{FD}(x), & \text{if } x \in (0,1) \\ 0 = \lim_{x \to 1^-} \mathcal{FD}(x), & \text{if } x = 0 \\ +\infty, & \text{otherwise.} \end{cases} \tag{29}$$

Then f is proper, lower semicontinuous and convex. In addition, it follows from [10, p. 50] that $f^*(x) = \ln(e^x + 1)$,

$$(f^*)'(x) = \frac{e^x}{e^x + 1} \quad \text{and} \quad (f^*)''(x) = \frac{e^x}{(e^x + 1)^2}.$$

Thus $(f^*)'$ is a 1-Lipschitz function and hence is uniformly continuous. From Proposition 3 we obtain that f^* is uniformly smooth. Now it follows from Proposition 2 that f is uniformly convex and so is $\mathcal{FD}$. ♢

The following result (kindly communicated to us by Liangjin Yao [32]) provides still another way to prove Proposition 13.

Proposition 14. *Let $f : \mathbb{R} \to (-\infty, +\infty]$ be a proper and convex function with bounded domain. Assume that for every open interval $(a, b) \subseteq \operatorname{dom} f$, the restriction of f to (a, b) is not an affine function. Then f is uniformly convex.*

A fairly detailed discussion of the relations between total and uniform convexity and of various other properties is to be found in [11]. We caution that these terms do not have uniform definitions throughout the literature. The next remark allows us to explicitly produce BFNE operators.

Remark 13. Let K be a nonempty subset of $(0, +\infty)$. By Theorem 1 we know that an increasing operator T is BFNE if S_T is also increasing. If T is also differentiable on $\operatorname{int} K$, then $S_T' = f'' - f''(T)\, T'$.

We conclude that T is BFNE on K with respect to an admissible twice-differentiable function f as soon as

$$0 \leq T'(x) \leq \frac{f''(x)}{f''(T(x))} \tag{30}$$

for all $x \in \operatorname{int} K$. $\diamondsuit$

4.1.1 The Boltzmann-Shannon entropy

We return to the *Boltzmann-Shannon entropy* (BS entropy), which we recall from (24) is the function $\mathcal{BS} : (0, +\infty) \to \mathbb{R}$ defined by

$$\mathcal{BS}(x) := x \log(x) - x.$$

It is clear that $\mathcal{BS}$ is differentiable on $(0, +\infty)$ and that $\mathcal{BS}'(x) = \log(x)$. We also

Figure 1: The Boltzmann-Shannon entropy

have

$$D_{\mathcal{BS}}(y, x) = y \log\left(\frac{y}{x}\right) - y + x$$

for any $x, y \in (0, +\infty)$. The $\mathcal{BS}$ distance is also well defined for $y = 0$ (see [10, 11]), but this does not concern us here.

The following result gives sufficient conditions for an operator T to be BFNE with respect to $\mathcal{BS}$. We shall write that T is $\mathcal{BS}$-BFNE.

Proposition 15 (Conditions for $\mathcal{BS}$-BFNE)**.** *Let K be a nonempty subset of $(0, +\infty)$ and let $T : K \to K$ be an operator. Assume that one of the following conditions holds:*

(i) T is increasing and $T(x)/x$ is decreasing for every $x \in \operatorname{int} K$;

(ii) T is differentiable on $\operatorname{int} K$ and its derivative T' satisfies

$$0 \leq T'(x) \leq \frac{T(x)}{x} \tag{31}$$

for every $x \in \operatorname{int} K$;

(iii) T *is decreasing and* $T(x)/x$ *is increasing for every* $x \in \mathrm{int}K$*;*

(iv) T *is differentiable on* $\mathrm{int}\, K$ *and its derivative* T' *satisfies*

$$\frac{T(x)}{x} \leq T'(x) \leq 0$$

for every $x \in \mathrm{int}K$*.*

Then T *is a* $\mathcal{BS}$*-BFNE operator on* K*.*

Proof. This result follows immediately from Theorem 1 and Remark 13. □

Remark 14. The only solution of the differential equation

$$T'(x) = \frac{T(x)}{x}$$

is

$$T(x) = \alpha x$$

for any $\alpha \in \mathbb{R}$, but in our case $\alpha \in (0, +\infty)$ since $T(x) \in (0, +\infty)$ for any $x \in (0, +\infty)$. ◊

Using the conditions provided in Proposition 15, we can give more examples of $\mathcal{BS}$-BFNE operators.

Example 1 (Examples of $\mathcal{BS}$-BFNE operators)**.** We provide:

(i) $T(x) = \alpha x + \beta$, $\alpha, \beta \in (0, +\infty)$, is a $\mathcal{BS}$-BFNE operator on any nonempty subset of $(0, +\infty)$. Therefore, if $\alpha = 0$ then $T(x) = \beta$ and if $\beta = 0$ then $T(x) = \alpha x$ are $\mathcal{BS}$-BFNE operators on any subset of $(0, +\infty)$.

Proof. Since $T'(x) = \alpha$ and

$$\frac{T(x)}{x} = \alpha + \frac{\beta}{x}$$

it is easy to see that $0 \leq T'(x) \leq T(x)/x$ if and only if $\alpha, \beta \in (0, +\infty)$. Therefore T is a $\mathcal{BS}$-BFNE operator for any $\alpha, \beta \in (0, +\infty)$. □

(ii) $T(x) = x^p$, $p \in (0, 1]$, is a $\mathcal{BS}$-BFNE operator on any subset of $(0, +\infty)$.

Proof. Since $T'(x) = px^{p-1}$ and $T(x)/x = x^{p-1}$, it is easy to see that $0 \leq T'(x) \leq T(x)/x$ if and only if $p \in (0,1]$. Therefore T is a $\mathcal{BS}$-BFNE operator for any $p \in (0,1]$. $\square$

(iii) $T(x) = \alpha x - x^p$, $p \in [1,+\infty)$ and $\alpha \in \mathbb{R}_{++}$, is a $\mathcal{BS}$-BFNE operator on any subset of $\left(0, \alpha^{1/(p-1)}\right)$.

Proof. The operator T is well-defined when $T(x) > 0$, which in this case happens when $\alpha \in \mathbb{R}_{++}$ and $x \in \left(0, \alpha^{1/(p-1)}\right)$. Since $T'(x) = \alpha - px^{p-1}$ and $T(x)/x = \alpha - x^{p-1}$, it is easy to see that $0 \leq T'(x) \leq T(x)/x$ if and only if $p \in [1,+\infty)$ and $\alpha \in \mathbb{R}_{++}$. Therefore T is a $\mathcal{BS}$-BFNE operator for any $p \in [1,+\infty)$ and $\alpha \in \mathbb{R}_{++}$. $\square$

(iv) $T(x) = \log(x)$ is a $\mathcal{BS}$-BFNE operator on any subset of $[e,+\infty)$.

Proof. Since $T'(x) = 1/x$ and $T(x)/x = \log(x)/x$, it is easy to see that $0 < T'(x) \leq T(x)/x$ if and only if $1 \leq \log(x)$ and this happens if and only if $x \in [e,+\infty)$. $\square$

(v) $T(x) = e^x$ is a $\mathcal{BS}$-BFNE operator on any subset of $(0,1]$.

Proof. Since $T'(x) = e^x$ and $T(x)/x = e^x/x$, it is easy to see that $0 < T'(x) \leq T(x)/x$ if and only if $1 \leq 1/x$ and this happens if and only if $x \in (0,1]$. $\square$

(vi) $T(x) = \sin(x)$ is a $\mathcal{BS}$-BFNE operator on any subset of $(0,\pi/2]$ (there are more such intervals).

Proof. Since $T'(x) = \cos(x)$ and $T(x)/x = \sin(x)/x$, it is easy to see that $0 < T'(x) \leq T(x)/x$ if and only if $x \leq \tan(x)$ and this happens if $x \in (0,\pi/2]$. $\square$

Needless to say, it is possible to produce many more such examples. $\Diamond$

Figure 2: The Fermi-Dirac entropy

4.1.2 The Fermi-Dirac entropy

Recall from (25) that the *Fermi-Dirac entropy* is the function $\mathcal{FD} : [0,1] \to \mathbb{R}$ given by

$$\mathcal{FD}(x) := x \log(x) + (1-x)\log(1-x).$$

It is clear that $\mathcal{FD}$ is differentiable on $(0,1)$ and $\mathcal{FD}'(x) = \log(x/(1-x))$. We also have

$$D_{\mathcal{FD}}(y,x) = y \log\left(\frac{y}{x}\right) + (1-y)\log\left(\frac{1-y}{1-x}\right)$$

for any $x, y \in (0,1)$. Once again, this function is also well defined for $y = 0, 1$ (see [10, 11]).

In this subsection we study the BFNE operators with respect to $\mathcal{FD}$. We shall denote them by $\mathcal{FD}$-BFNE.

Remark 15. Let K be a nonempty subset of $(0,1)$. From Theorem 1 we know that an increasing operator T is $\mathcal{FD}$-BFNE if S_T is increasing. If T is also differentiable on $\operatorname{int} K$, then $S_T' = \mathcal{FD}'' - \mathcal{FD}''(T)\,T'$.

Therefore T is $\mathcal{FD}$-BFNE on K if the following condition holds:

$$0 \leq T'(x) \leq \frac{\mathcal{FD}''(x)}{\mathcal{FD}''(T(x))}. \tag{32}$$

for all $x \in \operatorname{int} K$. ♢

The following result gives sufficient conditions for the operator T to be $\mathcal{FD}$-BFNE.

Proposition 16 (Conditions for $\mathcal{FD}$-BFNE)**.** *Let K be a nonempty subset of $(0,1)$ and let $T : K \to K$ be an operator. Assume that one of the following conditions holds:*

(i) T is increasing and

$$\frac{T(x)(1-x)}{x(1-T(x))}$$

is decreasing for every $x \in \operatorname{int} K$;

(ii) T *is differentiable and its derivative* T' *satisfies*

$$0 \leq T'(x) \leq \frac{T(x)(1-T(x))}{x(1-x)}$$

for every $x \in \mathrm{int}K$*;*

(iii) T *is decreasing and*

$$\frac{T(x)(1-x)}{x(1-T(x))}$$

is increasing for every $x \in \mathrm{int}K$*;*

(iv) T *is differentiable on* $\mathrm{int}\,K$ *and its derivative* T' *satisfies*

$$\frac{T(x)(1-T(x))}{x(1-x)} \leq T'(x) \leq 0$$

for every $x \in \mathrm{int}K$*.*

Then T *is an* $\mathcal{FD}$*-BFNE operator on* K*.*

Proof. This result follows immediately from Theorem 1 and Remark 15. □

Remark 16. The only solution of the differential equation

$$T'(x) = \frac{T(x)(1-T(x))}{x(1-x)}$$

is

$$T(x) = \frac{\alpha x}{(1-x+\alpha x)} \tag{33}$$

for any $\alpha \in \mathbb{R}$, but in our case $\alpha \in (0,+\infty)$ since $T(x) \in (0,1)$ for any $x \in (0,1)$. ♢

Using Proposition 16 we can give examples of $\mathcal{FD}$-BFNE operators.

Example 2 (Examples of $\mathcal{FD}$-BFNE operators)**.** We provide:

(i) $T(x) = \alpha$, $\alpha \in (0,1)$, is an $\mathcal{FD}$-BFNE operator on any subset of $(0,1)$.

(ii) $T(x) = \alpha x$, $\alpha \in (0,1)$, is an $\mathcal{FD}$-BFNE operator on any subset of $(0,1)$.

(iii) $T(x) = x^p$, $p \in (0,1)$, is an $\mathcal{FD}$-BFNE operator on any subset of $(0,+\infty)$.

Proof. Since $T'(x) = px^{p-1}$ and

$$\frac{T(x)(1-T(x))}{x(1-x)} = \frac{x^{p-1}(1-x^p)}{(1-x)},$$

it is easy to see that

$$0 \leq T'(x) \leq \frac{T(x)(1-T(x))}{x(1-x)}$$

if and only if $p \in (0,1)$. Thus, T is $\mathcal{FD}$-BFNE for any $p \in (0,1)$. $\square$

(vi) $T(x) = \sin(x)$ is an $\mathcal{FD}$-BFNE operator on any subset of $(0,1)$.

Proof. Since $T'(x) = \cos(x)$ and

$$\frac{T(x)(1-T(x))}{x(1-x)} = \frac{\sin(x)(1-\sin(x))}{x(1-x)},$$

we get

$$\frac{\sin(x)(1-\sin(x))}{x(1-x)\cos(x)} = \frac{\sin(x)\cos(x)}{x(1-x)(1+\sin(x))} = \frac{\sin(2x)}{2x(1-x)(1+\sin(x))}.$$

On the one hand, we have

$$\frac{\sin(2x)}{2x} > \frac{2x-(1/6)(2x)^3}{2x} = 1-\frac{4}{6}x^2 > 1-x^2.$$

On the other hand,

$$(1-x)(1+\sin(x)) < (1-x)(1+x) = 1-x^2.$$

Hence

$$\frac{\sin(x)(1-\sin(x))}{x(1-x)\cos(x)} > 1,$$

which means that $\sin(x)$ is indeed an $\mathcal{FD}$-BFNE operator on any subset of $(0,1)$. $\square$

Again, there are many more such examples. $\Diamond$

$f(x)$	Domain	Condition
$\mathcal{BS}$	$(0,+\infty)$	$0 \leq T'(x) \leq \frac{T(x)}{x}$
$\mathcal{FD}$	$(0,1)$	$0 \leq T'(x) \leq \frac{T(x)(1-T(x))}{x(1-x)}$
$\cosh x$	$\mathbb{R}$	$0 \leq T'(x) \leq \frac{\cosh(x)}{\cosh(T(x))}$
$x^2/2$	$\mathbb{R}$	$0 \leq T'(x) \leq 1$
$x^4/4$	$\mathbb{R}$	$0 \leq T'(x) \leq \frac{x^2}{(T(x))^2}$
e^x	$\mathbb{R}$	$0 \leq T'(x) \leq \frac{e^x}{e^{T(x)}}$
$-\log(x)$	$(0,+\infty)$	$0 \leq T'(x) \leq \frac{(T(x))^2}{x^2}$

Table 1: Conditions for T to be an f-BFNE operators

4.1.3 Other admissible functions

In Table 1 we summarize sufficient conditions on the operator T to be f-BFNE with respect to various choices of functions f.

4.2 Constructions in general Euclidean spaces

Assume that $X = \mathbb{R}^n$. In this case the *Boltzmann-Shannon entropy* is the function $\mathcal{BS}_n : \mathbb{R}^n_{++} \to \mathbb{R}$ defined by

$$\mathcal{BS}_n\,(x) := \sum_{i=1}^{n} x_i \log(x_i) - x_i,\ \ x \in \mathbb{R}^n_{++}.$$

The following result shows that $\mathcal{BS}_n$ is a totally convex function (*cf.* [14]).

Proposition 17 (Total convexity of $\mathcal{BS}_n$ in $\mathbb{R}^n$)**.** *The function $\mathcal{BS}_n$ is totally convex and its modulus of total convexity satisfies*

$$\upsilon_{\mathcal{BS}_n}\,(x,t) \geq \min_{1 \leq i \leq n} \left\{ x_i \left[\left(1 + \frac{t}{x_i\sqrt{n}}\right) \log\left(1 + \frac{t}{x_i\sqrt{n}}\right) - \frac{t}{x_i\sqrt{n}} \right] \right\}.$$

Proof. Let $\overline{\mathcal{BS}} : [0,+\infty) \to \mathbb{R}$ be the continuous and convex function defined by

$$\overline{\mathcal{BS}}\,(x) := \begin{cases} x\log(x) - x, & x > 0 \\ 0, & x = 0. \end{cases}$$

It is clear that the restriction of $\overline{\mathcal{BS}}$ to $(0,+\infty)$ is exactly $\mathcal{BS}$. The function $\overline{\mathcal{BS}_n}$: $\mathbb{R}^n_{++} \to \mathbb{R}$ defined by

$$\overline{\mathcal{BS}_n}\,(x) := \sum_{i=1}^{n} \overline{\mathcal{BS}}\,(x_i)$$

is convex, continuous and its restriction to $\mathbb{R}^n_{++}$ is exactly $\mathcal{BS}_n$. Let

$$\upsilon_{\overline{\mathcal{BS}_n}}(x,t) = \inf\left\{D_{\overline{\mathcal{BS}_n}}(y,x) : y \in \mathbb{R}^n_+,\ \ \|y-x\| = t\right\}.$$

Since the set $\left\{y \in \mathbb{R}^n_+ : \|y-x\| = t\right\}$ is compact in $\mathbb{R}^n$ and $D_{\overline{\mathcal{BS}_n}}(\cdot,x)$ is continuous on this set, there exists $\bar{y} \in \mathbb{R}^n_+$ such that $\|x-\bar{y}\| = t$ and

$$\upsilon_{\mathcal{BS}_n}(x,t) \geq \upsilon_{\overline{\mathcal{BS}_n}}(x,t) = D_{\overline{\mathcal{BS}_n}}(\bar{y},x) = \sum_{i=1}^{n} D_{\overline{\mathcal{BS}}}(\bar{y}_i,x_i).$$

The modulus of total convexity of $\mathcal{BS}$ is given by (26) and is continuous in t. Therefore we can apply Proposition 1(vi) and obtain that, for each $1 \leq i \leq n$,

$$D_{\overline{\mathcal{BS}}}(\bar{y}_i,x_i) \geq \upsilon_{\mathcal{BS}}(x_i,\|x_i-\bar{y}_i\|).$$

Hence,

$$\upsilon_{\mathcal{BS}_n}(x,t) \geq \sum_{i=1}^{n} \upsilon_{\mathcal{BS}}(x_i,\|x_i-\bar{y}_i\|). \tag{34}$$

When $t>0$, we have $\|x_i-\bar{y}_i\| > 0$ for at least one index i. As noted in Proposition 9, the function $\mathcal{BS}$ is totally convex. Consequently, $\upsilon_{\mathcal{BS}}(x_i,\|x_i-\bar{y}_i\|) > 0$ for at least one index i. This and (34) show that, if $t>0$, then $\upsilon_{\mathcal{BS}_n}(x,t) > 0$, i.e., $\mathcal{BS}_n$ is totally convex.

Since for at least one index i_0 we have $\|x_{i_0}-\bar{y}_{i_0}\| \geq t/\sqrt{n}$, we deduce from (34) that

$$\begin{aligned}\upsilon_{\mathcal{BS}_n}(x,t) &\geq \sum_{i=1}^{n} \upsilon_{\mathcal{BS}}(x_i,\|x_i-\bar{y}_i\|) \geq \upsilon_{\mathcal{BS}}(x_{i_0},\|x_{i_0}-\bar{y}_{i_0}\|)\\ &\geq \upsilon_{\mathcal{BS}}\left(x_{i_0},t/\sqrt{n}\right) \geq \min_{1\leq i\leq n}\left\{\upsilon_{\mathcal{BS}}\left(x_i,t/\sqrt{n}\right)\right\}.\end{aligned}$$

When combined with (26), this inequality completes the proof. □

Remark 17 (Product constructions)**.** For each $i = 1,2,\cdots,n$, let $f_i : \mathbb{R} \to \mathbb{R}$ be an admissible function, and define the function $F : \mathbb{R}^n \to \mathbb{R}^n$ by

$$F(x_1,\cdots,x_n) = (f_1(x_1),\cdots,f_n(x_n)).$$

For each $i = 1,2,\cdots,n$, let K_i be a nonempty subset of int dom f_i. Let $T : \prod_{i=1}^{n} K_i \to \prod_{i=1}^{n}$ int dom f_i be the operator which is defined by $T = (T_1,T_2,\ldots,T_n)$, where $T_i : K_i \to$ int dom f_i for each $1 \leq i \leq n$. If each T_i, $i = 1,\cdots,n$, satisfies the hypotheses of Theorem 1, then the operator T is BFNE with respect to F on $\prod_{i=1}^{n} K_i$.

In a similar way we can take the function F to be the sum of any n admissible functions. Then the operator T is BFNE with respect to F if each operator T_i is BFNE with respect to the chosen function f_i, $i = 1, \cdots, n$. $\Diamond$

5 Examples of f-resolvents

As we explained in the introduction, any BFNE operator is a resolvent of a monotone mapping, and the resolvent plays an important role in the analysis of optimization problems. Therefore, in the following subsection, we provide several explicit examples of resolvents with respect to different choices of the admissible function f, for example, the Boltzmann-Shannon entropy and the Fermi-Dirac entropy.

5.1 Resolvents with respect to $\mathcal{BS}$

Let $A : (0, +\infty) \to \mathbb{R}$ be a monotone mapping. Then the resolvent of A with respect to $\mathcal{BS}$ is

$$\mathrm{Res}_A^{\mathcal{BS}} := (\log + A)^{-1} \circ \log .$$

Remark 18. We can also write the resolvent as follows:

$$\begin{aligned}\mathrm{Res}_A^{\mathcal{BS}} &:= \left(\left((\log + A)^{-1} \circ \log\right)^{-1}\right)^{-1} = \left((\log)^{-1} \circ (\log + A)\right)^{-1} \\ &= \left(e^{(\log + A)}\right)^{-1},\end{aligned}$$

where $\left(e^{(\log + A)}\right)(x) = xe^{A(x)}$. This naturally leads us to the Lambert W function. $\Diamond$

Recall [10, 11] that the *Lambert W function*, W, is defined to be the inverse of $x \mapsto xe^x$ and is implemented in both Maple and Mathematica. Its principal branch on the real axis is shown in Figure 3. Like log, it is concave increasing, and its domain is $(-1/e, +\infty)$.

Figure 3: The Lambert W function

We now give several examples of $\mathcal{BS}$-resolvents.

Example 3. (i) If $A(x) = \alpha$, $\alpha \in \mathbb{R}$, then $\mathrm{Res}_A^{\mathcal{BS}}(x) = e^{-\alpha}x$ for all $x \in \mathbb{R}_{++}$.
In particular, if $\alpha = 0$ then $\mathrm{Res}_A^{\mathcal{BS}}(x) = x$, $x \in \mathbb{R}_{++}$.

(ii) If $A(x) = \alpha x + \beta$, $\alpha, \beta \in \mathbb{R}$, then $\mathrm{Res}_A^{\mathcal{BS}}(x) = (1/\alpha)\,\mathrm{W}\left(\alpha e^{-\beta}x\right)$ for all $x \in \mathbb{R}_{++}$.
Hence, if $\alpha = 1$ and $\beta = 0$ then $\mathrm{Res}_A^{\mathcal{BS}}(x) = \mathrm{W}(x)$, $x \in \mathbb{R}_{++}$.

(iii) If $A(x) = \alpha \log(x)$, $\alpha \in \mathbb{R}$, then $\mathrm{Res}_A^{\mathcal{BS}}(x) = x^{1/(1+\alpha)}$ for all $x \in \mathbb{R}_{++}$.
Therefore, if $\alpha = 1$ then $\mathrm{Res}_A^{\mathcal{BS}}(x) = \sqrt{x}$, $x \in \mathbb{R}_{++}$.

(iv) If $A(x) = x^p/p$, $p > 1$, then $\mathrm{Res}_A^{\mathcal{BS}}(x) = (\mathrm{W}(x^p))^{1/p}$ for all $x \in \mathbb{R}_{++}$.
Thus, if $p = 2$ then $\mathrm{Res}_A^{\mathcal{BS}}(x) = \sqrt{\mathrm{W}(x^2)}$, $x \in \mathbb{R}_{++}$.

(v) If $A(x) = \mathrm{W}(\alpha x^p)$, $\alpha \in \mathbb{R}$ and $p \geq 1$, then

$$\mathrm{Res}_A^{\mathcal{BS}}(x) = \left(\frac{x}{\alpha(p+1)}\right)^{\frac{1}{p+1}} (\mathrm{W}(\alpha(p+1)x^p))^{\frac{1}{p+1}}$$

for all $x \in \mathbb{R}_{++}$.
Therefore, if $\alpha = 2$ and $p = 1$, then $\mathrm{Res}_A^{\mathcal{BS}}(x) = \sqrt{\frac{x}{4}}\sqrt{\mathrm{W}(4x)}$, $x \in \mathbb{R}_{++}$.

$\diamondsuit$

We now present an example of a $\mathcal{BS}$-resolvent in $\mathbb{R}^2$.

Example 4. Let $\mathcal{BS}_2(x, y) := x\log(x) + y\log(y) - x - y$. Thus $\nabla\mathcal{BS}_2(x, y) = (\log(x), \log(y))$. Let $\theta \in [0, \pi/2]$ and consider the rotation mapping $A_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ defined by

$$A_\theta(x, y) := \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

In particular, the $\mathcal{BS}$-resolvent of the rotation mapping $A_{\pi/2}$ is the operator:

$$\mathrm{Res}_{A_{\pi/2}}^{\mathcal{BS}_2} := (\nabla\mathcal{BS}_2 + A_{\pi/2})^{-1}(\nabla\mathcal{BS}_2).$$

We claim that the inverse of $\nabla\mathcal{BS}_2 + A_{\pi/2}$ uniquely exists. To see this, note that for any $x, y \in (0, +\infty)$, we have

$$\left(\nabla\mathcal{BS}_2 + A_{\pi/2}\right) \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \log(x) - y \\ \log(y) + x \end{pmatrix}.$$

Thus we have to show that for any $(z, w) \in \mathbb{R}^2$, there exist unique $x, y \in (0, +\infty)$ such that $z = \log(x) - y$ and $w = \log(y) + x$. These two equations can be written as

$$x = e^{y+z} \quad \text{and} \quad y = e^{w-x}.$$

Therefore, $x = e^{e^{w-x}+z}$. This equation has indeed a unique solution in $(0, +\infty)$. To check this, define a function $f : [0, +\infty) \to \mathbb{R}$ by $f(x) = x - e^{e^{w-x}+z}$. Then it is easy to see that $f(0) = -e^{e^{w}+z} < 0$ and $\lim_{x\to+\infty} f(x) = +\infty$. Since the function f is continuous, it has at least one root. On the other hand,

$$f'(x) = 1 - e^{e^{w-x}+z}\left(-e^{w-x}\right) = 1 + e^{e^{w-x}+w-x+z} > 0.$$

This means that f has exactly one root, which is the unique solution of the equation $x = e^{e^{w-x}+z}$. The general case is similar but less explicit. $\Diamond$

5.2 Resolvents with respect to $\mathcal{FD}$

Let $A : (0, 1) \to \mathbb{R}$ be a monotone mapping. Then the resolvent of A with respect to $\mathcal{FD}$ is

$$\mathrm{Res}_A^{\mathcal{FD}} := (\mathcal{FD}' + A)^{-1} \circ \mathcal{FD}',$$

where in this case $\mathcal{FD}'(x) = \log(x/(1-x))$ and therefore $(\mathcal{FD}')^{-1}(x) = e^x/(1+e^x)$.

Remark 19. We can also write the resolvent in the following way

$$\begin{aligned}\mathrm{Res}_A^{\mathcal{FD}} &:= \left(\left((\mathcal{FD}' + A)^{-1} \circ \mathcal{FD}'\right)^{-1}\right)^{-1} = \left((\mathcal{FD}')^{-1} \circ (\mathcal{FD}' + A)\right)^{-1} \\ &= \left(\frac{e^{(\mathcal{FD}'+A)}}{1 + e^{(\mathcal{FD}'+A)}}\right)^{-1},\end{aligned}$$

where

$$\left(\frac{e^{(\mathcal{FD}'+A)}}{1 + e^{(\mathcal{FD}'+A)}}\right)(x) = \frac{xe^{A(x)}}{1 - x + xe^{A(x)}}.$$

$\Diamond$

Several examples of resolvents with respect to $\mathcal{FD}$ follow.

Example 5. (i) If $A(x) = \alpha$, $\alpha \in (0, +\infty)$, then

$$\mathrm{Res}_A^{\mathcal{FD}}(x) = \frac{x}{x + e^{\alpha}(1-x)}, \; x \in (0, 1).$$

If $\alpha = 0$, then $\mathrm{Res}_A^{\mathcal{FD}}(x) = x$, $x \in (0, 1)$.

(ii) If $A(x) = \log(x)$, then

$$\mathrm{Res}_A^{\mathcal{FD}}(x) = \frac{x - \sqrt{4x - 3x^2}}{2(x-1)}$$

for all $x \in (0,1)$.

(iii) If $A(x) = \log(1-x)$, then

$$\mathrm{Res}_A^{\mathcal{FD}}(x) = \frac{x}{1-x}$$

for all $x \in (0,1)$.

(vi) If $A(x) = 2 * \log(1-x)$, then

$$\mathrm{Res}_A^{\mathcal{FD}}(x) = \frac{1 - x - \sqrt{5x^2 - 6x + 1}}{2(x-1)}$$

for all $x \in (0, 1/5]$.

$\Diamond$

Finally, Table 2 lists resolvents with respect to various choices of the function f. Here $\mathrm{Res}_A^f = g^{-1}$.

$f(x)$	Domain	$g(x)$
$\mathcal{BS}$	$(0, +\infty)$	$xe^{A(x)}$
$\mathcal{FD}$	$(0,1)$	$\frac{xe^{A(x)}}{1-x+xe^{A(x)}}$
$x^2/2$	$\mathbb{R}$	$x + A(x)$
$x^4/4$	$\mathbb{R}$	$(x^3 + A(x))^{1/3}$
e^x	$\mathbb{R}$	$\log(e^x + A(x))$
$-\log(x)$	$(0, +\infty)$	$\frac{x}{1-xA(x)}$

Table 2: Examples of Resolvents

Acknowledgements. This research was supported by the Australian Research Council, the Israel Science Foundation (Grant 647/07), the Graduate School of the Technion, the Fund for the Promotion of Research at the Technion and by the Technion President's Research Fund. It began when the last author was a visiting student fellow at CARMA at the University of Newcastle. He is grateful to the University of Newcastle and his colleagues there for their warm hospitality and financial support. We are very grateful to Heinz H. Bauschke for his many helpful comments and corrections. We also thank Liangjin Yao for Proposition 14.

References

[1] Alber, Ya. I. and Reich, S.: An iterative method for solving a class of nonlinear operator equations in Banach spaces, *Panamer. Math. J.* **4** (1994), 39–54.

[2] Bauschke, H. H., Personal communication, 2011.

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

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

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

[6] Bauschke, H. H. and Combettes, P. L.: A weak-to-strong convergence principle for Fejér-monotone methods in Hilbert spaces, *Math. Oper. Res.* **26** (2001), 248–264.

[7] Bauschke, H. H. and Combettes, P. L.: Construction of best Bregman approximations in reflexive Banach spaces, *Proc. Amer. Math. Soc.* **131** (2003), 3757–3766.

[8] Bauschke, H. H., Wang, X. and Yao, L.: General resolvents for monotone operators: characterization and extension, *Biomedical Mathematics: Promising Directions in Imaging, Therapy Planning and Inverse Problems*, Medical Physics Publishing, Madison, WI, USA, 2010, pp. 57–74.

[9] Beck, A. and Teboulle, M.: Mirror descent and nonlinear projected subgradient methods for convex optimization, *Oper. Res. Lett.* **31** (2003), 167-175.

[10] Borwein, J. M. and Lewis, A. S.: Convex Analysis and Nonlinear Optimization. Theory and Examples, *CMS (Canadian Mathematical Society) Springer-Verlag*, New York. Second extended edition, 2005. Paperback, 2009.

[11] Borwein J. M. and Vanderwerff, J.: Convex Functions: Constructions, Characterizations and Counterexamples, Encyclopedia of Mathematics and Applications, *Cambridge University Press*, 2010.

[12] Borwein, J. M. and Vanderwerff, J.: Constructions of uniformly convex functions, preprint 2010.

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

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

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

[16] Butnariu, D. and Kassay, G.: A proximal-projection method for finding zeroes of set-valued operators, *SIAM J. Control Optim.* **47** (2008), 2096–2136.

[17] Butnariu, D. and Resmerita, E.: Bregman distances, totally convex functions and a method for solving operator equations in Banach spaces, *Abstr. Appl. Anal.* **2006** (2006), Art. ID 84919, 1–39.

[18] Calvert, B.: The range of T-monotone operators, *Boll. Un. Math. Ital.* **4** (1971), 132–144.

[19] Censor, Y. and Reich, S.: Iterations of paracontractions and firmly nonexpansive operators with applications to feasibility and optimization, *Optimization* **37** (1996), 323–339.

[20] Chen, G. and Teboulle, M.: Convergence analysis of a proximal-like minimization algorithm using Bregman functions, *SIAM J. Optim.* **3** (1993), 538–543.

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

[22] Klir, G. J. and Folger, T. A.: Fuzzy Sets, Uncertainty, and Information, Prentice Hall, Englewood Cliffs, NJ, 1988.

[23] Phelps, R. R.: Convex Functions, Monotone Operators, and Differentiability, Lecture Notes in Mathematics **1364**, 2nd Edition, *Springer Verlag*, Berlin, 1993.

[24] Reich, S.: Extension problems for accretive sets in Banach spaces, *J. Functional Analysis* **26** (1977), 376–395.

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

[26] Reich, S. and Sabach, S.: A strong convergence theorem for a proximal-type algorithm in reflexive Banach spaces, *J. Nonlinear Convex Anal.* **10** (2009), 471–485.

[27] Reich, S. and Sabach, S.: Two strong convergence theorems for a proximal method in reflexive Banach spaces, *Numer. Funct. Anal. Optim.* **31** (2010), 22–44.

[28] Reich, S. and Sabach, S.: Existence and approximation of fixed points of Bregman firmly nonexpansive mappings in reflexive Banach spaces, *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, Springer, New York, accepted for publication.

[29] Reich, S. and Sabach, S.: Two strong convergence theorems for Bregman strongly nonexpansive operators in reflexive Banach spaces, *Nonlinear Anal.* **73** (2010), 122–135.

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

[31] Teboulle, M.: Entropic proximal mappings with application to nonlinear programming, *Math. Oper. Res.* **17** (1992) 670-690.

[32] Yao, Liangjin, Personal communication, 2011.

[33] Zălinescu, C.: Convex analysis in general vector spaces, *World Scientific Publishing*, Singapore, 2002.