Int. J. Math. And Appl., 7(2)(2019), 23–38
ISSN: 2347-1557
Available Online: http://ijmaa.in/

International Journal of Mathematics And its Applications

# Entropic Approximation

**S. Kabbadj[1,*]**

1 Department of Mathematics, Faculty of Sciences of Meknes, B.P. 11201, Marocco.

**Abstract:** In this article, we study the theoretical aspect of the entropic approximations of a convex function on $R^p$, we put in evidence it properties of regulirization and approximation of those approximates, we find the most of the approximal properties of Moreau-Yosida.

**Keywords:** Convex minimization, Approximation of Moreau-Yosida, Bregman's function, Legendre's function, Entropic approximation.



## 1. Introduction

Let f be a lower semicontinuous, proper and convex function on a Hilbert H. Moreau introduced and studied the approximation Moreau-Yosida $f_\lambda$ defined by

$$f_\lambda(x) := \inf_z \left\{ f(z) + \frac{1}{2\lambda} \parallel x - z \parallel^2 \right\}, \ \forall\ x \in H, \ \forall\ \lambda > 0,$$

as well as the proximal mapping $prox_{\lambda f}$ defined by:

$$prox_{\lambda f}(x) := \arg\min_z \left\{ f(z) + \frac{1}{2\lambda} \parallel x - z \parallel^2 \right\}.$$

$f_\lambda$ regulates f and has several important properties that make it very useful in optimization theory; it provides numerical methods for solving convex optimization problems, including the algorithme of the proximal point [3,10]. In the same vein, where $H = R^p$, Teboulle [11,12] introduced a class of approximations, replacing the quadratic kernel with a so-called entropic kernel $D_h(.,.)$ defined by:

$$D_h(x, y) := h(x) - h(y) - \langle x - y, \nabla h(y) \rangle.$$

Thus, he defined the entropic approximation by:

$$f_{h\lambda}(x) := \inf_z \left\{ f(z) + \lambda^{-1} D_h(z, x) \right\},$$

and studied some properties of this function where h is a function of Legendre. On this labor we put in evidence the properties of regularization and approximation of those approximates , while $D_h(.,.)$ is not a distance.This study covers

* E-mail: kabbajsaid63@yahoo.com

most of the approximate properties of Moreau-Yosida, however the contraction of the operator $prox^h_{\lambda f}$ is realized only under reinforced conditions on h or on f. This study has a determining role for the study of the entropic proximal algorithms [7,11,12]

$$x^n := \arg\min_z \left\{ f(z) + \frac{1}{\lambda_n} D_h(z, x^{n-1}) \right\}, \lambda_n > 0.$$

In the section 2, we remind the fundamental properties of the approximate $f_\lambda$. and of the proximal mapping $prox_{\lambda f}$. In section the 3, we introduce the set of functions of Legendre $L(C)$; the set of Bregman's functions $B(S)$ as well as the set $C(S)$ which is very useful for the convergence of inexact entropic proximal algorithms. We give in section 4 some examples of functions of these three sets. In the section 5, we study $f_{h\lambda}$ and $prox^h_{\lambda f}$, especially in the case where h is such as $Im\nabla h = R^p$, condition realised for most interesting kernel. Our notation is fairly standart; $\langle ., . \rangle$ is the scalar product on H; and the associated norm $\|.\|$. The closure of the set C (interior, relative interior) by $\overline{C}$ (intC , riC, respectively). For any convex function f, we denote by :

(1). $domf = \{x \in H; f(x) < +\infty\}$ its effective domain,

(2). $f(.) = \sup_x\{\langle ., x\rangle - f(x)\}$ its conjugate,

(3). $\partial_\epsilon f(.) = \{v, f(y) \geq f(.) + \langle v, y - .\rangle - \epsilon, \forall\, y\}$ its $\varepsilon$-subdifferential

(4). $Arg \min_{x\in H} = \{x \in H; f(x) = \inf_H f\}$ its Argmin.

## 2. Approximation of Moreau-Yosida

Let $\Gamma_0(H)$ set of proper lower semicontinuous convex functions $f : H \to \overline{R}$. In this section we suppose that $f \in \Gamma_0(H)$, and we remind the properties of approximation of Moreau-Yosida.

**Proposition 2.1** ([1])**.** *Let $x \in H$ and $\lambda > 0$*

*(1). $prox_{\lambda f}(x)$ exist and unique.*

*(2). $J^f_\lambda(x) := prox_{\lambda f}(x) = (I + \lambda\partial f)^{-1}(x)$ where I is the identity operator of H.*

*(3). $A^f_\lambda(x) := \left(\frac{I-J^f_\lambda}{\lambda}\right)(x) \in \partial f(J^f_\lambda(x))$.*

*(4). The operator $J^f_\lambda$ is contractor, i.e. $\forall\, x, x' \in H$, $\| J^f_\lambda(x) - J^f_\lambda(x') \| \leq \| x - x' \|$.*

**Proposition 2.2** ([1])**.** *$f_\lambda$ is Frchet differentiable on H and $\nabla f_\lambda(x) = A^f_\lambda(x)$, $\forall x \in H$.*

**Proposition 2.3** ([1])**.** *Let $f \in \Gamma_0(H)$, for all $x \in domf$, we have the following properties:*

*(1). $J^f_\lambda(x)$ converges strongly to x if $\lambda \to 0$,*

*(2). $f(J^f_\lambda(x)) \to f(x)$ if $\lambda \to 0$.*

## 3. Entropic Distances

In this section, we introduce the set of functions of Legendre $L(C)$ for $H = R^p$, the set of Bregman's functions $B(S)$ thus a new set $C(S)$ where the role is determinant for the convergence of the inexact proximal algorithms. We then study the class of the kernels of $D_h(., .)$ defined by:

$$D_h(x, y) = h(x) - h(y) - \langle x - y, \nabla h(y)\rangle.$$

**Definition 3.1** ([9]). *Let $C$ be a convex not empty of $R^p$*

*(1). A convex function $h : R^p \rightarrow ]-\infty, +\infty]$ is of Legendre on $C$ if it verifies the three following conditions:*

*(a). $C = int\ (domh)$*

*(b). $h$ is differentiable on $C$*

*(c). $\lim \parallel \nabla h(x_i) \parallel = +\infty$, for any sequence $\{x_i\}$ of $C$ that converges towards a boundary point of $C$.*

*(2). The class of strictly convex functions verifying a, b and c is called the class of Legendre's functions on $C$ and denoted by $L(C)$.*

*(3). A regular convex function on $R^p$ (ie: finite and differentiable) is in particular essentially regular; the set of those functions are denoted by $L(R^p)$.*

*(4). $h$ is co-finite if : $domh^* = R^p$.*

**Proposition 3.2** ([9]). *Let $h \in L(R^p)$.*

$$h \text{ is co-finite} \Leftrightarrow \lim_{\|x_i\| \longrightarrow +\infty} \|\nabla h(x_i)\| = +\infty$$

**Proposition 3.3** ([9]). *Let $h \in L(C)$.*

*(1). $h^* \in L(C^*)$ where $C^* = int(domh^*)$.*

*(2). $\nabla h^* = (\nabla h)^{-1}$.*

*(3). $h^*(\nabla h(z)) = \langle z, \nabla h(z) \rangle - h(z)$.*

Let S be an convex open subset of $R^p$ and $h : \overline{S} \rightarrow R$. Let us consider the following hypotheses:

$H_1$: h is continuously differentiable on S.

$H_2$: h is continuous and strictly convex on $\overline{S}$.

$H_3$: $\forall\ r \geq 0, \forall x \in \overline{S}, \forall\ y \in S$, the sets $L_1(x, r)$ and $L_2(y, r)$ are bounded where:

$$L_1(x, r) = \{y \in S / D_h(x, y) \leq r\}$$
$$L_2(y, r) = \{x \in \overline{S} / D_h(x, y) \leq r\}.$$

$H_4$: If $\{y^k\}_k \subset S$ is such as $y^k \rightarrow y^* \in \overline{S}$ , so $D_h(y^*, y^k) \rightarrow 0$.

$H_5$: If $\{x^k\}_k \subset \overline{S}$ is such as $\{Y^k\}_k \subset S$ are such as:

$$y^k \longrightarrow y^* \in \overline{S}\ ,\ \{x^k\}_k \text{ is bounded, and } D_h(x^k, y^k) \longrightarrow 0, \text{ then } x^k \longrightarrow y^*.$$

$H_6$: If $\{x^k\}_k$ and $\{y^k\}_k$ are two sequences of S such as:

$$D_h(x^k, y^k) \longrightarrow 0 \text{ and } x^k \longrightarrow x^* \in S, \text{ then } y^k \longrightarrow x^*.$$

$H_7$: $Im \nabla h = R^p$.

**Definition 3.4.**

*(1). $h : \overline{S} \rightarrow R$ is a Bregman function on $S$ or "D-function" if $h$ verify $H_1, H_2, H_3, H_4$ and $H_5$.*

*(2).* $D_h(.,.) : \overline{S} X S \to R$ *such us :* $\forall\ x \in \overline{S}, \forall\ y \in S$

$$D_h(x, y) = h(x) - h(y) - \langle x - y, \nabla h(y)\rangle$$

*is called entropic distance if h is a Bregman function. We put:*

$$A(S) = \{h : \overline{S} \to R \ verifying\ H_1 \ and\ H_2\}$$

$$B(S) = \{h : \overline{S} \to R \ verifying\ H_1, H_2, H_3, H_4 \ and\ H_5\}$$

$$C(S) = \{h : \overline{S} \to R \ verifying\ H_1, H_2, H_3, H_4 \ and\ H_6\}.$$

*If* $h \in A(R^p)$*, then the hypotheses* $H_4$ *and* $H_5$ *are verified.*

**Proposition 3.5.** *Let us assume:*

*(1).* $h \in A(S)$,

*(2). h is strongly convex on S with parametr* $\alpha$.

*Then,*

*(a).* $\forall\ x \in \overline{S}, \forall\ y \in S\ ,\ D_h(x, y) \geq \frac{\alpha}{2}\|x - y\|^2$.

*(b). The hypotheses* $H_5$ *and* $H_6$ *are verified.*

*Proof.*

(a). By virtue of the differentiability of h and 2, we have:

$$\forall\ u, v \in S\ , \langle \nabla h(u) - \nabla h(v), u - v\rangle \geq \alpha\|u - v\|^2. \tag{1}$$

For $x \in \overline{S}$ and $y \in S$, for all $t \in [0, 1[\,, y + t(x - y) \in S$. Let $k : [0, 1] \to R$, the function defined by:

$$k(t) = h(y + t(x - y)). \tag{2}$$

k is a derivable convex and

$$k^{'}(t) = \langle \nabla h(y + t(x - y)), x - y\rangle.$$

Then

$$k(1) = k(0) + \int_0^1 k^{'}(t)dt,$$

that means;

$$h(x) - h(y) - \langle \nabla h(y), x - y\rangle = \int_0^1 \langle \nabla h(y + t(x - y)) - \nabla h(y), x - y\rangle dt. \tag{3}$$

From (1), $D_h(x, y) \geq \alpha \int_0^1 tdt\|x - y\|^2$, what establishes the wanted inequality.

(b). We get: $D_h(x, y) \geq \frac{\alpha}{2}\|x - y\|^2$. We replace x by $x^k$ and y by $y^k$ in the previous inequality; we obtain then:

$$D_h(x^k, y^k) \geq \frac{\alpha}{2}\|x - y\|^2.$$

If $D_h(x^k, y^k) \to 0$ and $y^k \to y^* \in \overline{S}$ then $x^k \to y^*$ i.e. $H_5$. If $D_h(x^k, y^k) \to 0$ and $x^k \to x^* \in S$ then $y^k \to y^*$ , i.e. $H_6$. □

**Proposition 3.6** ([5]). *If $h \in A(S)$; then:*

$$D_h(x,y) = \begin{cases} 0 & if \;\; x = y, \\ > 0 & if \;\; x \neq y \end{cases}$$

**Remark 3.7.** *$D_h(.,.)$ is not a distance because the properties of the symmetry and the triangle inequality are not verified. In [5], it was proven that $D_h(.,.)$ is symmetric in the unique case where h is defined by $h(x) = x^T Qx + q^T x,\; q \in R^p$ where $Q$ a squared matrix of order p symmetric and positive definite.*

**Proposition 3.8.** *Let h and $h'$ verify $H_1$.*

$$\forall\, \lambda, D_{\lambda h + h'}(.,.) = \lambda D_h(.,.) + D_{h'}(.,.).$$

**Proposition 3.9.** *Let $h \in A(s)$ such as;*

*(1). h strongly convex on S with parameter $\alpha$.*

*(2). It exist $\beta > 0$ such as: $\|\nabla h(x) - \nabla h(y)\| \leq \beta \|x - y\|,\; \forall x, y \in S$.*

*Then:*

$$\forall\; x \in \overline{S}, \forall\; y \in S,\; \frac{\alpha}{2}\|x-y\|^2 \leq D_h(x,y) \leq \frac{\beta}{2}\|x-y\|^2.$$

**Lemma 3.10** ([11]). $\forall\; h \in A(S), \forall\; a \in \overline{S}, \forall\; b, c \in S$

$$D_h(a,b) + D_h(b,c) - D_h(a,c) = \langle a - b, \nabla h(c) - \nabla h(b)\rangle.$$

**Corollary 3.11.**

*(1).* $\forall\; h \in A(S), \forall\; a, b \in S,$

$$D_h(a,b) + D_h(b,a) \leq \|a-b\|\|\nabla h(a) - \nabla h(b)\|$$

*(2). Let $h \in A(S)$ and let $\{x^k\} \subset S$ such as $x^k \longrightarrow x^* \in S$ , so $D_h(x^*, x^k) \longrightarrow 0$ and $D_h(x^k, x^*) \longrightarrow 0$.*

*Proof.*

(1). By replacing c by a in the Lemma 3.10, we obtain: $D_h(a,b) + D_h(b,a) = \langle a - b, \nabla h(a) - \nabla h(b)\rangle$, from when the result.

(2). By replacing a by $x^*$ and b by $x^k$ on (1), we have:

$$D_h(x^*, x^k) + D_h(x^k, x^*) \leq \|x^* - x^k\|\|\nabla h(x^k) - \nabla h(x^*)\|,$$

which results that $D_h(x^k, x^*) \longrightarrow 0$ and $D_h(x^*, x^k) \longrightarrow 0$. □

**Proposition 3.12.** *Let $h \in A(S)$. If $Im\nabla h = R^p$ then $h \in L(S)$.*

*Proof.* (a) and (b) of the Definition 3.1 being verified, let's demontrate that the condition (c) is verified too, which means:

$$\lim_{x_i \rightarrow x^* \in Fr(S)} \|\nabla h(x_i)\| = +\infty.$$

Let $\{x_i\}$ such as $x_i \to x^* \in Fr(S) = \overline{S}/S$. If $\lim\limits_{x_i \to x^* \in Fr(S)} \|\nabla h(x_i)\| \neq +\infty$ then it would exist a subsequence $\{\nabla h(x_{i_k})\}$ of $\{\nabla h(x_i)\}$ bounded, so it would exist a subsequence $\{\nabla h(x_{i_n})\}$ of $\{\nabla h(x_i)\}$ such as

$$\nabla h(x_{i_n}) \to u^*.$$

Since $Im \nabla h = R^p$ it exists that $u \in S$, such as $\nabla h(u) = u^*$. We have then;

$$\nabla h(x_{i_n}) \to \nabla h(u).$$

From another part , from the Corollary 3.11,

$$\begin{aligned} & D_h(u, x_{i_n}) + D_h(x_{i_n}, u) \leq \|u - x_{i_n}\|.\|\nabla h(u) - \nabla h(x_{i_n})\| \\ \Rightarrow \quad & \lim_{i_n \to \infty} D_h(x_{i_n}, u) = 0 \\ \Rightarrow \quad & D_h(x^*, u) = 0 \\ \Rightarrow \quad & x^* = u, \end{aligned}$$

thing which is contradictory with $x^* \in \overline{S} \backslash S$ since $u \in S$. □

# 4. Examples of Bregman Functions

**Example 4.1.** *If* $S_0 = R^p$ *and* $h_0(x) = \frac{1}{2} \parallel x \parallel^2$ *then* $D_{h_0}(x, y) = \frac{1}{2} \parallel x - y \parallel^2$.

**Example 4.2.** *If* $S_1 = R^p_{++} := \{x \in R^p / x_i > 0, i = 1, \ldots, p\}$ *and*

$$h_1(x) = \sum_{i=1}^{i=p} x_i \log x_i - x_i; \ \forall\ x \in \overline{S_1},$$

*with the convention :* $0 \log 0 = 0$, *then*

$$D_{h_1}(x, y) = \sum_{i=1}^{p} x_i \log \frac{x_i}{y_i} + y_i - x_i, \ \forall\ (x, y) \in \overline{S}_1 X S_1.$$

**Example 4.3.** *If* $S_2 = ]-1, 1[^p$ *and* $h_2(x) = -\sum\limits_{i=1}^{i=p} \sqrt{1 - x_i^2}$, *then:*

$$D_{h_2}(x, y) = h_2(x) + \sum_{i=1}^{p} \frac{1 - x_i y_i}{\sqrt{1 - y_i^2}}, \ \forall\ (x, y) \in \overline{S_2} X S_2.$$

**Proposition 4.4.** $h_i \in B(S_i) \cap C(S_i) \cap L(S_i), i = 1, 2, 3.$

# 5. Entropic Approximations

On this paragraph, we introduce the entropic approximation defined in a point $x \in S$ by:

$$\inf_{y \in \overline{S}} \{f(y) + \lambda^{-1} D_h(y, x)\},$$

thus the proximal entropic mapping defined by

$$\arg\min_{y \in \overline{S}} \{f(y) + \lambda^{-1} D_h(y, x)\}.$$

We study the properties of those functions for the class of the functions h belonging at $A(S)$ and verifying $H_7$, covering the most of the approximation properties of Moreau-Yosida reminded in 2.

**Proposition 5.1.** *Let $f \in \Gamma_0(R^p)$ and $h \in A(S)$ such as dom $f \cap \overline{S} \neq \phi$. Let $x \in S$ and $\lambda > 0$ such as:*

$$ri(dom\ f^*) \cap int(\lambda^{-1}(\nabla h(x) - dom\ h^*)) \neq \phi.$$

*Then, the function: $u \mapsto f(u) + \lambda^{-1} D_h(u, x)$ reaches a minimum in a unique point on $\overline{S}$.*

*Proof.* Uniqueness: $f(.) + \lambda^{-1} D_h(., x)$ is strictly convex thanks to $H_2$.

Existence: For that it's enough to demonstrate that: $\forall r \in R$,

$$L(x, r) = \{u : f(u) + \lambda^{-1} D_h(u, x) \leq r\},$$

which is closed, and bounded when it's not empty. Let $y \in ri(dom\ f^*) \cap int(\lambda^{-1}(\nabla h(x) - dom\ h^*))$. Since $ri(dom\ f^*) \subset Im\partial f$, $y \in Im\partial f$, which means: it exists z such as: $\forall\ u, f(u) \geq f(z) + \langle u - z, y\rangle$, it follows that ,

$$\begin{aligned} L(x, r) &\subset \{u : f(z) + \langle u - z, y\rangle + \lambda^{-1} D_h(u, x) \leq r\} \\ &= \{u : h(u) - \langle u, \nabla h(x) - \lambda y\rangle \leq K\}, \end{aligned}$$

where $K = \lambda(r - f(z) + \langle z, y\rangle) + h(x) - \langle x, \nabla h(x)\rangle$. Let

$$v := \nabla h(x) - \lambda y \text{ and } g(u) := h(u) - \langle u, v\rangle.$$

To show that $L(x, r)$ is bounded brings back then to prove that $0 \in int(dom\ g^*)$.

$$g^*(w) = \sup_u\{\langle w, u\rangle - g(u)\} = \sup_u\{\langle w + v, u\rangle - h(u)\} = h^*(w + v),$$

consequently,

$$dom\ g^* = dom\ h^* - v.$$

$0 \in int(dom\ g^*) \Leftrightarrow v \in int(dom\ h^*) \Leftrightarrow \nabla h(x) - \lambda y \in int(dom\ h^*) \Leftrightarrow y \in int(\lambda^{-1}(\nabla h(x) - dom\ h^*))$. □

**Theorem 5.2.** *Let $f \in \Gamma_0(R^p)$ and $h \in A(S)$ such as dom $f \cap \overline{S} \neq \phi$. If one of the two following conditions are verified:*

*(1). $\inf_{\overline{S}} f > -\infty$ and $h$ verify $H_3$.*

*(2). $Im\nabla h = R^p$.*

*Then for all $x \in S$, for all $\lambda > 0$, the function $u \longmapsto f(u) + \lambda^{-1} D_h(u, v)$ reaches it minimum in a unique point on $\overline{S}$.*

*Proof.*

(1). As previously, it 's enought to demonstrate that: $\forall\ r \in R$,

$$L(x, r) := \{u : f(u) + \lambda^{-1} D_h(u, x) \leq r\}$$

is bounded when it is not empty.

$$\begin{aligned} u \in L(x, r) \Rightarrow & f(u) + \lambda^{-1} D_h(u, x) \leq r \\ \Rightarrow & \quad D_h(u, x) \leq \lambda(r - \inf_{\overline{S}} f). \end{aligned}$$

It follows that

$$L(x, r) \subset L_2\left(x, \lambda\left(r - \inf_{\overline{S}} f\right)\right),$$

thanks to $H_3$, $L_2\left(x, \lambda\left(r - \inf_{\overline{S}} f\right)\right)$ is bounded , which leads that $L(x, r)$ is bounded too.

(2). $Im\nabla h = R^p$ and $Im\nabla h \subset domh^* \Rightarrow dom\ h^* = R^p$. Consequently, the condition of the Proposition 5.1 is verified, for all $x \in S$ and for all $\lambda > 0$, whence the desired result. □

**Definition 5.3.** *f and h verify the hypotheses of the Theorem 5.2.*

*(1). The entropic approximation of f compared to h, of parameter $\lambda(\lambda > 0)$ is the function defined by :*

$$f_{h\lambda}(x) := \inf_{y\in\overline{S}}\{f(y) + \lambda^{-1}D_h(y,x)\},\ \forall\ x \in S.$$

*(2). The application entropic proximal of f comparing to h, of parameter $\lambda$ is the operator defined by:*

$$h_\lambda^f(x) := prox_{\lambda f}^h(x) := \arg\min_{y\in\overline{S}}\{f(y) + \lambda^{-1}D_h(y,x)\},\ \forall\ x \in S.$$

*Now, we search at which conditions*

$$h_\lambda^f(x) \in S,\ \forall\ x \in S,$$

*this is in order to verify the effectiveness of algorithms [9].*

**Theorem 5.4** ([7])**.** *Let $f \in \Gamma_0(R^p)$ and $h \in B(S)$. If one of the two following conditions are verified:*

*(1). f is at finite values and $inf f > -\infty$*

*(2). dom $f \subseteq S$ and h verify $H_7$.*

*Then $h_\lambda^f(x) \in S,\ \forall\ x \in S$.*

From another approach, we are going to improve the condition 2. of this Theorem.

**Lemma 5.5.** *If $h \in L(S)$. Then $\forall\ u \in S$,*

$$\partial(D_h(.,u))(x^*) = \begin{cases} \{\nabla h(x^*) - \nabla h(u)\} & if\ \ x^* \in S \\ \emptyset & if\ not \end{cases}$$

*Proof.* Since h is a Legendre function on S, $D_h(.,u)$ it is too. By application of Theorem 26.1 [9], $\partial(D_h(.;u))$ verifies:
$-$If $x^* \in int(dom\ D_h(.,u)) = S$ then $\partial(D_h(.,u))(x^*) = \{\nabla(D_h(.,u)(x^*)\}$

$$D_h(x,u) = h(x) - h(u) - \langle x - u, \nabla h(u)\rangle,$$

h is differentiable on S, so $\nabla(D_h(.,u)(x^*) = \nabla h(x^*) - \nabla h(u)$.
$-$If $x^* \notin S$ then $\partial(D_h(.,u))(x^*) = \emptyset$. □

**Theorem 5.6** ([9])**.** *If $f_1, f_2, \ldots, f_m$ are convex and proper functions on $R^p$, then:*

$$\partial(f_1, f_2 + \cdots + f_m)(x) \supset \partial f_1(x) + \partial f_2(x) + \cdots + \partial f_m(x).$$

*If furthermore, $\cap(ri\ dom\ f_i) \neq \emptyset$, then;*

$$\partial(f_1, f_2 + \cdots + f_m)(x) = \partial f_1(x) + \partial f_2(x) + \cdots + \partial f_m(x), \forall\ x.$$

**Theorem 5.7.** *Let $h \in L(S)$ and $f \in \Gamma_0(R^p)$ such as $ri(dom\ f) \cap S \neq \emptyset$. Let $x \in S$ and $\lambda > 0$ such as $ri(dom\ f^*) \cap int(\lambda^{-1}(\nabla h(x) - dom\ h^*)) \neq \emptyset$. Then $h_\lambda^f(x) \in S$.*

*Proof.* From the Proposition 5.1. $h_\lambda^f(x) \in \overline{S}$. Let's suppose that $h_\lambda^f(x) \notin S$.

$$h_\lambda^f(x) = \arg\min_{y \in \overline{S}} \{f(y) + \lambda^{-1} D_h(y, x)\},$$

which leads to

$$0 \in \partial(f + \lambda^{-1} D_h(., x))(h_\lambda^f(x)).$$

$ri(dom\ f) \cap ri(dom\ D_h(., x)) = ri(dom\ f) \cap ri(\overline{S}) = ri(dom\ f) \cap S \neq \emptyset$, and from the Theorem 5.6;

$$0 \in \partial f(h_\lambda^f(x)) + \lambda^{-1} \partial(D_h(., x))(h_\lambda^f(x)).$$

It follows that: $\exists\ u : u \in \partial f(h_\lambda^f(x))$. Such as:

$$-\lambda u \in \partial(D_h(., x)(h_\lambda^f(x)).$$

Is in contradiction with the Lemma 5.5. □

**Corollary 5.8.** *Let $h \in A(S)$ and $f \in \Gamma_0(R^p)$ such as:*

*(1). $ri(dom\ f) \cap S \neq \phi$,*

*(2). $Im\nabla h = R^p$.*

*Then $h_\lambda^f(x) \in S$, $\forall\ x \in S$, $\forall\ \lambda > 0$.*

*Proof.* We get, $Im\nabla h \subset dom\ h^*$, since $Im\nabla h = R^p$, we deduce from this, that $dom\ h^* = R^p$. Consequently;

$$ri(dom\ f^*) \cap int(\lambda^{-1}(\nabla h(x) - dom\ h^*)) \neq \emptyset, \forall\ x \in S, \forall\ \lambda > 0.$$

By application of Theorem 5.7 and of the Proposition 3.10, we obtain the result. □

**Remark 5.9.** *The Theorem 5.4, 2. appears then as a consequence of Corollary 5.8 it brings a prove at the affirmation below of Chen and Teboulle on [11]:*

*If $ridom\ f \subset S$ and $Im\nabla h = R^p$, then $h_\lambda^f(x) \in S$, for all $x \in S$. We give the properties of the proximal entropic function $h_\lambda^f$ such as propositions.*

**Proposition 5.10.** *Let $A := \{x \in S, h_\lambda^f(x) \in S\}$. If $ri(dom\ f) \cap S \neq \emptyset$, then*

*(1). $\frac{\nabla h(x) - \nabla h(h_\lambda^f(x))}{\lambda} \in \partial f(h_\lambda^f(x))$, $\forall\ x \in A$;*

*(2). $h_\lambda^f = (\nabla h + \lambda \partial f)^{-1} o \nabla h$, on $A$.*

*Proof.*

(1). $h_\lambda^f(x) = \arg_u \min\{f(u) + \lambda^{-1} D_h(u, x)\} \Leftrightarrow 0 \in \partial\left[f(.) + \lambda^{-1} D_h(., x)\right](h_\lambda^f(x))$. As $ri(dom\ f) \cap S \neq \emptyset$, from the Theorem 5.6,

$$0 \in \partial f(h_\lambda^f(x)) + \lambda^{-1} \nabla(D_h(., x))(h_\lambda^f(x)),$$

thanks to Lemma 5.5, we deduct (1).

(2). (1) $\Leftrightarrow \nabla h(x) \in \nabla h(h_\lambda^f(x)) + \lambda \partial f(h_\lambda^f(x)) \Leftrightarrow h_\lambda^f(x) \in (\nabla h + \lambda \partial f)^{-1} \nabla h(x)$. f is convex and h is strictly convex, $\nabla h + \lambda \partial f$ is then a strictly monotone operator. $(\nabla h + \lambda \partial f)^{-1}$ is then a univocale operator and; $h_\lambda^f(x) = (\nabla h + \lambda \partial f)^{-1}(\nabla h(x))$. □

**Remark 5.11.** *Replacing $h$ by $h_0$ at the Proposition 5.10, we obtain the result of the Proposition 2.2.*

**Proposition 5.12.** *We suppose that $h$ and $f$ verify the conditions of Corollary 5.8.*

*(1). If Argmin $f \neq \emptyset$ then, for all $x^* \in$ Argmin $f$, for all $x \in S$, we get:*

$$D_h(x^*, h_\lambda^f(x)) + D_h(h_\lambda^f(x), x) \leq D_h(x^*, x) \tag{4}$$

*(2). If $\inf f > -\infty$ then, for all $\varepsilon > 0$, for all $x^*$ such as $0 \in \partial_\varepsilon f(x^*)$, for all $x \in S$, we get:*

$$D_h(x^*, h_\lambda^f(x)) + D_h(h_\lambda^f(x), x) \leq D_h(x^*, x) + \varepsilon. \tag{5}$$

*Proof.*

(1). From (1) of the Proposition 5.10,

$$\frac{\nabla h(x) - \nabla h(h_\lambda^f(x))}{\lambda} \in \partial f(h_\lambda^f(x)) \quad and \quad 0 \in \partial f(x^*)$$

$\partial f$ is the monotone operator, so;

$$\langle \nabla h(x) - \nabla h(h_\lambda^f(x)), x^* - h_\lambda^f(x) \rangle \leq 0, \tag{6}$$

and by vertue of Lemma 3.10, we get:

$$D_h(x^*, h_\lambda^f(x)) + D_h(h_\lambda^f(x), x) \leq D_h(x^*, x).$$

(2). From a similar way to (1),

$$\langle \nabla h(x) - \nabla h(h_\lambda^f(x)), x^* - h_\lambda^f(x) \rangle \leq \varepsilon,$$

whence the inequality (5). □

**Corollary 5.13.** *We suppose that $h$ anf $f$ verify the conditions of Corollary 5.8. If $\inf(f) > -\infty$ and $h$ verifies $H_3$, then $h_\lambda^f : S \longrightarrow S$ is a continued application.*

*Proof.* Let $x \in S$ and $x^n \in S$ such as $x^n \longrightarrow x$, let's show that $h_\lambda^f(x^n) \longrightarrow h_\lambda^f(x)$. Let $x^*$ such as $0 \in \partial_\varepsilon f(x^*)$, by replacing x by $x^n$ on (5), we obtain:

$$D_h(x^*, h_\lambda^f(x^n)) + D_h(h_\lambda^f(x^n), x^n) \leq D_h(x^*, x^n) + \varepsilon.$$

We get then

$$D_h(x^*, h_\lambda^f(x^n)) \leq D_h(x^*, x^n) + \varepsilon.$$

As $x^n \longrightarrow x \in S$, $D_h(x^*, x^n) \longrightarrow D_h(x^*, x)$, so the previous inequality leads that the sequence $\{D_h(x^*, h_\lambda^f(x^n))\}$ is bounded. From $H_3$, we deduce that $\{h_\lambda^f(x^n)\}$ is bounded too. Let $\{h_\lambda^f(x^{n_i})\}_n$ a subsequence of $\{h_\lambda^f(x^n)\}_n$ such as $h_\lambda^f(x^{n_i}) \longrightarrow u$, we get,

$$f(h_\lambda^f(x^{n_i})) + \lambda^{-1} D_h(h_\lambda^f(x^{n_i}), x^{n_i}) \leq f(v) + \lambda^{-1} D_h(v, x^{n_i})$$

Passing by at the limit, $f(u) + \lambda^{-1} D_h(u, x) \leq f(v) + \lambda^{-1} D_h(v, x)$, which means

$$u = h_\lambda^f(x).$$

$h_\lambda^f(x)$ is unique, whence $h_\lambda^f(x^n) \longrightarrow h_\lambda^f(x)$ with an enhancement of conditions on h or f, we can establish the contradiction of the operator $h_\lambda^f$ for $\lambda$ large enough. □

**Proposition 5.14.** *We suppose that h and f verify the conditions of the Corollary 5.8. If furthermore h or f is strongly convex with parameter $\alpha$, we have then,*

$$\|h_\lambda^f o(\nabla h)^{-1}(x) - h_\lambda^f \circ (\nabla h)^{-1}(y)\| \leq \frac{1}{\alpha}\|x - y\|.$$

*Proof.* $h_\lambda^f = (\nabla h + \lambda\partial f)^{-1} \circ \nabla h \Rightarrow h_\lambda^f \circ (\nabla h)^{-1} = (\nabla h + \lambda\partial f)^{-1}$. $\nabla h + \lambda\partial f$ is a strongly convex operator, from [10] we have the inequality. □

**Proposition 5.15.** *We suppose that h and f verify the conditions of the Corollary 5.8. furthermore, h verifies:*

$$\exists \beta >, \forall\ x, y \in S, \|\nabla h(x) - \nabla h(y)\| \leq \beta\|x - y\|.$$

*(a). If f is strongly convex with parameter $\alpha(\alpha > 0)$, then,*

$$\forall\ x, y \in S, \|h_\lambda^f(x) - h_\lambda^f(y)\| \leq \frac{\beta}{\alpha\lambda}\|x - y\|.$$

*If $\frac{\beta}{\alpha} \leq \lambda$, then $h_\lambda^f$ is a contraction.*

*(b). If h is strongly convex on S with parameter $\alpha(\alpha > 0)$, then:*

$$\forall\ x, y \in S, \|h_\lambda^f(x) - h_\lambda^f(y)\| \leq \frac{\beta}{\alpha}\|x - y\|.$$

*If $\beta = \alpha$, then $h_\lambda^f$ is a contraction.*

*Proof.*

(a). We put $h_\lambda^f(x) = x^*$ and $h_\lambda^f(y) = y^*$. From the Proposition 5.10,

$$\frac{\nabla h(x) - \nabla h(x^*)}{\lambda} \in \partial f(x^*),$$
$$\frac{\nabla h(y) - \nabla h(y^*)}{\lambda} \in \partial f(y^*).$$

f is strongly convex with parameter $\alpha$, then

$$\langle \nabla h(x) - \nabla h(x^*) - \nabla h(y) + \nabla h(y^*), x^* - y^* \rangle \geq \alpha\lambda\|x^* - y^*\|^2$$

Which equals at;

$$\begin{aligned}
&\langle \nabla h(x) - \nabla h(y), x^* - y^* \rangle \geq \alpha\lambda\|x^* - y^*\|^2 + \langle \nabla h(x^*) - \nabla h(y^*), x^* - y^* \rangle \\
\Rightarrow &\langle \nabla h(x) - \nabla h(y), x^* - y^* \rangle \geq \alpha\lambda\|x^* - y^*\|^2 \\
\Rightarrow &\qquad \alpha\lambda\|x^* - y^*\|^2 \leq \|\nabla h(x) - \nabla h(y)\|.\|x^* - y^*\| \\
\Rightarrow &\qquad \|x^* - y^*\|^2 \leq \frac{\beta}{\lambda\alpha}\|x^* - y^*\|\|x - y\| \\
\Rightarrow &\qquad \|x^* - y^*\| \leq \frac{\beta}{\lambda\alpha}\|x - y\|
\end{aligned}$$

If $\frac{\beta}{\alpha} \leq \lambda$ then $\frac{\beta}{\lambda\alpha} \leq 1$ and then $\|h_\lambda^f(x) - h_\lambda^f(y)\| \leq \|x - y\|$.

(b). $\partial f$ is a monoyone operator, then;

$$\begin{aligned} &\langle \nabla h(x) - \nabla h(x^*) - \nabla h(y) + \nabla h(y^*), x^* - y^* \rangle \geq 0 \\ \Rightarrow \quad & \langle \nabla h(x) - \nabla h(y), x^* - y^* \rangle \geq \langle \nabla h(x^*) - \nabla h(y^*), x^* - y^* \rangle \\ \Rightarrow \quad & \langle \nabla h(x) - \nabla h(y), x^* - y^* \rangle \geq \alpha \|x^* - y^*\|^2. \end{aligned}$$

As the same as previous, we conclude that : $\|x^* - y^*\| \leq \frac{\beta}{\lambda}\|x - y\|$. □

**Theorem 5.16.**

*(a). Let $h \in A(S)$ and $f \in \Gamma_0(R^p)$ such as $ri(dom\ f) \cap S \neq \emptyset$. Then: $\forall\ x \in S, \forall\ \lambda > 0$, we get :*

$$f_{h\lambda}(x) + (f^* \diamond (\lambda^{-1}h)^*)(\lambda^{-1}\nabla h(x)) = \lambda^{-1}(\langle x, \nabla h(x)\rangle - h(x)). \tag{7}$$

*Where $\diamond$ is inf concvolution.*

*(b). Let $h \in L(S), f \in \Gamma_0(R^p), x \in S$ and $\lambda > 0$ verifying the theorem conditions 5.7, we get :*

*(i). $\inf_u\{f(u) + \lambda^{-1}D_h(u,x)\} + \inf_v\{f^*(v) + \lambda^{-1}h^*(\nabla h(x) - \lambda v)\} = \lambda^{-1}h^*(\nabla h(x))$. Those two infima are finite and achieved respectively in $u^*$ and $v^*$ such as :*

$$\nabla h(x) = \nabla h(u^*) + \lambda v^*. \tag{8}$$

*(ii). If dom $h^* = Im\nabla h$, then the second infimum is achieved in a unique point $v^*$ verifying (8).*

*Proof.*

(a).

$$\begin{aligned} f_{h\lambda}(x) &= \inf_u\{f(u) + \lambda^{-1}D_h(u,x)\} \\ &= -\sup\{\langle \lambda^{-1}\nabla h(x), u\rangle - (f(u) + \lambda^{-1}h(u))\} + \lambda^{-1}(\langle x, \nabla h(x)\rangle - h(x)) \end{aligned}$$

$\Rightarrow f_{h\lambda}(x) + (f + \lambda^{-1}h)^*(\lambda^{-1}\nabla h(x)) = \lambda^{-1}(\langle x, \nabla h(x)\rangle - h(x))$.

As $ri(dom\ f) \bigcap S \neq \emptyset$, from the Theorem 16, 4 [9], we have

$$(f + \lambda^{-1}h)^*(\lambda^{-1}\nabla h(x)) = (f^* \diamond (\lambda^{-1}h)^*)(\lambda^{-1}\nabla h(x)),$$

which leads the equality (7)

(b). (i). By application of the Proposition 5.10, we get :

$$\begin{aligned} u^* = h_\lambda^f(x) &\Leftrightarrow \frac{\nabla h(x) - \nabla h(u^*)}{\lambda} \in \partial f(u^*) \\ &\Leftrightarrow u^* \in \partial f^*\left(\frac{\nabla h(x) - \nabla h(u^*)}{\lambda}\right) \\ &\Leftrightarrow 0 \in \partial f^*\left(\frac{\nabla h(x) - \nabla h(u^*)}{\lambda}\right) - \nabla h^*(\nabla h(u^*)) \end{aligned}$$

($\nabla h^* = (\nabla h)^{-1}$ because $h \in L(S)$). Let $v^*$ such as : $\nabla h(x) = \nabla h(u^*) + \lambda v^*$,

$$\begin{aligned} u^* = h_\lambda^f(x) &\Rightarrow 0 \in \partial f^*(v^*) - \nabla h^*(\nabla h(x) - \lambda v^*) \\ &\Rightarrow v^* \in\ Arg\min_v\{f^*(v) + \lambda^{-1}h^*(\nabla h(x) - \lambda v)\} \end{aligned}$$

which establishes (8).

(ii). Let $v^*$ such as :

$$v^* \in \; Arg\min_v \{f^*(v) + \lambda^{-1} h^*(\nabla h(x) - \lambda v)\}.$$

We deduct that

$$0 \in \partial f^*(v^*) - \nabla h^*(\nabla h(x) - \lambda v^*).$$

Since $\nabla h(x) - \lambda v^* \in dom\ h^* = Im\nabla h$, it exists $u^* \in S$ such as

$$\nabla h(x) - \lambda v^* = \nabla h(u^*)$$

We have then :

$$\begin{aligned} & 0 \in \partial f^* \left( \frac{\nabla h(x) - \nabla h(u^*)}{\lambda} \right) - \nabla h^*(\nabla h(u^*)) \\ \Rightarrow\ & u^* \in \partial f^* \left( \frac{\nabla h(x) - \nabla h(u^*)}{\lambda} \right) \\ \Rightarrow\ & u^* = h^f_\lambda(x). \end{aligned}$$

Which result the uniqueness of $v^*$.

□

Until now, we study the properties of the entropic approximation $f_{h\lambda}$.

**Proposition 5.17.**

*(1). If $h \in A(S)$, then; $\forall\ \lambda \geq \mu > 0, \forall\ x \in S, f_{h\lambda}(x) \leq f_{h\mu}(x) \leq f(x)$.*

*(2). If $h$ and $f$ verify the hypotheses of Corollary 5.8, then: $\inf_S f_{h\lambda} = \inf_S f$.*

*Proof.*

(1). $\forall\ y \in \overline{S}, \forall\ x \in S, D_h(y,x) \geq 0$. Therefore

$$\begin{aligned} \mu \leq \lambda \Rightarrow & \lambda^{-1} D_h(y,x) \leq \mu^{-1} D_h(y,x), \forall\ y \in \overline{S}, \forall\ x \in S \\ \Rightarrow \quad & f_{h\lambda}(x) \leq f_{h\mu}(x). \end{aligned}$$

Moreover:

$$f_{h\mu}(x) \leq f(y) + \mu^{-1} D_h(y,x), \forall\ y \in \overline{S}.$$

Replacing y by x, we obtain

$$f_{h\mu}(x) \leq f(x), \forall\ x \in S.$$

(2).
$$\begin{aligned} \inf_{x\in S} f_{h\lambda}(x) &= \inf_{x\in S} \{ \inf_{u\in \overline{S}} (f(u) + \lambda^{-1} D_h(u,x))\} \\ &= \inf_{x\in S} \{ \inf_{u\in S} (f(u) + \lambda^{-1} D_h(u,x))\} \\ &= \inf_{u\in S} \inf_{x\in S} \{(f(u) + \lambda^{-1} D_h(u,x)\} \\ &= \inf_{u\in S} \{(f(u) + \inf_{x\in S} \lambda^{-1} D_h(u,x)\}. \end{aligned}$$

$\inf_{x\in S} \lambda^{-1} D_h(u,x) = 0$ for $u \in S$, whence $\inf_S f_{h\lambda} = \inf_S f$. □

**Proposition 5.18.** *We suppose that h and f verify the hypotheses of Corollary 5.13. If h is twice Continuously Differentiable on S and $D_h(.,.)$ and jointly convex; then $f_{h\lambda}$ is continually differentiable, convex and such as:*

$$\forall\, x \in S, \nabla f_{h\lambda}(x) = \lambda^{-1} H(x)(x - h^f_\lambda(x)) \text{ where } H = \nabla^2 h$$

*Proof.* $D_h(.,.)$ is jointly convex and f is convex , $f_{h\lambda}$ is then convex. Lets show that:

$$\forall\, x \in S, \partial f_{h\lambda}(x) \subset \partial(\lambda^{-1} D_h(x^*, .))(x) \text{ when } x^* = h^f_\lambda(x). \tag{9}$$

$\forall\, x \in S$, $f_{h\lambda}(x) = f(x^*) + \lambda^{-1} D_h(x^*, x)$. Let $y \in S$, we get:

$$f_{h\lambda}(y) = f(x^* + \lambda^{-1} D_h(x^*, y).$$

Let $u \in \partial f_{h\lambda}(x)$, we have:

$$f_{h\lambda}(y) \geq f_{h\lambda}(x) + \langle u, y - x\rangle$$

$x^* = h^f_\lambda(x) \;\in dom\ f \Rightarrow \lambda^{-1} D_h(x^*, y) \geq \lambda^{-1} D_h(x^*, x) + \langle u, y - x\rangle$ which means: $\lambda u \in \partial(D_h(x^*, .))(x)$, which shows (9). h is two times conditionally differentiable, therefore:

$$\begin{aligned}
\lambda u &= \nabla(D_h(x^*, .))(x) \\
\lambda u &= -\nabla^2 h(x)(x^* - x) \\
u &= \lambda^{-1} H(x)(x - x^*).
\end{aligned}$$

Consequently, $\nabla f_{h\lambda}(x) = \lambda^{-1} H(x)(x - h^f_\lambda(x))$. ☐

**Proposition 5.19.** *We suppose that h and f verify the hypotheses of the Corollary 5.8. If furthermore,*

*(1). h is twice Continuously Differentiable on S and $D_h(.,.)$ is convex jointly,*

*(2). H is defined positive.*

*Then $Arg \min_S f = \; Arg \min_S f_{h\lambda}$.*

*Proof.* Let $u^* \in \; Arg \min_S f_{h\lambda}$.

$$\begin{aligned}
f_{h\lambda}(u^*) = \inf_S f_{h\lambda} \Leftrightarrow &\qquad 0 \in \partial f_{h\lambda}(u^*) \\
\Leftrightarrow &\qquad 0 = \nabla f_{h\lambda}(u^*) \\
\Leftrightarrow &\lambda^{-1} H(u^*)(u^* - h^f_\lambda(u^*)) = 0.
\end{aligned}$$

Since H is defined positive, we from then deduct that $u^* = h^f_\nabla(u^*)$. From the Proposition 5.10, we have:

$$u^* = h^f_\lambda(u^*) \Rightarrow 0 \in \partial f(u^*) \Rightarrow u^* \in \arg\min_S f.$$

We get then: $Arg \min_S f_{h\lambda} \subset \; Arg \min_S f$ reciprocally, let $x^*$ such that $f(x^*) = \inf_S f$.

$$f(x^*) = \inf_S f_{h\lambda} \leq f_{h\lambda}(x^*) \leq f(x^*).$$

Thus we have $f(x^*) = \inf_S f_{h\lambda} = f_{h\lambda}(x^*)$, which complete the demonstration. ☐

**Proposition 5.20.** *Let $f \in \Gamma_0(R^p)$ and $h \in A(S)$ verifying $H_3$ and $H_7$.*

*(a).* $\forall\ x \in S \cap dom f, \lim\limits_{\lambda \longrightarrow 0} prox^h_{\lambda f}(x) = x$

*(b).* $\forall\ x \in S, \lim\limits_{\lambda \longrightarrow 0} f_{h\lambda}(x) = f(x).$

*Proof.*

(a). From the Theorem 5.2, $prox^h_{\lambda f}(x) := x_\lambda \in \overline{S} \cap dom\ f$, we have: $f_{h\lambda}(x) = f(x_\lambda) + \lambda^{-1} D_h(x_\lambda, x)$, and $f_{h1}(x) \leq f(u) + D_h(u, x)$. Replacing u by $x_\lambda$ on the previous inequality, we deduce that: $f_{h1}(x) - D_h(x_\lambda, x) + \lambda^{-1} D_h(x_\lambda, x) \leq f_{h\lambda}(x)$ or

$$D_h(x_\lambda, x)(\lambda^{-1} - 1) \leq f_{h\lambda}(x) - f_{h1}(x)$$

$$D_h(x_\lambda, x)(1 - \lambda) \leq \lambda\, [f_{h\lambda}(x) - f_{h1}(x)]$$

For $0 < \lambda < 1$,

$$0 \leq D_h(x_\lambda, x) \leq \frac{\lambda}{1 - \lambda}\, [f(x) - f_{h1}(x)]$$

When $\lambda \longrightarrow 0, D_h(x_\lambda, x) \longrightarrow 0$. From $H_3$, the generalized sequence $\{x_\lambda\}_{\lambda \in I}$ is bounded. Let $x^*$ an adherence value of $\{x_\lambda\}_{\lambda \in I}$, it exists then a sub-sequence $\{x_{\alpha(\lambda)}\}$ such as $x_{\alpha(\lambda)} \longrightarrow x^*$. We get:

$$D_h(x_{\alpha(\lambda)}, x) = h(x_{\alpha(\lambda)}) - h(x) - \langle x - x_{\alpha(\lambda)}, \nabla h(x) \rangle,$$

Whence, by passage to the limit, $D_h(x^*, x) = 0$. That means that $x^* = x$ and therefore, $x_\lambda \longrightarrow x$ as $\lambda \longrightarrow 0$.

(b). We get: $f(x_\lambda) \leq f_{h\lambda}(x) \leq f(x)$.

$$x_\lambda \longrightarrow x \text{ and } f s.c.i. \Rightarrow f(x) \leq \underline{lim} f(x_\lambda) \leq \underline{lim} f_{h\lambda}(x) \leq \overline{lim} f_{h\lambda}(x) \leq f(x) \Rightarrow \lim_{\lambda \longrightarrow 0} f_{h\lambda}(x) = f(x).$$

If h is strongly convex on the module 1, then the approximation of f by $f_{h\lambda}$ is better than by $f_\lambda$, as the following proposition shows. □

**Proposition 5.21.** *Let $h, h' \in A(S)$*

*(1). If $h - h'$ is a convex function, then:* $\forall\ x \in S, \forall\ \lambda > 0,\ f_{h'\lambda}(x) \leq f_{h\lambda}(x) \leq f(x).$

*(2). If h is strongly convex of module 1, then:* $\forall\ x \in S, \forall\ \lambda > 0,\ f_\lambda(x) \leq f_{h\lambda}(x) \leq f(x).$

*Proof.*

(1). Let $x \in S$ and $\lambda > 0$. By value of the Proposition 3.8, $\forall\ y \in \overline{S}, D_{h-h'}(y, x) = D_h(x, y) - D_{h'}(y, x)$. Since $h - h'$ is a convex function, we get: $\forall\ y \in \overline{S}, D_{h-h'}(y, x) \geq 0$. We deduct that $\forall\ y \in \overline{S}, D_{h'}(y, x) \leq D_h(y, x)$. From this inequality, we obtain $\forall\ y \in \overline{S}, f(y) + \lambda^{-1} D_{h'}(y, x) \leq f(y) + \lambda^{-1} D_h(y, x)$. We have then $f_{h'\lambda}(x) \leq f_{h\lambda}(x)$. According to the Proposition 5.17 (1), we deduct the wanted inequality.

(2). h is strongly convex on the module 1, means that h - $h_0$ is convex . Consequently, from (1), we have $f_{h_0\lambda}(x) \leq f_{h\lambda}(x) \leq f(x)$. $f_{h_0\lambda} = f_\lambda$, whence the result. □

# 6. Conclusion

Replacing h by $h_0$ in all the result developed previously, we find all the properties of regularity and approximation given by Moreau and Yosida in spite of the non-symmetry of $D_h(.,.)$ and the absence of the triangular inequality. These results make it easy to establish the convergence of the algorithmic type :

$$x^n := \arg\min_z \left\{ f(z) + \frac{1}{\lambda_n} D_h(z, x^{n-1}) \right\}, \lambda_n > 0.$$

This sequence converges towards a minimum of f.

## References

[1] H. Attouch, *Variational convergence for functions and operator*, Pitman Advanced Publishing Program, (1984).

[2] H. Attouch and R. J. B. Wets, *Isometrie for the Legendre-Fencheltransform*, Trans .A.M.S., 296(1986), 33-60.

[3] M. A. Bahraoui, *Diagonal Bundle Methods for Convex Minimization, Theoritical Aspect*, Int. J. Math. And Appl., 6(4)(2018), 5157.

[4] H. Brezis, *Oprateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, Math. Studies, 5(1973).

[5] Y. Censor and A. Lent, *An iterative row action method for interval convex programming*, Journal of Optimization Theory and Applications, 34(3)(1981), 321-353.

[6] A. R. De Pierro and A. N. Iusem, *A relaxed version of Bregman's method for convex programming*, Journal of Optimization theory and Application, 5(3)(1986), 421-440.

[7] J. Eckstein, *Nonlinear proximal point algorithms using Bregman functions, with applications to convex programming*, Mathematics of Operations Research, 18(1(1993), 202-226.

[8] J. J. Moreau, *Proximit et dualit dans un espace Hilbertien*, Bull.Soc. Math. France, 93(1965), 273-299.

[9] R. T. Rockafellar, *Convex Analysis*, Princeton, New Jersey, Princeton University Press, (1970).

[10] R. T. Rockafellar, *Monotone operators and the proximal point algorithm*, SIAM J. Control and Optimization, 14(5)(1975), 877-898.

[11] M. Teboulle and Gong Chen, *Convergence analysis of a proximal-like minimization algorithm using Bregman function*, SIAM Journal on Optimization, 3(3)(1993), 538-543.

[12] M. Teboulle, *Entropic proximal mappings with applications to nonlinear programming*, Mathematics of Operations Research, 17(3)(1992), 670-690.