

Barrier operators and associated gradient-like dynamical systems for constrained minimization problems.

JÉRÔME BOLTE¹ AND MARC TEBoulLE²

Abstract. We study some continuous dynamical systems associated with constrained optimization problems. For that purpose, we introduce the concept of elliptic barrier operators and develop a unified framework to derive and analyze the associated class of gradient-like dynamical systems, called A-Driven Descent Method (A-DM). Prominent methods belonging to this class include several continuous descent methods studied earlier in the literature such as steepest descent method, continuous gradient projection methods, Newton type methods as well as continuous interior descent methods such as Lotka-Volterra type differential equations, and Riemannian gradient methods. Related discrete iterative methods such as proximal interior point algorithms based on Bregman functions and second order homogeneous kernels can also be recovered within our framework and allow for deriving some new and interesting dynamics. We prove global existence and strong viability results of the corresponding trajectories of (A-DM) for a smooth objective function. When the objective function is convex, we analyze the asymptotic behavior at infinity of the trajectory produced by the proposed class of dynamical systems (A-DM). In particular, we derive a general criterion ensuring the global convergence of the trajectory of (A-DM) to a minimizer of a convex function over a closed convex set. This result is then applied to several dynamics built upon specific elliptic barrier operators. Throughout the paper, our results are illustrated with many examples.

Key words: Dynamical systems, continuous gradient-like systems, elliptic barrier operators, Lotka-Volterra differential equations, asymptotic analysis, viability, Lyapunov functionals, explicit and implicit discrete schemes, interior proximal algorithms, global convergence, constrained convex minimization, Riemannian gradient methods.

1 Introduction

This paper proposes to study some continuous dynamical systems in relation with the constrained optimization problem

$$(\mathcal{P}) \quad \inf\{f(x) : x \in \overline{C}\},$$

where C is a nonempty *open* convex subset of \mathbb{R}^n , $n \geq 1$, $f : \mathbb{R}^n \mapsto \mathbb{R}$ is a convex function and \overline{C} denotes the closure of C .

Our first aim is to give a unified framework to smooth continuous interior descent methods studied earlier in the literature: steepest descent method, Lotka-Volterra type equations,

¹Département de Mathématiques, Case 51, ACSIOM-CNRS FRE 2311, Université Montpellier II, 34095 Montpellier, Cedex 5, France
email: bolte@math.univ-montp2.fr.

²School of Mathematical Sciences, Tel-Aviv University, Ramat-Aviv 69978, Israel
email: teboull@post.tau.ac.il

continuous Newton method, continuous gradient projection method. Another goal of this study is to enlighten the local geometric aspects of some discrete *implicit* dynamics related to (\mathcal{P}) (particularly proximal type algorithms), by associating them to some adequate vector fields. More precisely, we will also show that one of our continuous models can be cast as a specific Riemannian gradient method.

This has led us to introduce the following class of gradient-like dynamical systems

$$(A - DM) \quad \begin{cases} \dot{x}(t) + A_{x(t)} \nabla f(x(t)) = 0, \forall t \geq 0, \\ x(0) \in C, \end{cases}$$

with

$$A : \begin{cases} C \times \mathbb{R}^n & \mapsto \mathbb{R}^n \\ (x, v) & \mapsto A_x v. \end{cases} \quad (1.1)$$

The notation $(A - DM)$ stands for A -driven descent method. To make $(A - DM)$ an interior descent method, we introduce a class of mappings of the type (1.1) called *elliptic barrier operators*. This is an alternative approach to the classical barrier methods, see for instance Auslender-Cominetti-Haddou [9], since the penalization does not act on the objective function f but on its gradient. Roughly speaking this implies two major requirements on the map A :

- the mapping $x \in C \rightarrow A_x \nabla f(x)$ must preserve the local optimality information given by $\nabla f(\cdot)$,
- the operator A has to vanish on $\{(x, -\nu), x \in \overline{C}, \nu \in N_{\overline{C}}(x)\}$, where $N_{\overline{C}}(x)$ is the normal cone to \overline{C} at $x \in \overline{C}$.

In the next section, a formal definition and the basic properties of elliptic barrier operators are given. The relevance of this notion is first illustrated by the general properties of $(A - DM)$ systems. We prove existence and viability results. If ∇f is locally Lipschitz continuous, then the trajectories of $(A - DM)$ are defined for all $t \geq 0$, and remain in C . Let us emphasize the fact that, unlike in Nagumo-type theorems used in viability theory (Aubin-Cellina [7]), the trajectories never encounter the boundary of C , and thus making $(A - DM)$ an *interior* method.

In Section 3, we propose a general and unifying framework to generate in a systematic way elliptic barrier operators. This is achieved by developing an abstract setting, with the help of proximal-like maps involving appropriately defined distance-like functions. Given a convenient distance-like function $d : \mathbb{R}^n \times C \mapsto \mathbb{R} \cup \{+\infty\}$ closed, proper, and convex with respect of its first variable, we introduce the following class of mappings

$$A_x^d v = x - \arg \min \{ \langle u, v \rangle + d(u, x) \mid u \in \mathbb{R}^n \}, \quad (x, v) \in C \times \mathbb{R}^n \quad (1.2)$$

where $\langle \cdot, \cdot \rangle$ stands for the Euclidean inner product of \mathbb{R}^n . Besides the fact that slight assumptions on d allow to make A^d an elliptic barrier operator, the associated A^d -driven descent method can be seen as another step towards a unified approach to both continuous

and discrete gradient-like dynamics. Indeed, one of the main fact underlying the introduction of d operator is that $(A^d - DM)$ systems can be reformulated as the following differential inclusion

$$\partial_1 d(\dot{x}(t) + x(t), x(t)) + \nabla f(x(t)) \ni 0, t \geq 0 \quad (1.3)$$

where for each $t \geq 0$, $\partial_1 d(\cdot, x(t))$ denotes the subdifferential of $d(\cdot, x(t))$.

This structure is at the heart of the so called proximal-like methods, (see the examples below),

$$\partial_1 d(x^{k+1}, x^k) + \nabla f(x^{k+1}) \ni 0, x^0 \in C, k \geq 0. \quad (1.4)$$

For instance, with $d(u, x) = 2^{-1}|u - x|^2$, where $|\cdot|$ denotes the Euclidean norm, the inclusion (1.4) reduces to the proximal minimization algorithm, see e.g., Martinet [32], Lemaire [29], and references therein. Then, according to the classical idea that consists in interpreting an iterative scheme as some discretization of a continuous dynamical system, the differential inclusion (1.3), i.e. $(A^d - DM)$, can be proposed as a continuous model for the proximal method (1.4). This opens new perspectives of crossed investigations and from that viewpoint it is important to realize that the interplay between discrete and continuous dynamical systems goes far beyond the fruitful finite-time approximation aspects. For instance in Alvarez-Attouch [2], Antipin [4] crucial features of the asymptotic analysis appear also as closely related matters.

To give the reader a concrete idea on the type of operators A that will emerge in this study, we outline below some specific models.

(a) *The gradient projection operator*

The first natural example is given by

$$A^P : \begin{cases} C \times \mathbb{R}^n & \mapsto \mathbb{R}^n \\ (x, v) & \mapsto x - P_{\overline{C}}(x - v), \end{cases} \quad (1.5)$$

where $P_{\overline{C}}$ is the orthogonal projection on \overline{C} . $(A^P - DM)$ is the continuous gradient projection method as introduced in [4],

$$\dot{x}(t) + x(t) - P_{\overline{C}}[x(t) - \nabla f(x(t))] = 0, x(0) \in C, \forall t \geq 0. \quad (1.6)$$

The operator A^P ruling (1.6) can be recovered thanks to (1.2) with a distance-like function of the type $d : \mathbb{R}^n \times C \ni (u, x) \mapsto \frac{1}{2}|u - x|^2 + \delta_{\overline{C}}(u)$ where $\delta_{\overline{C}}$ is the indicator function of \overline{C} . Let us emphasize the fact that the trajectory of the continuous system (1.6) is interior, which is not the case for the following well known explicit discretization

$$x^{k+1} = P_{\overline{C}}[x^k - \mu_k \nabla f(x^k)], x^0 \in C, \mu_k > 0$$

see e.g., [30], [18].

(b) *The Bregman operators*

The Bregman proximal method (*BPM*) is obtained by replacing the quadratic kernel in the proximal minimization algorithm by a distance-like based on a Bregman function $h : \overline{C} \rightarrow \mathbb{R}$. Defining

$$\forall (x, y) \in \overline{C} \times C, \quad D_h(x, y) = h(x) - h(y) - \langle \nabla h(y), x - y \rangle, \quad (1.7)$$

it leads to the scheme

$$(BPM) \quad x^{k+1} \in \arg \min \left\{ f(x) + c_k D_h(x, x_k) \mid x \in \overline{C} \right\}, \quad c_k > 0, \quad x^0 \in \overline{C}.$$

(*BPM*) has been studied and generalized from many viewpoints, see for instance Censor-Zenios [16], Chen-Teboulle [17], Eckstein [19], Kiwiel [26], Teboulle [36]. One of the corresponding continuous model that is proposed here is given by barrier operators A^{q_h} of the type

$$A^{q_h} : \begin{cases} C \times \mathbb{R}^n & \mapsto \mathbb{R}^n \\ (x, v) & \mapsto \nabla^2 h(x)^{-1} v, \end{cases}$$

where $\nabla^2 h(x)$ is the Hessian of some convenient Bregman function with zone C and with $q_h(u, x) = \langle \nabla^2 h(x)(u-x), u-x \rangle$, $(u, x) \in \mathbb{R}^n \times C$. The A^{q_h} -driven descent method –actually a Riemannian gradient method– is then given by

$$(A^{q_h} - DM) \quad \dot{x}(t) + \nabla^2 h(x(t))^{-1} \nabla f(x(t)) = 0, \quad x(0) \in C.$$

Besides its links with (*BPM*) developed in Section 4, the latter system allows to recover several dynamics. With $h_1(x) = \frac{\alpha}{2}|x|^2 + \beta \sum_{i=1..N} x_i \log x_i$, $\alpha, \beta > 0$ on $C = \mathbb{R}_{++}^n := \{x \in \mathbb{R}^n, x_i > 0\}$, we obtain the regularized Lotka-Volterra equation recently proposed, from a completely different viewpoint, in Attouch-Teboulle [6]:

$$(A^{q_{h_1}} - DM) \quad \dot{x}_i(t) + \frac{x_i(t)}{\beta + \alpha x_i(t)} \frac{\partial f}{\partial x_i}(x(t)) = 0, \quad \forall i = 1 \dots, n, \quad x(0) \in \mathbb{R}_{++}^n, \quad (1.8)$$

where f is to be optimized on \mathbb{R}_+^n .

If $h(x) = \frac{\alpha}{2}|x|^2$ and $C = \mathbb{R}^n$, ($A^{q_h} - DM$) is the classical continuous steepest descent method $\dot{x}(t) + \nabla f(x(t)) = 0$, $t \geq 0$, see Brézis [13].

For $h(x) = f(x)$ and $C = \mathbb{R}^n$, we obtain the continuous Newton descent method, studied in Alvarez-Pérez [3], see also [7]

$$(A^{q_f} - DM) \quad \dot{x}(t) + \nabla^2 f(x(t))^{-1} \nabla f(x(t)) = 0. \quad (1.9)$$

Another surprising fact of this dynamics is to be physically meaningful in infinite-dimensional spaces. Naturally those problems are out of the scope of the present paper, but the reader interested by thermodynamical evolution equations of the form ($A^{q_h} - DM$) is referred to Kenmochi-Pawlow [25] and references therein.

(c) *Barrier operators based on interior methods for the positive orthant*

Another line of research pursued by Auslender, Teboulle, and Ben Tiba [8] concerning proximal interior methods is based on the distance-like function

$$\forall (x, y) \in (\mathbb{R}_{++}^n)^2 \quad d_\varphi(x, y) = \sum_{i=1}^n y_i^2 \varphi\left(\frac{x_i}{y_i}\right), \quad (1.10)$$

where $\varphi : \mathbb{R}_{++} \rightarrow \mathbb{R}$ is some relevant convex function.

The associated iterative proximal interior method is given by

$$(RIPM) \quad x^{k+1} \in \arg \min \{f(x) + c_k d_\varphi(x, x_k) \mid x \in \mathbb{R}_+^n\}, \quad c_k > 0, \quad x^0 \in \mathbb{R}_{++}^n$$

where (*RIPM*) stands for regularized interior proximal method. Like (*BPM*) this algorithm can be applied to a minimize a general closed convex function. However, it enjoys stronger convergence properties, particularly when applied to a dual problem of a convex program, see [8], for further details and results.

Our continuous approach to (*RIPM*) is obtained by considering barrier operators of the form

$$A^{d_\varphi} : \begin{cases} \mathbb{R}_{++}^n \times \mathbb{R}^n & \mapsto \mathbb{R}^n \\ (x, v) & \mapsto \left(x_i - x_i(\varphi^*)'(x_i^{-1}v_i)\right)_{i=1, \dots, n} \end{cases}$$

where φ^* is the Legendre-Fenchel conjugate of the function φ used in (*RIPM*).

All these continuous models are derived and analyzed in Section 4. Section 5 of this paper is devoted to the asymptotic analysis of (*A-DM*) in the convex case. We derive a general criterion ensuring the global convergence of the trajectories of (*A-DM*) to a minimizer of f over \bar{C} . We then apply this general result to the dynamics built upon A^P , A^{q_h} , and A^{d_φ} . The proof relies on the existence of Lyapounov functionals measuring a sort of distance between the state variable and the set of equilibria. This approach is inspired at the same time by Opial's lemma [34] and the techniques used in monotone optimization algorithms. We also prove a general localisation result for the limit point of the trajectories produced by (*A-DM*), which extends results of the same type obtained recently in [6], and in [29] for the classical continuous gradient descent scheme. Throughout this paper we give many examples exhibiting some explicit and new systems of the type (*A-DM*). For instance with $C = \mathbb{R}_{++}^n$, one obtains the systems,

$$(A^{q_h} - DM) \quad \dot{x}_i(t) + \frac{2x_i(t)^{3/2}}{x_i(t)^{3/2} + 1} \frac{\partial f}{\partial x_i}(x(t)) = 0, \quad x_i(0) > 0, \quad \forall i \in \{1, \dots, n\}.$$

or

$$(A^{d_\varphi} - DM) \quad \dot{x}_i(t) + x_i(t) + \frac{1}{2} \frac{\partial f}{\partial x_i}(x(t)) - \sqrt{\frac{1}{4} \frac{\partial f}{\partial x_i}(x(t))^2 + x_i(t)^2} = 0,$$

with $i \in \{1, \dots, n\}$, $t \geq 0$ and $x(0) \in \mathbb{R}_{++}^n$. The first equation is given by the Bregman function $h(s) = s^2/4 - 2\sqrt{s}$, $s \geq 0$ while the second one corresponds to a continuous model

of the logarithmic-quadratic method [8], obtained with the choice $\varphi(s) = 1/2(s - 1)^2 - \log s + s - 1$, $s > 0$.

NOTATIONS. Our notations are fairly standard. The Euclidean space \mathbb{R}^n is equipped with the scalar product $\langle \cdot, \cdot \rangle$; the related norm is denoted $|\cdot|$. The boundary of C is denoted $\text{bd } C$. $N_{\overline{C}}(x)$ and $T_{\overline{C}}(x)$ denote respectively the normal cone and the tangent cone of \overline{C} at $x \in \overline{C}$. We recall that $N_{\overline{C}}(x) = \{v \in \mathbb{R}^n \mid \langle v, z - x \rangle \leq 0, \forall z \in \overline{C}\} = \{v \in \mathbb{R}^n \mid \forall u \in T_{\overline{C}}(x), \langle v, u \rangle \leq 0\}$. If $\phi : \mathbb{R}^p \rightarrow \mathbb{R} \cup \{+\infty\}$, $p \geq 1$ is a closed proper convex function, its domain is defined by $\text{dom } \phi = \{x \in \mathbb{R}^p \mid \phi(x) < +\infty\}$ and its Legendre-Fenchel conjugate, $y \in \mathbb{R}^p \rightarrow \sup \{\langle y, x \rangle - \phi(x) \mid x \in \mathbb{R}^p\}$, is denoted ϕ^* . If S is a closed convex subset of \mathbb{R}^n , the set of minimizers of ϕ on S is denoted $\arg \min_S \phi$. The indicator function of \overline{C} is denoted by $\delta_{\overline{C}}$. Other notations and definitions not explicitly stated here can be found in the classical book of Rockafellar [35].

2 Elliptic barrier operators and viability results

In this section, the definition and the first properties of elliptic barrier operators are introduced. Then, in view of constrained minimization, we study the corresponding A -driven descent methods, proving in particular that the obtained trajectories $\{x(t)\}$ are *interior* and defined for any $t \in [0, +\infty)$.

2.1 Elliptic barrier operators: Definition and Properties

Definition 2.1 $A : C \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ is an elliptic barrier operator on C if it satisfies:

- (r1) A is Lipschitz continuous on every compact subset of $C \times \mathbb{R}^n$.
- (r2) There exists $\alpha > 0$, such that for every $(x, v) \in C \times \mathbb{R}^n$, $\langle A_x v, v \rangle \geq \alpha |A_x v|^2$.
- (r3) For all $x \in C$, $A_x v = 0$ implies $v = 0$.
- (v) $\forall b \in \text{bd } C, \forall \nu \in N_{\overline{C}}(b), \forall M > 0, \exists \epsilon, K > 0$ such that $|x - b| < \epsilon, x \in C, |v| \leq M$ implies

$$\langle -A_x v, \nu \rangle \leq K \langle b - x, \nu \rangle. \quad (2.1)$$

This definition is motivated by the study of $(A - DM)$ systems. The *regularity* assumption (r1) naturally meets the conditions of the Cauchy-Lipschitz theorem. The *ellipticity* condition (r2) and the *non degeneracy* assumption (r3) allow to obtain a proper descent method. An important consequence of (r2) is that the term $1/\alpha$ can be seen as an upper bound for the gradient stepsize in $(A - DM)$. Indeed, it follows readily from (r2) that $|A_x v| \leq \alpha^{-1} |v|$, and therefore a trajectory $x(\cdot)$ of $(A - DM)$ satisfies

$$|\dot{x}(t)| \leq \alpha^{-1} |\nabla f(x(t))|,$$

whenever $x(t)$ is defined and belongs to C .

The *normal boundary* property (v) is required to control the outwards normal impulses near the boundary of C , making the trajectories of $(A - DM)$ strongly viable, i.e., $x(t) \in C, t \geq 0$. The choice of the term $\langle b - x, \nu \rangle$ in (2.1) has also a regularizing effect. Indeed, as it will be

proved in Theorem 2.1 below (see also Remark 2.1 (b)), it contributes to the fact that the trajectories of $(A - DM)$ are defined on $[0, +\infty)$.

Remark 2.1 (a) A natural extension of Definition 2.1 can be obtained by replacing assumptions (r2) and (r3) respectively by

$$\begin{aligned} (r2)' & \text{ For every } (x, v) \in C \times \mathbb{R}^n, v \neq 0 \langle A_x v, v \rangle > 0, \\ (r3)' & \text{ For all } x \in C, v = 0 \text{ implies } A_x v = 0. \end{aligned}$$

Observing that (r2)' and (r3)' imply (r3) it follows that an elliptic barrier operator satisfies this new definition. This widened concept opens new perspectives but also raises some difficulties in the study of $(A - DM)$: finite-time solutions, loss of regularity (see Theorem 2.1 in the elliptic case), no upper bound for the gradient step-sizes, etc... The study of such a class of mappings will not be carried out in the present paper, but appears as an interesting matter for future research.

(b) If the left term in (2.1) is replaced for instance by $\langle b - x, \nu \rangle^{1-\theta}$, $\theta \in (0, 1)$ the well-posedness of $(A - DM)$ may fail: take for instance $A : (x, v) \in \mathbb{R}_+ \times \mathbb{R} \rightarrow x^{1-\theta} \cdot v$, $\theta \in (0, 1)$, $f(x) = x + 1$ and observe that the maximal solutions of $(A - DM)$ are not defined on $[0, +\infty)$.

In what follows it is of interest to strengthen (r1) by assuming the additional hypothesis,

$$(r4) \text{ } A \text{ is continuous on } \overline{C} \times \mathbb{R}^n.$$

The following result shows that an elliptic barrier operator on C can be continuously extended to

$$C \times \mathbb{R}^n \cup \{(x, v) | x \in \text{bd } C, v \in -N_{\overline{C}}(x)\},$$

by setting $A_x v = 0$, if $x \in \text{bd } C$, $v \in -N_{\overline{C}}(x)$.

Proposition 2.1 *Let $A : C \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ be an elliptic barrier operator. Assume that $(x^k, v^k), k \in N$ is a sequence in $C \times \mathbb{R}^n$ such that $x^k \rightarrow x \in \overline{C}$ and $v^k \rightarrow v \in -N_{\overline{C}}(x)$ as $k \rightarrow +\infty$. Then*

(i) $A_{x^k} v^k \rightarrow 0$ as $k \rightarrow +\infty$.

(ii) In addition, if A satisfies (r4) then for all $x \in \overline{C}$ one has,

$$A_x^{-1}(\{0\}) \supset -N_{\overline{C}}(x). \quad (2.2)$$

Proof. If $x \in C$ the conclusion follows from (r1) and (r3). Else $x \in \text{bd } C$. (r2) and the Cauchy Schwarz inequality yield $|A_{x^k} v^k| \cdot |v^k| \geq \alpha |A_{x^k} v^k|^2$, for all $k \in N$ and some $\alpha > 0$. Since the sequence $v^k, k \in N$ is bounded so is $A_{x^k} v^k, k \in N$. From (v) it follows that for k large enough $\langle -A_{x^k} v^k, -v \rangle \leq K \langle x - x^k, -v \rangle$ and therefore

$$\limsup_{k \rightarrow +\infty} \langle A_{x^k} v^k, v \rangle \leq 0. \quad (2.3)$$

On the other hand we have

$$\langle A_{x^k} v^k, v \rangle = \langle A_{x^k} v^k, v - v^k \rangle + \langle A_{x^k} v^k, v^k \rangle, \forall k \in N,$$

and since $A_{x^k} v^k, k \in N$ is bounded we obtain

$$\liminf_{k \rightarrow +\infty} \langle A_{x^k} v^k, v \rangle = \liminf_{k \rightarrow +\infty} \langle A_{x^k} v^k, v^k \rangle \geq 0. \quad (2.4)$$

From (2.3) and (2.4), we deduce that $\lim_{k \rightarrow +\infty} \langle A_{x^k} v^k, v \rangle = \liminf_{k \rightarrow +\infty} \langle A_{x^k} v^k, v^k \rangle = 0$, and thus by (r2), $\lim_{k \rightarrow +\infty} |A_{x^k} v^k|^2 = 0$. \square

Remark 2.2 For simplicity, assume that f is convex, with $\arg \min_{\overline{C}} f \neq \emptyset$ and that A satisfies (r4). Subdifferential calculus, (see e.g., [35]) allows to associate to (\mathcal{P}) the following variational characterization

$$x^* \text{ solves } (\mathcal{P}) \text{ iff } \nabla f(x^*) + N_{\overline{C}}(x^*) = 0.$$

Using (2.2), we know that the solutions of (\mathcal{P}) are contained in the set of zeros of the gradient-like map $x \in \overline{C} \rightarrow A_x \nabla f(x)$. This is only a necessary condition for optimality and it can be written,

$$\text{if } x^* \text{ solves } (\mathcal{P}) \text{ then } A_{x^*} \nabla f(x^*) = 0. \quad (2.5)$$

The important point here, is to realize that our approach to optimization is given *throughout* $(A-DM)$ *dynamics* and thus x^* is obtained as a limit point of some descent method. Indeed, as we shall see, most of the systems and examples of Section 4 satisfy (2.2) with a *strict* inclusion, yet their orbits converge to a minimizer of f on \overline{C} , see Section 5.

We conclude these introductory notions by stating a useful criterion implying assumption (v) of Definition 2.1.

Lemma 2.1 *Let $A : C \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, $m > 0$, and $k : C \times \mathbb{R}^n \rightarrow [m, +\infty)$ be such that*

$$x - k(x, v) A_x v \in \overline{C}, \forall (x, v) \in C \times \mathbb{R}^n.$$

Then A satisfies (v).

Proof. It relies on the fact that $x - k(x, v) A_x v - b \in T_{\overline{C}}(b)$ for every (x, v) in $C \times \mathbb{R}^n$ and for every $b \in \overline{C}$. By definition we have for all $\nu \in N_{\overline{C}}(b)$, $\langle x - k(x, v) A_x v - b, \nu \rangle \leq 0$, and therefore

$$\langle -A_x v, \nu \rangle \leq \frac{1}{k(x, v)} \langle b - x, \nu \rangle \leq \frac{1}{m} \langle b - x, \nu \rangle. \square$$

2.2 Global existence and viability results.

From now on, the function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is C^1 and satisfies

- (\mathcal{H}_1) ∇f is Lipschitz continuous on bounded sets,
- (\mathcal{H}_2) $\inf_{\bar{C}} f > -\infty$.

Observe that for the moment the function f is not supposed to be convex.

Theorem 2.1 *Let A be an elliptic barrier operator. Then,*

- (i) *The system $(A - DM)$ admits a unique C^1 solution x defined on $[0, +\infty)$.*

Moreover,

- (ii) $\forall t \geq 0, x(t) \in C$.
- (iii) *The function $t \in [0, +\infty) \rightarrow f(x(t))$ is nonincreasing and has a limit as $t \rightarrow +\infty$,*
- (iv) $\dot{x} \in L^2(0, +\infty; \mathbb{R}^n)$.
- (v) *If A satisfies (r4) and $x(\cdot)$ is bounded then $\dot{x}(t) \rightarrow 0$ as $t \rightarrow 0$, and all limit point x^* of $x(\cdot)$ satisfies the weak optimality condition*

$$A_{x^*} \nabla f(x^*) = 0.$$

Proof. Fix $T > 0$ and consider the assertion $E(T)$:

“There exists a solution of $(A - DM)$ defined on $[0, T]$, and such that $x(t) \in C$ for all $t \in [0, T]$.”

Set $T_{max} := \sup\{T \mid E(T) \text{ is satisfied}\}$. From (r1), (\mathcal{H}_1) and the fact that $x(0) \in C$, it follows by Cauchy-Lipschitz Theorem that $T_{max} > 0$ and that the solution of $(A - DM)$ defined on $[0, T_{max})$ is unique.

Let us derive some a priori estimates. Let $T \in (0, T_{max})$, by the $(A - DM)$ system we have for all $t \in [0, T]$

$$\langle \dot{x}(t), \nabla f(x(t)) \rangle + \langle A_{x(t)} \nabla f(x(t)), \nabla f(x(t)) \rangle = 0,$$

and thus by (r2) and $(A - DM)$ again,

$$\frac{d}{dt} f(x(t)) + \alpha |\dot{x}(t)|^2 \leq 0. \quad (2.6)$$

Integrating over some interval $(0, t)$, with $t \leq T$ this gives

$$f(x(t)) - f(x(0)) + \alpha \int_0^t |\dot{x}|^2 \leq 0. \quad (2.7)$$

Note that if $T_{max} = +\infty$, (iii) and (iv) follow from (2.6), (2.7) and (\mathcal{H}_2). Let us argue by contradiction and assume that $T_{max} < +\infty$.

Using Cauchy Schwarz inequality and the fact that $\dot{x} \in L^2(0, T_{max}; \mathbb{R}^n)$, we obtain that x is a Cauchy net at T_{max} . Therefore x can be continuously extended by an application still denoted by x . Set $x(T_{max}) := b \in \bar{C}$.

By definition of T_{max} , b necessarily belongs to $\text{bd} C$. The function $t \in [0, T_{max}] \rightarrow \nabla f(x(t))$ is bounded by a positive constant M . Owing to the continuity of x and (v), there exists $t_0 \in (0, T_{max})$, $\epsilon > 0$, $K > 0$ and $\nu \in N_{\bar{C}}(b)$, $\nu \neq 0$, such that for all $t \in (t_0, T_{max})$

$$\langle -A_{x(t)} \nabla f(x(t)), \nu \rangle \leq K \langle b - x(t), \nu \rangle. \quad (2.8)$$

Let us project $(A - DM)$ on $\mathbb{R}\nu := \{\tau\nu \mid \tau \in \mathbb{R}, 0 \neq \nu \in N_{\overline{C}}(b)\}$, this gives for all $t \in (t_0, T_{max})$

$$\frac{d}{dt}\langle x(t), -\nu \rangle + \langle A_{x(t)}\nabla f(x(t)), -\nu \rangle = 0,$$

and using (2.8) we obtain

$$\frac{d}{dt}\langle b - x(t), \nu \rangle + K\langle b - x(t), \nu \rangle \geq 0.$$

Multiplying the above inequality by $\exp Kt$ and integrating over (t_0, T_{max}) it follows that

$$\langle b - x(T_{max}), \nu \rangle \geq \exp[-K(T_{max} - t_0)]\langle b - x(t_0), \nu \rangle.$$

Observe that by definition, $b = x(T_{max})$, hence to draw a contradiction from the latter we just have to prove that the second term of the inequality is positive. Indeed, $x(t_0) \in C$ which is open convex and $0 \neq \nu \in N_{\overline{C}}(b)$, thus there exists $\eta > 0$ such that $x(t_0) + \eta\nu \in C$, and a fortiori $x(t_0) + \eta\nu - b \in T_{\overline{C}}(b)$. This implies $\langle x(t_0) + \eta\nu - b, \nu \rangle \leq 0$ or equivalently, $\langle b - x(t_0), \nu \rangle \geq \eta|\nu|^2 > 0$, and (i) is proved.

Let us prove the last statement (v). From the boundedness property of x , along with (r4) and (\mathcal{H}_1) , it follows that \dot{x} is bounded and therefore x is a Lipschitz continuous map. The properties (r4), (\mathcal{H}_1) imply that $t \geq 0 \rightarrow A_{x(t)}\nabla f(x(t))$ is uniformly continuous and therefore so is $\dot{x}(\cdot)$. Combining this fact with (iv), it follows by a classical argument that $\dot{x}(t) \rightarrow 0$ as $t \rightarrow +\infty$. Using (r4), it ensues that a cluster point x^* of x satisfies $A_{x^*}\nabla f(x^*) = 0$. \square

3 A general abstract framework for dynamical systems with elliptic barrier operators

In this section, we propose with the help of proximal maps, a systematic and unifying way to generate elliptic barrier operators. We start with an informal motivation. Given a convenient distance-like function $d : \mathbb{R}^n \times C \mapsto \mathbb{R} \cup \{+\infty\}$, the idea is to realize the descent direction $-A_x\nabla f(x)$, $x \in C$ as a vector based on x and pointing on some proximal point $u^d(x, \nabla f(x))$.

Indeed, assume that d is convex with respect to its first variable, and for $x \in C$ define formally

$$u^d(x, \nabla f(x)) \in \arg \min \{\langle u, \nabla f(x) \rangle + d(u, x) \mid u \in \mathbb{R}^n\}. \quad (3.1)$$

In this definition the objective function has been replaced by its first order approximation at the point x , the constraints are supposed to be naturally taken into account by $d(\cdot, \cdot)$ and the descent direction obtained is $-A_x^d\nabla f(x) := u^d(x, \nabla f(x)) - x$. It is of interest to notice that this approach is akin to the following well known fixed point reformulation of the optimization problem (\mathcal{P}) :

$$x^* \text{ solves } (\mathcal{P}) \text{ iff } x^* \in \arg \min \{\langle u, \nabla f(x^*) \rangle \mid u \in \overline{C}\}, \quad (3.2)$$

whenever f is convex. From that viewpoint, the formal definition (3.1), may appear as a proximal regularization of some possibly ill-posed problem. On the other hand, the corresponding A^d -driven descent method can be written as a fixed point like dynamics

$$\dot{x}(t) + x(t) = u^d[x(t), \nabla f(x(t))], \quad x(0) \in C, \quad \forall t \geq 0. \quad (3.3)$$

The solution of (3.3) is then expected to provide asymptotically a solution of $x^* = u^d(x^*, \nabla f(x^*))$, and when it makes sense, this last problem corresponds to another formulation of (3.2).

As a first example, consider $d(u, x) = 1/2|u - x|^2 + \delta_{\overline{C}}(u)$, $(u, x) \in \mathbb{R}^n \times C$. The definition of u^d writes

$$\nabla f(x) + u^d(x, \nabla f(x)) - x + N_{\overline{C}}[u^d(x, \nabla f(x))] \ni 0,$$

which in turns is equivalent to

$$u^d(x, \nabla f(x)) \in (I + N_{\overline{C}})^{-1}(x - \nabla f(x)).$$

Recalling that $(I + N_{\overline{C}})^{-1} = P_{\overline{C}}$, the proximal point is thus given by $u^d(x, \nabla f(x)) = P_{\overline{C}}(x - \nabla f(x))$. This gives rise to the descent direction $-A_x^d \nabla f(x) = P_{\overline{C}}(x - \nabla f(x)) - x$, and the projected-gradient dynamics (1.6) is recovered. As mentioned in the above discussion note that the reformulation of (3.2) throughout $d(\cdot, \cdot)$, that is $x^* = u^d(x^*, \nabla f(x^*))$, leads to the fixed point problem $x^* = P_{\overline{C}}(x^* - \nabla f(x^*))$.

Let us now develop an abstract setting that shall be illustrated in the next section with various useful kernels $d(\cdot, \cdot)$.

Let $d_0 : \mathbb{R}^n \times C \mapsto \mathbb{R}_+ \cup \{+\infty\}$ be such that

(P1) d_0 is C^1 on $C \times C$,

(P2) $\nabla_1 d_0(u, u) = 0$ for all $u \in C$,

(P3) For every $x \in C$, the mapping $u \in \mathbb{R}^n \mapsto d_0(u, x)$ is a closed convex function.

In (P1), $\nabla_1 d_0(\cdot, u)$ is the gradient of $d(\cdot, u)$; (more generally its subdifferential is denoted by $\partial_1 d_0(\cdot, u)$). Note that, since C is nonempty, (P1) ensures that $u \in \mathbb{R}^n \mapsto d_0(u, x)$ is also proper.

Denote by \mathcal{D} the set of mappings $d : \mathbb{R}^n \times C \mapsto \mathbb{R}_+ \cup \{+\infty\}$ that can be written

$$d(u, x) = \frac{\alpha}{2}|u - x|^2 + d_0(u, x), \quad (3.4)$$

with $\alpha > 0$ and with d_0 satisfying (P1), (P2) and (P3).

Definition 3.1 *Let d be in \mathcal{D} . For all $(x, v) \in C \times \mathbb{R}^n$ set*

$$u^d(x, v) \in \arg \min \{\langle u, v \rangle + d(u, x) | u \in \mathbb{R}^n\} \quad (3.5)$$

and define A^d by

$$A_x^d v = x - u^d(x, v). \quad (3.6)$$

The following proposition justifies the second part (3.6) of this definition (u^d could be multivalued), and describes some of the properties of the operator A^d .

Proposition 3.1 *Let $d \in \mathcal{D}$.*

(i) *For each $x \in C$, the map $v \in \mathbb{R}^n \mapsto u^d(x, v)$ is a single valued α^{-1} -Lipschitz continuous map.*

(ii) *A^d satisfies (r2), (r3), and for each $x \in C$, $v \in \mathbb{R}^n \mapsto A_x^d v$ is Lipschitz continuous.*

(iii) *Moreover if d satisfies the property*

$$(p) \quad \forall x \in C, \text{ dom } d(\cdot, x) \subset \overline{C}$$

then A^d satisfies (v) of Definition 2.1.

Proof. Let $(x, v) \in C \times \mathbb{R}^n$. From (P3) and the fact that $\alpha > 0$ it follows that $u \in \mathbb{R}^n \mapsto \langle u, v \rangle + d(u, x)$ is strongly convex and has a nonempty bounded lower level set. This implies that $u^d(x, v)$ exists and is unique. Using (P1) and (P3), allows to write the optimality condition in (3.5) as

$$v + \partial_1 d(\cdot, x)(u^d(x, v)) \ni 0,$$

and therefore by uniqueness of $u^d(x, v)$, (recalling (cf. [35]) that for any closed proper convex function F , one has $(\partial F)^{-1} = \partial F^*$), it follows that

$$u^d(x, v) = \partial_1 d^*(\cdot, x)(-v). \quad (3.7)$$

Denoting by I the identity map of \mathbb{R}^n , we observe using the definition of $d \in \mathcal{D}$ that $\partial_1 d^*(\cdot, x)$ can also be written

$$(\alpha I + \partial_1 d_0(\cdot, x) - \alpha x)^{-1}$$

or equivalently as the composition,

$$(I + \alpha^{-1} \partial_1 d_0(\cdot, x) - x)^{-1} \circ \alpha^{-1} I.$$

By (P3), the operator $\alpha^{-1} \partial_1 d_0(\cdot, x) - x$ is maximal monotone and therefore by [13, Proposition 2.2], $(I + \frac{1}{\alpha} \partial_1 d_0(\cdot, x) - x)^{-1}$ is a contraction defined on \mathbb{R}^n . Recalling that $u^d(x, v) = (I + \alpha^{-1} \partial_1 d_0(\cdot, x) - x)^{-1} \circ \alpha^{-1} I$ and $A_x^d v = x - u^d(x, v)$, the above arguments prove (i) and the second part of statement (ii).

Assume that d complies with the property (p). By definition of u^d , this implies that $u^d(x, v) = x - A_x^d v \in \overline{C}$ and therefore (iii) is a consequence of Lemma 2.1. It remains to prove the first two assertions of (ii). Let us prove that A^d satisfies (r3). Let $(x, v) \in C \times \mathbb{R}^n$, be such that $A_x v = 0$. Then by (3.7), $x = \partial_1 d^*(\cdot, x)(-v)$, which implies that $\partial_1 d(x, x) = \nabla_1 d(x, x) = -v$. Therefore, by (P2) one has $v = 0$. Now to prove that (r2) is also satisfied, we use the following

Lemma 3.1 (*Baillon-Haddad [10]*)

Let H , $\langle \cdot, \cdot \rangle$ be a Hilbert space whose norm is denoted $|\cdot|$, $\phi : H \mapsto \mathbb{R}$ a C^1 convex function and $L > 0$. The following statements are equivalent,

- (i) $\forall (x, y) \in H^2$, $|\nabla\phi(x) - \nabla\phi(y)| \leq L|x - y|$
- (ii) $\forall (x, y) \in H^2$, $\langle \nabla\phi(x) - \nabla\phi(y), x - y \rangle \geq \frac{1}{L}|\nabla\phi(x) - \nabla\phi(y)|^2$.

In view of (3.7) and (i), this result can be applied to $\phi := d^*(\cdot, x)$. Hence, for x fixed in C , and for all $(v_1, v_2) \in \mathbb{R}^n \times \mathbb{R}^n$ it gives

$$\langle \partial_1 d^*(\cdot, x)(v_1) - \partial_1 d^*(\cdot, x)(v_2), v_1 - v_2 \rangle \geq \alpha |\partial_1 d^*(\cdot, x)(v_1) - \partial_1 d^*(\cdot, x)(v_2)|^2.$$

Now, letting $v_1 = 0$, and $v_2 = -v$ in the latter yields

$$\langle x - u^d(x, v), v \rangle \geq \alpha |x - u^d(x, v)|^2,$$

which, according to (3.6), is exactly (r2). \square

4 Elliptic barrier operators and continuous models for proximal algorithms: Examples and Properties

In this section we show that for various minimization algorithms one can derive an elliptic barrier operator and construct the associated $(A-DM)$ -dynamical system. It is worthwhile mentioning that many of the examples to follow will generate convergent trajectories to the minimizer of a convex function f over the closed convex set \overline{C} . From now on α will always denote the positive parameter involved in the definition of the class \mathcal{D} , cf. (3.4).

4.1 Projection-like methods

Let $h_0 : \mathbb{R}^n \mapsto \mathbb{R}$ be a C^1 convex function whose gradient is Lipschitz continuous on bounded sets, and set

$$\tilde{D}_h : \begin{cases} \mathbb{R}^n \times C & \rightarrow \mathbb{R}_+ \cup \{+\infty\} \\ (u, x) & \mapsto D_h(u, x) + \delta_{\overline{C}}(u). \end{cases}$$

with $h(u) = \frac{\alpha}{2}|u|^2 + h_0(u)$, $u \in \mathbb{R}^n$ and where D_h is given by (cf. (1.7)):

$$\forall (x, y) \in \mathbb{R}^n \times C, \quad D_h(x, y) = h(x) - h(y) - \langle \nabla h(y), x - y \rangle. \quad (4.1)$$

Proposition 4.1 *Let \tilde{D}_h as defined above. Then $A^{\tilde{D}_h}$ is an elliptic barrier operator that satisfies (r4). Moreover, we have for all $(x, v) \in C \times \mathbb{R}^n$,*

$$A_x^{\tilde{D}_h} v = x - (\nabla h + N_{\overline{C}})^{-1}(\nabla h(x) - v). \quad (4.2)$$

Proof. An easy computation gives $\tilde{D}_h(u, x) = \frac{\alpha}{2}|u - x|^2 + D_{h_0}(u, x) + \delta_{\overline{C}}(u)$. Letting $d_0(u, x) = D_{h_0}(u, x) + \delta_{\overline{C}}(u)$, we obtain that d_0 satisfies (P1) and (P3). For $(u, x) \in C \times C$, we have $\nabla_1 d_0(u, x) = \nabla h_0(u) - \nabla h_0(x)$, and as a consequence (P2) is satisfied as well. Therefore \tilde{D}_h is in \mathcal{D} , and clearly verifies (p). Now applying Proposition 3.1, it follows that $A^{\tilde{D}_h}$ satisfies (r2), (r3) and (v). The explicit formula of $A^{\tilde{D}_h}$ follows from (3.7). To obtain (r1) and (r4), we just have to observe that $(\nabla h + N_{\overline{C}})^{-1}$ and ∇h_0 are locally Lipschitz continuous on \mathbb{R}^n . \square

The terminology of projection relies on the fact that (4.2) can be seen as some twisted projection in the Bregman sense. Indeed, defining the projection of $z \in \mathbb{R}^n$ on \overline{C} by

$$P_{\overline{C}}^h(z) := \arg \min \{D_h(u, z) \mid u \in \overline{C}\},$$

we obtain that $P_{\overline{C}}^h(z) = (\nabla h + N_{\overline{C}})^{-1}(\nabla h(z))$ (recall that $\alpha > 0$) and therefore since $\nabla h^* = (\nabla h)^{-1}$, one can write

$$A_x^{\tilde{D}_h} v = x - P_{\overline{C}}^h(\nabla h^*(\nabla h(x) - v)), \quad \forall (x, v) \in C \times \mathbb{R}^n.$$

It is worthwhile noticing that in the framework of convex minimization, the gradient-like map $x \mapsto A_x^{\tilde{D}_h} \nabla f(x)$ enjoys remarkable properties. As a matter of fact, assume that the objective function f is convex, and observe that the following characterization holds

$$x^* \text{ solves } (\mathcal{P}) \text{ iff } A_{x^*}^{\tilde{D}_h} \nabla f(x^*) = 0.$$

The associated $A^{\tilde{D}_h}$ -driven descent method leads to the following differential equation

$$\dot{x}(t) + x(t) - P_{\overline{C}}^h(\nabla h^*[\nabla h(x(t)) - \nabla f(x(t))]) = 0, \quad x(0) \in C, \quad \forall t \geq 0. \quad (4.3)$$

Note that with $h_0 = 0$ and $\alpha = 1$, the corresponding dynamical system $(A^{\tilde{D}_h} - DM)$ (with corresponding operator A^P) is nothing else but the continuous gradient projection method (1.6), that is

$$\dot{x}(t) + x(t) - P_{\overline{C}}[x(t) - \nabla f(x(t))] = 0, \quad x(0) \in C, \quad \forall t \geq 0.$$

We remark that if $x(0) \notin C$ we still obtain convergent trajectories (with f convex), see [4] or Bolte [12], but the dynamical system is neither a descent, nor an interior method.

4.2 Continuous models for Bregman proximal minimization algorithms.

In this section, we give two quite different continuous models associated with proximal methods based on Bregman distances.

Continuous model I: A Riemannian gradient method

Our model appears as a particular case of Riemannian gradient methods on the smooth manifold C . Let us make precise the setting. Denote by $S_n^{++}(\mathbb{R})$ the cone of real definite

positive symmetric matrices and let $\mathcal{T}_x C$ be the tangent space to C at $x \in C$. In the sequel we make the usual identification $\mathcal{T}_x C \simeq \mathbb{R}^n$ for all $x \in C$. If g is some differentiable metric on C , there exists a unique differentiable application $\lambda : C \rightarrow S_n^{++}(\mathbb{R})$ such that for all $(x, u, v) \in C \times \mathbb{R}^n \times \mathbb{R}^n$

$$g_x(u, v) = \langle \lambda(x)u, v \rangle.$$

The gradient of a smooth function ϕ with respect to the metric g , is then given by the formula $\nabla_g \phi(x) = \lambda(x)^{-1} \nabla \phi(x)$, $\forall x \in C$, and the corresponding gradient method is

$$\begin{cases} \dot{x}(t) + \nabla_g \phi(x(t)) = 0, \\ x(0) \in C. \end{cases} \quad (4.4)$$

For $C = \mathbb{R}^n$, ϕ real analytic and g continuously differentiable, a deep result of Lojasiewicz [31] allows to prove that all bounded trajectories defined on $[0, +\infty)$ converge to some critical point of ϕ .

Readers interested in the use of geometric tools in optimization are referred to Bayer-Lagarias [11] in the context of Linear Programming and for more general results to the recent monograph of Helmke-Moore [21] and references therein.

Remark 4.1 Although our primary concerns in this paper are far removed from the complexity analysis of optimization algorithms, let us mention that there exists an intimate relation between Riemannian geometry and the complexity analysis of interior point optimization methods, see e.g., the work of Karmarkar [24] in the context of linear programming. More generally, in the context of convex programming, Nemirovsky and Nesterov [33] introduced the fundamental concept of self-concordant barrier functions for a constraint set C , which plays a central role in the design and analysis of interior methods with polynomial complexity. Thus, an interesting topic which is left for future research, would be to study a Riemannian metric defined on C , based for example on the Hessian of a self-concordant barrier, and which could lead to further insights on the complexity of barrier methods.

We focus here on the special choice of the application $\lambda : C \rightarrow S_n^{++}(\mathbb{R})$ defined by $\lambda = \nabla^2 h$ where h is some C^3 Bregman function with zone C , see Definition 4.1, below. The idea is to penalize the Euclidean scalar product, rather than the objective function, and to study the corresponding Riemannian gradient method

$$\dot{x}(t) + \nabla^2 h(x(t))^{-1} \nabla f(x(t)) = 0, \quad (4.5)$$

or *equivalently*

$$\frac{d}{dt} \nabla h(x(t)) + \nabla f(x(t)) = 0. \quad (4.6)$$

When the objective function is linear this differential equation has been considered in Iusem-Svaiter-Da Cruz [23], however their approach to the asymptotic behaviour strongly relies on the linear properties of f , see Remark 5.3 (b) for an insight. Observe that this dynamics

has, in its first form (4.5), the structure of A -driven descent methods. We shall see actually that most of classical Bregman functions can generate a barrier operator. Moreover, as shown below, the general framework developed in Section 3 allows to recover those methods by considering families of quadratic forms.

For the moment, let us compare (4.6) with (BPM) as given in the introduction. By an Euler implicit discretization we formally obtain

$$\frac{1}{\Delta t_k} [\nabla h(x^{k+1}) - \nabla h(x^k)] + \nabla f(x^{k+1}) = 0, \Delta t_k > 0. \quad (4.7)$$

Now observe that (BPM) has exactly the form of (4.7), provided that the iterates remain in C [16, 17, 19].

Before going further, we need to recall some of the basic facts concerning Bregman functions. Their definition relies mainly on their D function, as specified in (1.7),

Definition 4.1 *A function $h : \overline{C} \rightarrow \mathbb{R}$ is called a Bregman function with zone C if it satisfies the following:*

- (i) h is C^1 on C .
- (ii) h is continuous and strictly convex on \overline{C} .
- (iii) For every $r \in \mathbb{R}$, the partial level subset $L_h(x_0, r) = \{y \in C \mid D_h(x_0, y) \leq r\}$ is bounded for every $x_0 \in \overline{C}$.
- (iv) Let $(y^k, k \in N)$ be a sequence in C and $x \in \overline{C}$. If $y^k \rightarrow x$ as $k \rightarrow +\infty$, then $D_h(x, y^k) \rightarrow 0$ as $k \rightarrow +\infty$.

This definition weakens the usual definition of Bregman function proposed by Censor and Lent in [15], and is actually inspired by the more general notion of B function introduced by Kiwiel in [27]. Because of (iv) and the smoothness property of h , we have kept the terminology of Bregman function.

For the asymptotic analysis of (4.6) which will be developed in Section 5, we already record here the following useful lemma due to Kiwiel ([27, Lemma 2.16]).

Lemma 4.1 *Let h be a Bregman function with zone C and $x \in \overline{C}$. If $y^k, k \in N$ is a bounded sequence in C such that $D_h(x, y^k) \rightarrow 0$ as $k \rightarrow +\infty$ then $y^k \rightarrow x$ as $k \rightarrow +\infty$.*

In relation with the barrier operators to follow, let us define now a subclass of Bregman functions with zone C .

For $h : \overline{C} \rightarrow \mathbb{R}$, we consider the following assumptions:

(r_h) There exist $\alpha > 0$ and a C^3 Bregman function with zone C denoted by h_0 , such that for all $x \in \overline{C}$

$$h(x) = \frac{\alpha}{2} |x|^2 + h_0(x).$$

(v_h) For every $b \in \text{bd } C$ and every $\nu \in N_{\overline{C}}(b)$ there exists $K, \epsilon > 0$ such that for every $x \in C, |x - b| < \epsilon$,

$$|\nabla^2 h(x)^{-1} \nu| \leq K \langle b - x, \nu \rangle.$$

The set of such functions is denoted by \mathcal{B}_C , and for each $h \in \mathcal{B}_C$ we define a family of quadratic forms by

$$q_h : \begin{cases} \mathbb{R}^n \times C & \rightarrow \mathbb{R}^n \\ (u, x) & \mapsto \langle \nabla^2 h(x)(u - x), u - x \rangle. \end{cases}$$

Proposition 4.2 *For every $h \in \mathcal{B}_C$, A^{q_h} is an elliptic barrier operator on C . Moreover, for all $(x, v) \in C \times \mathbb{R}^n$ the following formula holds*

$$A_x^{q_h} v = \nabla^2 h(x)^{-1} v. \quad (4.8)$$

Proof. To prove that $q_h \in \mathcal{D}$, it suffices to notice that by (r_h) ,

$$q_h(u, x) = \alpha/2|u - x|^2 + \langle \nabla^2 h_0(x)(u - x), u - x \rangle,$$

where $\langle \nabla^2 h_0(x)(u - x), u - x \rangle$ satisfies $(P1), (P2), (P3)$. This implies by Proposition 3.1, that the operator A^{q_h} satisfies $(r2), (r3)$. Note that q_h never satisfies the property (p) , which precludes the use of Proposition 3.1 (iii) .

Applying Definition 3.1, formula (4.8) can be derived easily from,

$$\nabla^2 h(x)[u^{q_h}(x, v) - x] + v = 0, \quad \forall (x, v) \in C \times \mathbb{R}^n.$$

Since the mapping $M \in S_n^{++}(\mathbb{R}) \rightarrow M^{-1}$ is C^∞ , we obtain by (r_h) that A^{q_h} satisfies $(r1)$. Let us prove that A^{q_h} complies with (v) of Definition 2.1. Take $b \in \text{bd } C$ and ν in $N_{\overline{C}}(b)$, and let us apply (v_h) . There exist $K, \epsilon > 0$ such that for every $v \in \mathbb{R}^n, x \in C, |x - b| < \epsilon$,

$$\langle -A_x^h v, \nu \rangle = -\langle \nabla^2 h(x)^{-1} v, \nu \rangle = -\langle v, \nabla^2 h(x)^{-1} \nu \rangle \leq K|v| \langle b - x, \nu \rangle.$$

Therefore, if v is bounded, the latter exactly amounts to (v) . \square

The next lemma gives a practical means to prove that a Bregman function is in the class \mathcal{B}_C .

For $a < b$ in $\overline{\mathbb{R}}$, $\varphi : (a, b) \rightarrow \mathbb{R}$ a C^2 Bregman function with zone (a, b) , consider the assumptions,

(v_l) If a is finite, there exist a neighborhood U of a in \mathbb{R} and a positive constant K_l such that

$$\forall u \in U \cap (a, b) \quad \varphi''(u) \geq K_l/(u - a),$$

(v_r) If b is finite, there exist a neighborhood V of b in \mathbb{R} and a positive constant K_r such that

$$\forall u \in V \cap (a, b) \quad \varphi''(u) \geq K_r/(b - u).$$

Lemma 4.2 *Let $\varphi_1, \dots, \varphi_n$ be some C^3 Bregman functions on \mathbb{R} with zones $(a_1, c_1), \dots, (a_n, c_n)$, $a_i < c_i$, $a_i, c_i \in \overline{\mathbb{R}}, \forall i \in \{1, \dots, n\}$. Assume that $\varphi_1, \dots, \varphi_n$ satisfy $(v_l), (v_r)$ on their respective zones, and for $\alpha > 0$ set,*

$$h(x) = \frac{\alpha}{2}|x|^2 + \sum_{i=1}^n \varphi_i(x_i).$$

Then h belongs to \mathcal{B}_K , where $K = \prod_{i=1}^n (a_i, c_i)$, and A^{q_h} is an elliptic barrier operator that satisfies (r4).

Proof. The fact that h is a C^3 Bregman function with zone K follows from [27, Lemma 2.8,(d)], and therefore (r_h) is satisfied.

To simplify the notations, let us assume that for all $i \in \{1, \dots, n\}$, $a_i = 0$ and $c_i = +\infty$ (which implies $K = \mathbb{R}_+^n$). For $b = (b_1, \dots, b_n) \in \text{bd } \mathbb{R}_+^n$, set $I(b) = \{i \in \{1, \dots, n\} | b_i = 0\} \neq \emptyset$ and $J(b) = \{i \in \{1, \dots, n\} | b_i \neq 0\}$. For each $i \in I(b)$, (v_l) yields the existence of a neighborhood U_i of 0 in \mathbb{R} and $K_i > 0$ such that

$$\forall u \in U_i \cap (0, +\infty) \quad \varphi''(u) \geq K_i/u. \quad (4.9)$$

Set $U_i = \mathbb{R}^n$ for each $i \in J(b)$, and $U = \mathbb{R}_{++}^n \cap \prod_{i=1..n} U_i$. Let $\nu \in N_{\overline{K}}(b)$, and observe that $\nu_i = 0$ for all $i \in J(b)$ and that $\nu_i < 0$ for all $i \in I(b)$. Therefore, for $x \in \mathbb{R}^n$ an easy computation gives

$$|\nabla^2 h(x)^{-1} \nu| \leq \sum_{i \in I(b)} \frac{-\nu_i}{|\alpha + \varphi_i''(x_i)|}.$$

Now if $x \in U$, (4.9) implies that

$$\begin{aligned} |\nabla^2 h(x)^{-1} \nu| &\leq \sum_{i \in I(b)} -\frac{1}{K_i} \nu_i \cdot x_i \\ &\leq \sup_{i \in I(b)} \frac{1}{K_i} \langle b - x, \nu \rangle. \end{aligned}$$

A direct computation gives for all $x \in K$, $i, j \in \{1, \dots, n\}$,

$$\left(\nabla^2 h(x)^{-1} \right)_{i,j} = \frac{\delta_{ij}}{\alpha + \varphi_i''(x_i)},$$

where $\delta_{ij} = 1$ if $i = j$, and $\delta_{ij} = 0$ otherwise. Applying again (v_l), we see that A^{q_h} can be continuously extended on \overline{K} . Hence A^{q_h} satisfies (r4). \square

Example 4.1 *Bregman-based Barrier operators and their dynamics.*

The list of examples below shows thanks to Lemma 4.2 that many classical Bregman functions can be used to provide an elliptic barrier operator. In what follows α is the positive regularizing term as defined in (v_h), and β is a positive parameter. For a Bregman function h with zone $I \subset \mathbb{R}$, set $h_n(x) = \sum_{i=1}^n h(x_i)$ for all $x \in I^n$.

(a) For $\theta \in (0, 1)$ consider $h(s) = \frac{\alpha}{2} s^2 - \beta \frac{s^\theta}{\theta}$, $s \in \mathbb{R}_+$. Then $h \in \mathcal{B}_{\mathbb{R}_{++}}$, $h_n \in \mathcal{B}_{\mathbb{R}_{++}^n}$ and the corresponding $(A^{q_{h_n}} - DM)$ system is

$$\dot{x}_i(t) + \frac{x_i(t)^{2-\theta}}{\alpha x_i(t)^{2-\theta} + \beta(1-\theta)} \frac{\partial f}{\partial x_i}(x(t)) = 0, \quad x_i(0) > 0, \quad \forall i \in \{1, \dots, n\}. \quad (4.10)$$

(b) $h(s) = \frac{\alpha}{2}s^2 + \beta s \log s$ on \mathbb{R}_+ is in $\mathcal{B}_{\mathbb{R}_{++}}$, $h_n \in \mathcal{B}_{\mathbb{R}_{++}^n}$ and the associated system is

$$\dot{x}_i(t) + \frac{x_i(t)}{\alpha x_i(t) + \beta} \frac{\partial f}{\partial x_i}(x(t)) = 0, \quad x_i(0) > 0, \quad \forall i \in \{1, \dots, n\}.$$

This system is exactly the regularized Lotka-Volterra equation (1.8) recently proposed in [6]. However, it is worthwhile noticing, that (1.8) was introduced there as a continuous model not based on (BPM), but on the proximal-like method,

$$x^{k+1} \in \arg \min \left\{ f(x) + c_k d_\varphi(x, x^k) \mid x \in \mathbb{R}_+^n \right\}, \quad c_k > 0,$$

where $\varphi(s) = s - \log s - 1$ and $d_\varphi(x, y) = \frac{\alpha}{2}|x - y|^2 + \beta \sum_{i=1}^n y_i \varphi(y_i^{-1} x_i)$ for all x, y in \mathbb{R}_{++}^n . For more results and applications on classical Lotka-Volterra systems see, e.g., Hofbauer-Sigmund [22].

(c) $h(s) = \frac{\alpha}{2}s^2 - \beta\sqrt{1-s^2}$ on $[-1, 1]$ is in $\mathcal{B}_{(-1,1)}$, $h_n \in \mathcal{B}_{(-1,1)^n}$ and the corresponding system is

$$\dot{x}_i(t) + \frac{(1-x_i(t)^2)^{3/2}}{\alpha(1-x_i(t)^2)^{3/2} + \beta} \frac{\partial f}{\partial x_i}(x(t)) = 0, \quad x_i(0) \in (-1, 1), \quad \forall i \in \{1, \dots, n\}.$$

(d) $h(s) = \frac{\alpha}{2}s^2 - \beta\sqrt{s(1-s)}$ on $[0, 1]$ is in $\mathcal{B}_{(0,1)}$, $h_n \in \mathcal{B}_{(0,1)^n}$ and the corresponding system is

$$\dot{x}_i(t) + \frac{4x_i(t)^{3/2}(1-x_i(t))^{3/2}}{4\alpha x_i(t)^{3/2}(1-x_i(t))^{3/2} + \beta} \frac{\partial f}{\partial x_i}(x(t)) = 0, \quad x_i(0) \in (0, 1), \quad \forall i \in \{1, \dots, n\}.$$

Remark 4.2 For $\epsilon, \gamma \geq 0$, and $f \in C^3(\mathbb{R}^n, \mathbb{R})$ set $h_{\epsilon, \gamma}(x) = \frac{\epsilon}{2}|x|^2 + \gamma f(x)$, $\forall x \in \mathbb{R}^n$. Then we have $h_{\epsilon, \gamma} \in \mathcal{B}_{\mathbb{R}^n}$, under one of the following assumptions:

- (\star) f is strongly convex, i.e., $\nabla^2 f - \lambda I$ is positive semi-definite, with $\lambda > 0$.
- (\star) f is convex and $\epsilon > 0$
- (\star) $\gamma = 0$, $\epsilon > 0$.

Letting $\epsilon = 0$, $\gamma = 1$ in the first case yields the continuous Newton descent method (1.9). The second version can be seen, for ϵ small, as a regularized Newton method

$$(A^{q_{h_{\epsilon, \gamma}}} - DM) \quad \dot{x}(t) + [\epsilon Id + \gamma \nabla^2 f(x(t))]^{-1} \nabla f(x(t)) = 0.$$

The last point with $\gamma = 0$, $\epsilon > 0$ gives rise to the classical steepest descent method.

In the examples just described, the $A^{q_{h_{\epsilon, \gamma}}}$ are elliptic barrier operators *on* \mathbb{R}^n so that the feasible set \overline{C} is the whole space \mathbb{R}^n , and $(v)_h$ holds vacuously. It actually raises another interesting aspect of barrier operators: they can be used also as a geometrical means to

improve convergence rate as well as well-posedness properties. This suggests, for instance, to go further in the study of the following Newton-Barrier methods

$$\dot{x}(t) + [\lambda \nabla^2 h(x(t)) + \mu \nabla^2 f(x(t))]^{-1} \nabla f(x(t)) = 0, \quad t \geq 0$$

with $\lambda, \mu > 0$ and where h is a C^3 Bregman function.

Continuous Model II

The Bregman distances appearing in the definition of projection methods (Section 4.1), can be used in a quite different way in order to provide some other continuous model of (BPM). Indeed, replacing the kernel h_0 defined on the whole space \mathbb{R}^n by some essentially smooth convex function (see definition below) allows to get rid of the normal cone and to reformulate (4.3) as

$$\nabla h(x(t) + \dot{x}(t)) - \nabla h(x(t)) + \nabla f(x(t)) = 0, \quad \forall t \geq 0.$$

This can be discretized as follows

$$\nabla h(x_{k+1}) - \nabla h(x_k) + \nabla f(x_{k+1}) = 0, \quad \forall k \in N,$$

and (BPM) is recovered with a sequence of stepsizes satisfying $c_k = 1, \forall k \in N$.

This model will be derived from our general framework developed in Section 3. First, we recall now the definition of essentially smooth convex functions, see [35].

Definition 4.2 *A proper convex function $\phi : \mathbb{R}^n \mapsto \mathbb{R} \cup \{+\infty\}$ is essentially smooth if it satisfies*

- (i) *the interior of $\text{dom } \phi$ is nonempty, i.e., $\text{int dom } \phi \neq \emptyset$.*
- (ii) *ϕ is differentiable on $\text{int dom } \phi$.*
- (iii) *For all b in the boundary of $\text{int dom } \phi$, and all sequence $x_k, k \in N$ in $\text{int dom } \phi$ such that $x_k \rightarrow b$ as $k \rightarrow +\infty$, we have $|\nabla \phi(x_k)| \rightarrow +\infty$ as $k \rightarrow +\infty$.*

As in subsection 4.1 we study now operators of the form A^{D_h} (cf (4.1)) for some relevant kernels h .

Let $h_0 : \mathbb{R}^n \mapsto \mathbb{R} \cup \{+\infty\}$ be a closed proper convex function such that,

- (i) _{h_0} h_0 is essentially smooth with in addition $\text{int dom } h_0 = C$,
- (ii) _{h_0} ∇h_0 is Lipschitz continuous on compact subsets of C .

For such a function h_0 , we set $h(u) = \alpha/2|u|^2 + h_0(u), \forall u \in \mathbb{R}^n$. In the following proposition, it is important to recall that D_h is an extended real function defined on the whole of $\mathbb{R}^n \times C$.

Proposition 4.3 *Let h be as above. Then A^{D_h} is an elliptic barrier operator on C , and for all $(x, v) \in C \times \mathbb{R}^n$ we have*

$$A_x^{D_h} v = x - \nabla h^*(\nabla h(x) - v). \quad (4.11)$$

Proof. From $(i)_{h_0}$ it ensues that $D_h \in \mathcal{D}$. Using the fact that h is essentially smooth with $\text{int dom } h = C$ we deduce that (p) is satisfied. By Proposition 3.1, we see that A^{D_h} verifies $(r2)$, $(r3)$, and (v) . The formula (4.11), follows from (3.6), and $(r1)$ from $(ii)_{h_0}$. \square

The associated A^{D_h} -driven descent method is thus given by

$$\dot{x}(t) + x(t) - \nabla h^*[\nabla h(x(t)) - \nabla f(x(t))] = 0, \quad x(0) \in C, \quad \forall t \geq 0, \quad (4.12)$$

or using $\nabla h^* = (\nabla h)^{-1}$ equivalently as

$$\nabla h(x(t) + \dot{x}(t)) - \nabla h(x(t)) + \nabla f(x(t)) = 0, \quad x(0) \in C, \quad \forall t \geq 0.$$

Example 4.2 Consider the regularized Burg's entropy obtained with, $g(s) = (\alpha/2)s^2 - \beta \log s$, $s > 0$, where β is a positive parameter. For $x \in \mathbb{R}_{++}^n$ set $h(x) = \sum_{i=1}^n g(x_i)$. The function h satisfies the requirements of Proposition 4.3. A direct computation shows that

$$(g^*)'(u) = \frac{u + \sqrt{u^2 + 4\alpha\beta}}{2\alpha}, \quad \forall u \in \mathbb{R}.$$

Substituting in (4.12), the following descent method is derived. For all $i = 1, \dots, n$,

$$\dot{x}_i(t) + x_i(t)/2 + (2\alpha)^{-1} \left(\beta/x_i(t) + \frac{\partial f}{\partial x_i}(x(t)) - \sqrt{[\alpha x_i(t) - \beta/x_i(t) - \frac{\partial f}{\partial x_i}(x(t))]^2 + 4\alpha\beta} \right) = 0,$$

for all $t \geq 0$ and with $x_i(0) > 0$, $\forall i \in \{1, \dots, n\}$.

It is interesting to notice that as $\alpha \rightarrow 0$ we do not recover here the Lotka-Volterra system; compare with the system given in Example 4.1 (b).

4.3 A continuous model for proximal algorithms with second order kernels

The class of operators A^{d_φ} defined in this section are built upon the kernels φ used to realize the $(RIPM)$ method introduced in [8], and which we now recall. Let $\varphi : \mathbb{R} \rightarrow \mathbb{R} \cup \{+\infty\}$ be a closed proper function whose domain $\text{dom } \varphi$ is a subset of $[0, +\infty)$. Consider the following assumptions on φ

- $(i)_\varphi$ φ is finite and C^2 on $(0, +\infty)$,
- $(ii)_\varphi$ φ is strictly convex on $(0, +\infty)$,
- $(iii)_\varphi$ $\lim_{s \rightarrow 0, s > 0} \varphi'(s) = -\infty$,
- $(iv)_\varphi$ $\varphi(1) = \varphi'(1) = 0$ and $\varphi''(1) > 0$,
- $(v)_\varphi$ for all $s > 0$, $\varphi''(1)(1 - \frac{1}{s}) \leq \varphi'(s) \leq \varphi''(1)(s - 1)$.

Now for $\alpha, \beta > 0$ set

$$\varphi(s) = \frac{\alpha}{2}(s - 1)^2 + \beta\varphi_0(s), \quad (4.13)$$

where φ_0 satisfies $(i)_\varphi - (v)_\varphi$, and denote by Φ the class of such functions. For $\varphi \in \Phi$, set

$$\forall (u, x) \in \mathbb{R}^n \times \mathbb{R}_{++}^n \quad d_\varphi(u, x) = \sum_{i=1}^n x_i^2 \varphi(x_i^{-1} u_i). \quad (4.14)$$

It is proved in [8], that the associated proximal method,

$$(RIPM) \quad x^{k+1} \in \arg \min \{f(x) + c_k d_\varphi(x, x_k) | x \in \mathbb{R}_+^n\}, \quad c_k > 0,$$

generates a positive sequence $\{x^k\}$ provided that $x^0 \in \mathbb{R}_{++}^n$. As a consequence an equivalent formulation of $(RIPM)$ is

$$c_k \partial_1 d_\varphi(x^{k+1}, x^k) + \nabla f(x^{k+1}) = 0, \quad \forall k \geq 1. \quad (4.15)$$

Under the additional assumptions that $\arg \min_{\mathbb{R}_+^n} f \neq \emptyset$, $\sum_{k=1}^{+\infty} c_k = \infty$ and

$$\alpha \geq \beta \varphi_0''(1), \quad (4.16)$$

it is proved in [8] that the sequence x^k , $k \in N$ converges to a minimizer of f .

Following the general framework developed in Section 3, we generate the elliptic barrier operator and dynamical system associated with $(RIPM)$.

Proposition 4.4 *Let $\varphi \in \Phi$. Then A^{d_φ} is an elliptic barrier operator, and one has for all $(x, v) \in \mathbb{R}_{++}^n \times \mathbb{R}^n$,*

$$\left(A_x^{d_\varphi} v\right)_i = x_i - x_i (\varphi^*)'(-x_i^{-1} v_i), \quad \forall i = 1, \dots, n. \quad (4.17)$$

Proof. For all $(u, x) \in \mathbb{R}_+^n \times \mathbb{R}_{++}^n$ we have $d_\varphi(u, x) = \alpha/2 |u - x|^2 + \beta d_{\varphi_0}(u, x)$, and therefore to prove that $d_\varphi \in \mathcal{D}$, we need to show that βd_{φ_0} satisfies $(P1)$, $(P2)$, and $(P3)$. $(P1)$ follows from $(i)_\varphi$, while $(P3)$ is a consequence of the definition of φ_0 . Using $(iv)_\varphi$, we see by a direct computation that $(P2)$ is satisfied and thus that $d_\varphi \in \mathcal{D}$.

Using Definition 3.1 with $d := d_\varphi \in \mathcal{D}$ the optimality conditions for (3.5) yields

$$v_i + x_i \varphi'(u_i x_i^{-1}) = 0, \quad \forall i = 1, \dots, n$$

from which formula (4.17) follows easily using $(\varphi^*)' = (\varphi')^{-1}$. Since $\text{dom } \varphi \subset \mathbb{R}_+$, we have for all $x \in \mathbb{R}_{++}^n$, $\text{dom } d_\varphi(\cdot, x) \subset \mathbb{R}_+^n$ and therefore by Proposition 3.1 A^{d_φ} satisfies $(r2)$, $(r3)$, and (v) .

It remains to prove that $(r1)$ holds. Using formula (4.17), and since $x \in \mathbb{R}_{++}^n$, it thus suffices to show that $(\varphi^*)'$ is Lipschitz continuous. But since here φ is a smooth α -strongly convex function, one has

$$(t - s)(\varphi'(t) - \varphi'(s)) \geq \alpha(t - s)^2, \quad \forall t, s > 0,$$

and thus recalling that $(\varphi^*)' = (\varphi')^{-1}$, one easily deduces the required Lipschitz property for $(\varphi^*)'$ and $(r1)$ follows. \square

Remark 4.3 (a) Requirement $(v)_\varphi$ allows acute controls on d_φ in the asymptotic analysis of $(RIPM)$ and $(A^{d_\varphi} - DM)$, (see, Section 5, Theorem 5.1), and is actually not needed for the above result. Technically those controls are the reason why our operator is based on φ and not on φ^* .

(b) The assumption $(iii)_\varphi$ reduces the computation of $(\varphi^*)'$ to the inversion of $\varphi'_{(0,+\infty)}$.

(c) Note also that A^{d_φ} does not satisfy (r4) in general, but as we shall see in the next section it has no consequence on the asymptotic study of $(A^{d_\varphi} - DM)$ when f is convex.

(d) One could also develop a similar construction with “regularized φ -divergence” distance-like functions, that is:

$$d(u, x) = \frac{\alpha}{2}|u - x|^2 + \sum_{i=1}^n x_i \varphi(u_i x_i^{-1}), \quad u, x \in \mathbb{R}_+^n$$

where $\varphi : \mathbb{R} \mapsto \mathbb{R} \cup \{+\infty\}$ is an essentially smooth convex function such that $(0, +\infty) \subset \text{dom } \varphi \subset [0, +\infty)$. Unfortunately the parameter α forbids the computation of the Legendre conjugates of $\partial d(\cdot, x)$, $x \in C := \mathbb{R}_+^n$, and leads to purely theoretical conclusions. This gives a new motivation to study barrier operators for which $\alpha = 0$ (see Remark 2.1 (a)).

The $(A^{d_\varphi} - DM)$ system is given by

$$\dot{x}_i(t) + x_i(t) - x_i(t)(\varphi^*)'(-x_i(t)^{-1} \frac{\partial f}{\partial x_i}(x(t))) = 0, \quad \forall t \geq 0,$$

or equivalently as

$$x_i(t) \varphi' \left(\frac{\dot{x}_i(t) + x_i(t)}{x_i(t)} \right) + \frac{\partial f}{\partial x_i}(x(t)) = 0, \quad t \geq 0.$$

To recover $(RIPM)$ by some discretization of $(A^{d_\varphi} - DM)$, the latter can be reformulated in the following way

$$\partial_1 d_\varphi(x(t) + \dot{x}(t), x(t)) + \nabla f(x(t)) = 0, \quad x(0) \in \mathbb{R}_{++}^n, \quad \forall t \geq 0. \quad (4.18)$$

Now, if we perform an implicit discretization of (4.18), it yields

$$\partial_1 d_\varphi(x^{k+1}, x^k) + \nabla f(x^{k+1}) = 0, \quad x^0 = x(0), \quad k \in N.$$

which is exactly (4.15), with $c_k = 1$.

Example 4.3 It is a delicate matter to build a function in Φ whose Fenchel conjugate is easily computable. As in [8] we focus on the important special choice given by a logarithmic-quadratic kernel,

$$\varphi(s) = \frac{\alpha}{2}(s - 1)^2 + \beta(-\log s + s - 1), \quad s > 0,$$

which admits (see [8, p.665]) an explicit conjugate $\varphi^* \in C^\infty(\mathbb{R})$, and with

$$(\varphi^*)'(s) = \frac{1}{2\alpha}[\alpha - \beta + s + \sqrt{(\alpha - \beta + s)^2 + 4\alpha\beta}], \forall s \in \mathbb{R}.$$

The corresponding $(A^\varphi - DM)$ system is then given by

$$\dot{x}_i(t) + \frac{\alpha + \beta}{2\alpha}x_i(t) + \frac{1}{2\alpha} \frac{\partial f(x(t))}{\partial x_i} - \sqrt{\frac{1}{4\alpha^2}[(\alpha - \beta)x_i(t) + \frac{\partial f(x(t))}{\partial x_i}]^2 + 4\alpha\beta x_i(t)^2} = 0, \quad (4.19)$$

with $i \in \{1, \dots, n\}$, $t \geq 0$ and $x(0) \in \mathbb{R}_{++}^n$. An interesting fact to notice is that (4.19) has a sense for any $x(0) \in \mathbb{R}^n$; this suggests like in [12] a study of its properties for *non feasible* initial data.

5 Asymptotic analysis for a convex objective function

In the sequel f satisfies the additional assumptions

$$(\mathcal{H}') : \begin{cases} f \text{ is convex,} \\ \arg \min_{\bar{C}} f \neq \emptyset. \end{cases}$$

This section proposes a criterion concerning elliptic barrier operators to obtain the convergence of the trajectories of $(A - DM)$. It is based on Lyapounov functionals and to their (theoretical) decreasing rate. This natural approach is inspired by the classical result of Bruck [14] on the generalized steepest descent method, and by the notions of Fejer or quasi-Fejer sequences which go back to the work of Ermoliev [20] and arise in monotone and generalized gradient optimization algorithms. Such techniques have also been applied succesfully to second order in time systems by Alvarez [1], and Alvarez-Attouch [2]. Before stating the main result of this section, let us describe the typical properties of those Lyapunov functionals, sometimes called *relative entropy*, when working on systems in the nonnegative orthant, see e.g., [22]. In what follows S should be understood as the set of equilibria of some convex function.

We suggest the following general definition for viable Lyapunov functionals.

Definition 5.1 *Let $S \subset \bar{C}$ be a nonempty set. A family of functions $\{e_a, a \in S\}$ is Lyapunov viable if it satisfies*

- (i)_e *For all $a \in S$, $e_a : C \rightarrow \mathbb{R}$ is C^1 .*
- (ii)_e *The functions e_a are nonnegative for all $a \in S$.*
- (iii)_e *For all $a \in S$, e_a is inf bounded. That is for every $r \in \mathbb{R}$, the set $\{y \in C | e_a(y) \leq r\}$ is bounded.*
- (iv)_e *Let $x^k, k \in N$ be a sequence in C . Then for all $a \in S$,*

$$e_a(x^k) \rightarrow 0 \text{ as } k \rightarrow +\infty \iff x^k \rightarrow a \text{ as } k \rightarrow +\infty.$$

The next result is a key lemma that can be used to establish convergence of trajectories of $(A - DM)$. First, we recall the following classical result (see e.g., [1, Lemma 2.2]) which will be useful to us.

Lemma 5.1 *Let $h : \mathbb{R} \rightarrow \mathbb{R}^+$ a C^1 function. If $(h')^+ := \max(0, h')$ is in $L^1(0, +\infty; \mathbb{R})$ then $\lim_{t \rightarrow +\infty} h(t)$ exists.*

Let us set $S := \arg \min_{\overline{C}} f$.

Lemma 5.2 *Let A be an elliptic barrier operator on C and f a function satisfying (\mathcal{H}_1) , (\mathcal{H}_2) , (\mathcal{H}') . Assume that there exists $\lambda > 0$, $\mu \in \mathbb{R}$ and a family of functions $\{e_a, a \in S\}$ that is Lyapunov viable (i.e., satisfying $(i)_e - (iv)_e$). Suppose in addition that for all $x \in C$,*

$$\langle -A_x \nabla f(x), \nabla e_a(x) \rangle + \lambda \langle \nabla f(x), x - a \rangle \leq \mu |A_x \nabla f(x)|^2. \quad (5.1)$$

If $x(t)$ is the solution of $(A - DM)$, then

(i) $f(x(t)) \rightarrow \inf_{\overline{C}} f$ as $t \rightarrow +\infty$, with the estimation

$$f(x(t)) - \inf_{\overline{C}} f \leq Mt^{-1}, \text{ for some } M > 0.$$

(ii) $\dot{x}(t) \rightarrow 0$ as $t \rightarrow +\infty$.

(iii) There exists $x^* \in S$ such that $x(t) \rightarrow x^*$ as $t \rightarrow +\infty$.

Proof. Let $a \in S$, by (5.1) and $(A - DM)$ we obtain

$$\frac{d}{dt} e_a(x(t)) + \lambda \langle \nabla f(x(t)), x(t) - a \rangle \leq \mu |\dot{x}(t)|^2, t \geq 0. \quad (5.2)$$

From the convex inequality it follows that for all $y \in \overline{C}$,

$$0 \geq f(a) - f(y) \geq \langle \nabla f(y), a - y \rangle. \quad (5.3)$$

Combining (ii) of Theorem 2.1, (5.3), and (5.2) yields $[\frac{d}{dt} e_a(x(t))]^+ \leq \mu |\dot{x}(t)|^2$, $t \geq 0$. From $(ii)_e$ and Lemma 5.1, we deduce that $e_a(x(t))$ converges as $t \rightarrow +\infty$. Hence, by $(iii)_e$, $x(\cdot)$ is bounded.

Coming back to (5.2), we obtain for all $T \geq 0$,

$$\lambda \int_0^T \langle \nabla f(x(t)), x(t) - a \rangle dt \leq \int_0^T |\dot{x}(t)|^2 dt + e_a(x(0)) - e_a(x(T)),$$

and since $\lambda > 0$,

$$\langle \nabla f(x(\cdot)), x(\cdot) - a \rangle \in L^1(0, \infty; \mathbb{R}). \quad (5.4)$$

From (5.4), (\mathcal{H}_1) , and the boundedness property of x we obtain that there exist $x^* \in \overline{C}$, and a nondecreasing sequence $t_k, k \in \mathbb{N}$ such that $\langle \nabla f(x(t_k)), x(t_k) - a \rangle \rightarrow 0$ and $x(t_k) \rightarrow x^*$ as $k \rightarrow +\infty$. Using (5.3) it ensues $f(x^*) \leq f(a)$ and thus $x^* \in S$.

By Theorem 2.1, (iii) and the continuity of f , we see that the latter argument implies $f(x(t)) \rightarrow \inf_{\overline{C}} f$ as $t \rightarrow +\infty$ and that all limit points of x are in S .

To prove the second part of (i), we first deduce from (5.2) and (5.3) that

$$\frac{d}{dt}e_a(x(t)) + \lambda(f(x(t)) - f(a)) \leq \mu|\dot{x}(t)|^2, \quad t \geq 0.$$

By integration it follows from Theorem 2.1 (iii) that for $t \geq 0$, $t\lambda[f(x(t)) - \inf_{\overline{C}} f] \leq e_a(x(0)) - e_a(x(t)) + \mu \int_0^t |\dot{x}|^2$. Using (iii)_e we obtain for all $t > 0$

$$\lambda[f(x(t)) - \inf_{\overline{C}} f] \leq \frac{1}{t}[e_a(x(0)) + \mu \int_0^t |\dot{x}|^2]. \quad (5.5)$$

The estimate announced in (i) is then a consequence of Theorem 2.1 (iv).

Let x_1^* and x_2^* be two cluster points of $x(\cdot)$ and $t_k, \tau_k, k \in N$ increasing sequences in \mathbb{R}^+ , such that $x(t_k) \rightarrow x_1^*$, $x(\tau_k) \rightarrow x_2^*$ as $k \rightarrow +\infty$. From (iv)_e, we deduce $e_{x_1^*}(x(t_k)) \rightarrow 0$ as $k \rightarrow +\infty$. But since the function $e_{x_1^*}(x(\cdot))$ has a limit as $t \rightarrow +\infty$, we also have $e_{x_1^*}(x(\tau_k)) \rightarrow 0$ as $k \rightarrow +\infty$, and by applying (iv)_e again we obtain $x_1^* = x_2^*$.

Let x^* be the limit point of $x(\cdot)$, it verifies the classical relation $\nabla f(x^*) \in -N_{\overline{C}}(x^*)$, and therefore (\mathcal{H}_1) implies that $(x(t), \nabla f(x(t)))$ has its limit point in $\{x^*\} \times -N_{\overline{C}}(x^*)$. Applying Proposition 2.1, it follows that $\dot{x}(t) \rightarrow 0$ as $t \rightarrow +\infty$. \square

Remark 5.1 (a) If $\mu \leq 0$, we have by (5.5)

$$f(x(t)) - \inf_{\overline{C}} f \leq \frac{1}{\lambda t} e_a(x(0)), \quad \forall t > 0$$

(b) Note that Lemma 5.2 allows to handle the case $\mu > 0$ in (5.1), which corresponds to quasi-Fejer convergence.

(c) The property (r3) has not been used, but it is implicitly contained in (5.1).

(d) Note also that the above result holds for an elliptic barrier operator which is possibly undefined on $\text{bd } C \times \mathbb{R}^n$.

Let us apply this result to some of the operators defined in Section 4. In what follows it is implicitly assumed that $C = \mathbb{R}_{++}^n$ when dealing with operators of the type A^{d_φ} , $\varphi \in \Phi$, while A^P is the gradient projection operator (cf. subsection 4.1).

Theorem 5.1 *Let $\varphi \in \Phi$ such that $\alpha \geq \beta\varphi_0''(1)$, $h \in \mathcal{B}_C$, and assume that f satisfies (\mathcal{H}_1) , (\mathcal{H}_2) , (\mathcal{H}') . Then the trajectories of $(A^P - DM)$, $(A^{q_h} - DM)$, and $(A^{d_\varphi} - DM)$ converge to some minimizer of f on \overline{C} . Moreover, for all trajectories x the following properties hold:*

(i) $f(x(t)) \rightarrow \inf_{\overline{C}} f$ as $t \rightarrow +\infty$, with the estimation

$$f(x(t)) - \inf_{\overline{C}} f \leq Mt^{-1}, \quad \text{where } M > 0.$$

(ii) $\dot{x}(t) \rightarrow 0$ as $t \rightarrow +\infty$.

Proof. By Propositions 4.1, 4.2, and 4.4, we know that A^P , A^{qh} , and $A^{d\varphi}$ are elliptic barrier operators. For every $a \in S$ and for all $x \in C$ set

$$\begin{aligned} e_a^P(x) &= f(x) - f(a) + \frac{1}{2}|x - a|^2, \\ e_a^h(x) &= D_h(a, x) = \frac{\alpha}{2}|x - a|^2 + D_{h_1}(a, x), \\ e_a^{d\varphi}(x) &= f(x) - f(a) + \theta|x - a|^2, \end{aligned}$$

where $\theta = (\alpha + \varphi_0''(1))/2$. Naturally the idea is to apply Lemma 5.2 to the operators A^P , A^{qh} , and $A^{d\varphi}$. Let $a \in S$. The functions $e_a^{d\varphi}$, e_a^h , e_a^P , satisfy clearly (i)_e, (ii)_e. To obtain (iii)_e, just notice that in the three cases, the structure of the functions has the following form

$$\xi_a(x) = k|x - a|^2 + \rho_a(x), \quad \forall x \in C,$$

with $\rho_a \geq 0$, $k > 0$. By definition of a Bregman function and by Lemma 4.1, e_a^h verifies (iv)_e. To prove that e_a^P and $e_a^{d\varphi}$ satisfy (iv)_e, we just have to combine (H), and the fact that a is a minimizer of f on \overline{C} . Let us prove that the property (5.1) holds for the couples (e_a^P, A^P) , (e_a^h, A^{qh}) , and $(e_a^{d\varphi}, A^{d\varphi})$.

- The continuous gradient projection method has already been studied from different viewpoints in [12], but for the sake of completeness we recall the argument. Let $x \in C$ and $a \in S$. The optimality property of the orthogonal projection operator gives for all $\xi \in \overline{C}$, $\langle x - \nabla f(x) - P_{\overline{C}}(x - \nabla f(x)), \xi - P_{\overline{C}}(x - \nabla f(x)) \rangle \leq 0$. Therefore if $\xi = a$, we obtain

$$\langle -\nabla f(x) + A_x^P \nabla f(x), a - x + A_x^P \nabla f(x) \rangle \leq 0,$$

or equivalently $\langle -A_x^P \nabla f(x), x - a + \nabla f(x) \rangle + |A_x^P \nabla f(x)|^2 + \langle \nabla f(x), x - a \rangle \leq 0$, which is (5.1) with $\mu = -1$.

- Now, let us consider A^{qh} where h is Bregman function that belongs to \mathcal{B}_C . Let us compute the gradient of e_a^h for all $a \in S$. For all $x \in C$, we have

$$\begin{aligned} \nabla e_a^h(x) &= \nabla[h(a) - h(\cdot) - \langle \nabla h(\cdot), a - \cdot \rangle](x) \\ &= \nabla^2 h(x)(x - a). \end{aligned}$$

And therefore $\langle -A_x^{qh} \nabla f(x), \nabla e_a^h(x) \rangle = -\langle \nabla^2 h(x)^{-1} \nabla f(x), \nabla^2 h(x)(x - a) \rangle = -\langle \nabla f(x), x - a \rangle$, which verifies (5.1) with $\mu = 0$ and $\lambda = 1$.

- Finally, let us deal with $e_a^P, A^{d\varphi}$. Our approach relies on the following key lemma proven in [8, Lemma 3.4]

Lemma 5.3 *For every $y_1 \in \mathbb{R}_+^n$ and for every $(y_1, y_2) \in \mathbb{R}_{++}^n \times \mathbb{R}_{++}^n$, we have*

$$\langle y_1 - y_2, \partial_1 d_\varphi(y_2, y_3) \rangle \leq \theta \left(|y_1 - y_3|^2 - |y_1 - y_2|^2 \right).$$

Note that it is here that the property (v)_φ, is needed. Indeed, the proof of this lemma is based on that assumption, together with the condition $\alpha \geq \beta\varphi_0''(1)$.

For all $i \in \{1 \dots, n\}$ and all $x \in \mathbb{R}_{++}^n$, set $(v_x)_i = -\left(A_x^{d_\varphi} \nabla f(x)\right)_i$. The A^{d_φ} -driven descent method can be rewritten as,

$$\partial_1 d_\varphi(x(t) + v_x, x) + \nabla f(x) = 0, \forall x \in \mathbb{R}_{++}^n. \quad (5.6)$$

Observe that $x \in \mathbb{R}_{++}^n$ implies $x + v_x \in \mathbb{R}_{++}^n$. Now for $a \in \arg \min_{\mathbb{R}_+^n} f$ and for all $x \in \mathbb{R}_{++}^n$, let us multiply (5.6) by $a - x - v_x$, this gives

$$\langle a - (v_x + x), \partial_1 d_\varphi(x + v_x, x) \rangle + \langle \nabla f(x), a - x - v_x \rangle = 0,$$

and therefore by Lemma 5.3

$$\theta \left(|a - x|^2 - |a - x - v_x|^2 \right) + \langle \nabla f(x), a - x - v_x \rangle \geq 0.$$

After direct algebra this reduces to

$$\langle v_x, 2\theta(x - a) + \nabla f(x) \rangle + \langle \nabla f(x), x - a \rangle + |v_x|^2 \leq 0, \forall x \in \mathbb{R}_{++}^n.$$

Recalling that $v_x = -A_x^{d_\varphi} \nabla f(x)$, we easily see that (5.1) is satisfied. \square

Remark 5.2 The convergence of the orbits generated by the other operators proposed in Section 4 remains an open question.

Localization of the limit point

Let A be an elliptic barrier operator, and e_a be a family of viable Lyapounov functionals satisfying (5.1) with $\mu \leq 0$. We assume moreover that for all a in $S \subset \overline{C}$, there exist a nonnegative convex function $\rho_a : C \mapsto \mathbb{R}$ and $k > 0$ such that

$$e_a(x) = k|x - a|^2 + \rho_a(x), \forall x \in C. \quad (5.7)$$

As in Lemaire [28], and inspired by the recent non Euclidean extension given in [6], the limit point of the trajectory produced by $(A - DM)$ can be localized.

Proposition 5.1 *Let A be an elliptic barrier operator on C , and let $\{e_a, a \in S\}$ be as defined in (5.7). Then the trajectory of $(A - DM)$, with $x(0) \in C$, converges to a minimizer x_∞ of f on \overline{C} , with the following estimation*

$$|x_\infty - x(0)|^2 \leq \inf \{ 4|x(0) - a|^2 + \frac{2}{k} \rho_a(x(0)) \mid a \in S \}$$

Proof. The convergence result of the trajectory $x(t)$ to $x_\infty \in S = \arg \min_{\overline{C}} f$ is a direct consequence of Lemma 5.2. To prove the estimation, let us come back to the inequality (5.2), proven in Lemma 5.2:

$$\frac{d}{dt} e_a(x(t)) + \lambda \langle \nabla f(x(t)), x(t) - a \rangle \leq \mu |\dot{x}(t)|^2, t \geq 0.$$

The convexity property of f , and the fact that $\mu \leq 0$ imply that $\mathbb{R}_+ \ni t \mapsto e_a(x(t))$ is nonincreasing. Therefore, for all $a \in S$ we have $e_a(x(t)) \leq e_a(x(0))$, where $t \geq 0$. Since $\rho_a \geq 0$, by letting $t \rightarrow +\infty$, (5.7) yields

$$k|x_\infty - a|^2 \leq k|x(0) - a|^2 + \rho_a(x(0)). \quad (5.8)$$

Now for all $a \in S$, we have

$$\begin{aligned} |x_\infty - x(0)|^2 &\leq [|x_\infty - a| + |a - x(0)|]^2 \\ &\leq 2|x_\infty - a|^2 + 2|a - x(0)|^2 \\ &\leq 4|x(0) - a|^2 + \frac{2}{k}\rho_a(x(0)) \end{aligned}$$

where the third inequality is a consequence of (5.8). The desired result is then obtained by taking the infimum overall $a \in S$. \square

As a consequence, we then have

Corollary 5.1 *Under the assumptions of Theorem 5.1, we have*

$$|x_\infty - x(0)|^2 \leq 4 \inf\{|x(0) - a|^2 + \frac{1}{\alpha}D_{h_1}(a, x(0)) \mid a \in S\},$$

if $A = A^{q_n}$, $h(\cdot) = \alpha/2|\cdot|^2 + h_1(\cdot)$.

Defining $s : \mathbb{R}^n \mapsto \mathbb{R}$ as $s(y) := \inf\{|y - a|^2 \mid a \in S\}$, then we also have

$$|x_\infty - x(0)|^2 \leq 4 \left(s(x(0)) + f(x(0)) - \frac{\inf f}{C} \right)$$

if $A = A^P$, and

$$|x_\infty - x(0)|^2 \leq 4s(x(0)) + \frac{2}{\theta} \left(f(x(0)) - \frac{\inf f}{C} \right)$$

if $A = A^{d_\varphi}$.

Proof. The families $\{e_a^h, e_a^P, e_a^\varphi, a \in S\}$ introduced in the beginning of the proof of Theorem 5.1 satisfy the assumptions of Proposition 5.1, and thus the claimed results follow easily. \square

Remark 5.3 (a) The estimations given in Corollary 5.1 for $A = A^{q_n}$ allow to recover the results obtained in [6, 28].

(b) Assume that f is a linear function, that is $f(x) = \langle c, x \rangle$, $\forall x \in \mathbb{R}^n$ where $c \in \mathbb{R}^n$. Take h as in Theorem 5.1. A straightforward integration of $(A^{q_n} - DM)$ in its form given in (4.6) yields

$$\nabla h(x(t)) - \nabla h(x(0)) + tc = 0, \quad \forall t \geq 0. \quad (5.9)$$

As already noticed in [23], the trajectory of $(A^{q_n} - DM)$ can be viewed as an optimal path relatively to the barrier function D_h . Indeed since for all $(y, z) \in C \times C$, $\nabla_1 D_h(y, z) = \nabla h(y) - \nabla h(z)$, (5.9) can be reformulated as

$$x(t) \in \arg \min \left\{ \langle c, u \rangle + \frac{1}{t}D_h(u, x(0)) \mid u \in \mathbb{R}^n \right\}, \quad t > 0.$$

The convergence techniques developed in [23], but also the viscosity methods studied in Attouch [5], allow then to fully characterize the limit point as

$$x_\infty \in \arg \min \{D_h(a, x(0)) \mid a \in S\}.$$

References

- [1] ALVAREZ, F., *On the minimizing property of a second order dissipative system in Hilbert space*, SIAM J. of Control and Optimization, 38 (2000), N.4, pp. 1102-1119.
- [2] ALVAREZ, F., ATTOUCH, H., *An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping*, Wellposedness in Optimization and related topics (Gargagno, 1999), Set Valued Analysis, 9 (2001), N.1-2, pp 3-11.
- [3] ALVAREZ, F., PÉREZ, J.M., *A dynamical system associated with Newton's method for parametric approximations of convex minimization problems*, Appl. Math. Optim., 38 (1998), N.2, pp. 193- 217.
- [4] ANTIPIN, A.S., *Minimization of convex functions on convex sets by means of differential equations*, Differential Equations, 30 (1994), pp.
- [5] ATTOUCH, H., *Viscosity solutions of minimization problems*, SIAM J.Optim, 6 (1996), N.3, pp 769-806.
- [6] ATTOUCH, H., TEBoulLE, M., *A regularized Lotka Volterra dynamical system as a continuous proximal-like method in optimization*, December 2001. Submitted for publication.
- [7] AUBIN, J.P., CELLINA, A., "Differential inclusions", Springer Verlag, 1984.
- [8] AUSLENDER, A., TEBoulLE, M., BEN-TIBA, S., *Interior proximal and multiplier method based on second order homogeneous kernels*, Mathematics of Operation Research, 24 (1999), pp 645-668.
- [9] AUSLENDER, A., COMINETTI, R., HADDOU, M., *Asymptotic analysis for penalty and barrier methods in convex and linear programming*, Mathematics of Operations Research, 22 (1997), N.1, pp 43-62.
- [10] BAILLON J.B., HADDAD G., *Quelques propriétés des opérateurs angle-bornés et n -cycliquement monotones*, Israel J. of Mathematiques, 26 (1977), pp 137-150.
- [11] BAYER, D., A., LAGARIAS, J.C., *The nonlinear geometry of linear programming: Part I, Affine and projective scaling trajectories; Part II, Legendre Transform coordinates and central trajectories*, Trans. of AMS, 314, (1989), pp 499-526 and pp 527-581.

- [12] BOLTE, J., *On the continuous gradient projection method in Hilbert spaces*, (2001), to appear in J. of Opt. Theory and Applications.
- [13] BRÉZIS, H., “Opérateurs maximaux monotones”, Lecture notes N.5, North Holland, 1973.
- [14] BRUCK, R.E., *Asymptotic convergence of nonlinear contraction semi-groups in Hilbert space*, Journal of Functional Analysis, 18 (1974), pp. 15-26.
- [15] CENSOR, Y., LENT, A., *An iterative row-action method for interval convex programming*, J. Opt. Theory and Applications, 34 (1981), N.3, pp. 321-353.
- [16] CENSOR, Y., ZENIOS, S., *The proximal minimization algorithm with D-functions*, J. of Optim. Theory and Applications, 73, (1992), pp. 451-464.
- [17] CHEN, G., TEBoulLE, M., *Convergence analysis of a proximal-like minimization algorithm using Bregman functions*, SIAM J. Optimization, 3, (1993), pp. 538-543.
- [18] DUNN, J.C., *Global and asymptotic convergence rate estimates for a class of projected gradient processes*, SIAM J. on Control, 19, (1981), pp. 368-400.
- [19] ECKSTEIN, J., *Nonlinear proximal point algorithms using Bregman functions, with application to convex programming*, Mathematics of Operations Research, 18 (1993), N.1, pp. 202-226.
- [20] ERMOLIEV, Y., M., *On the method of generalized stochastic gradients and Quasi-Fejer sequences*, Cybernetics, 5, (1969), pp. 208-220.
- [21] HELMKE, U., MOORE, J.B., “Optimization and Dynamical Systems”, Springer-Verlag, London, 1994.
- [22] HOFBAUER, J., SIGMUND, K., “The Theory of Evolution and Dynamical Systems”, Cambridge University Press, 1988.
- [23] IUSEM A.N., SVAITER B.F., DA CRUZ J.X., *Central paths, generalized proximal point methods, and Cauchy trajectories in Riemannian manifolds*, SIAM J. Control Optim, Vol.37, N.2 (1999), pp. 566-588.
- [24] KARMARKAR, N.K., *Riemannian geometry underlying interior point methods for linear programming*, in Mathematical developments arising from linear programming, J.C. Lagarias and M. Todd, Editors, Copmtemporary Mathematics, 114, (1990), pp. 51–75.
- [25] KENMOCHI, N., PAWLOW, I., *A class of doubly nonlinear elliptic-parabolic equations with time dependent constraints*, Nonlinear Analysis, 10 (1986),pp 1181-1202.
- [26] KIWIEL, K.C., *Proximal mimization methods with generalized Bregman functions*, SIAM J. of Control and Optimization, 35 (1997), N.4, pp. 1142-1168.

- [27] KIWIŁ, K.C., *Free-steering relaxation methods for problems with strictly convex costs*, M. of Operations Research. , 22 (1997), N.2, pp. 326-349.
- [28] LEMAIRE, B., *An asymptotical variational principle associated with the steepest descent method for a convex function*, J.Convex Analysis, 3, 1996, pp. 63-70.
- [29] LEMAIRE, B., *The proximal point algorithm*, International series of numerical mathematics 1987 (J.P. Penot, Ed.), Birkhauser, Verlag, Basel, pp. 73-87.
- [30] LEVITIN, E. S., POLYAK, B.T., *Constrained minimization methods*, USSR Comp. Math. and Phys., 6 (1965), pp. 787-823.
- [31] LOJASIEWICZ, S., *Une propriété topologique des sous-ensembles analytiques réels*, Les équations aux dérivées partielles (Paris 1962), pp. 87-89, Editions du centre National de la Recherche Scientifique, Paris.
- [32] MARTINET, B., *Perturbation des méthodes d'optimisation, application*, R.A.I.R.O, Analyse numérique, 12 (1978), pp. 153-171.
- [33] NESTEROV, Y., NEMIROVSKII, A., "Interior-point polynomial algorithms in convex programming", SIAM Studies in Applied Mathematics, Philadelphia, PA, 1994.
- [34] OPIAL, Z., *Weak convergence of the sequence of successive approximations for nonexpansive mappings*, Bull. of the American Mathematical Society, 73 (1967), pp 591-597.
- [35] ROCKAFELLAR, R.T., *Convex Analysis*, Princeton University Press, 1970.
- [36] TEBoulLE, M., *Entropic proximal mappings with applications to nonlinear programming*, Mathematics of Operations Research, 17 (1992), N.3, pp. 670-690.