TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 370, Number 2, February 2018, Pages 755–790
http://dx.doi.org/10.1090/tran/6965
Article electronically published on July 13, 2017

ASYMPTOTIC BEHAVIOR OF NONAUTONOMOUS MONOTONE AND SUBGRADIENT EVOLUTION EQUATIONS

HEDY ATTOUCH, ALEXANDRE CABOT, AND MARC-OLIVIER CZARNECKI

ABSTRACT. In a Hilbert setting H, we study the asymptotic behavior of the trajectories of nonautonomous evolution equations $\dot{x}(t) + A_t(x(t)) \ni 0$, where for each $t \geq 0$, $A_t : H \rightrightarrows H$ denotes a maximal monotone operator. We provide general conditions guaranteeing the weak ergodic convergence of each trajectory $x(\cdot)$ to a zero of a limit maximal monotone operator A_∞ as the time variable t tends to $+\infty$. The crucial point is to use the Brézis-Haraux function, or equivalently the Fitzpatrick function, to express at which rate the excess of $\mathrm{gph}A_\infty$ over $\mathrm{gph}A_t$ tends to zero. This approach gives a sharp and unifying view of this subject. In the case of operators $A_t = \partial\varphi_t$ which are subdifferentials of proper closed convex functions φ_t, we show convergence results for the trajectories. Then, we specialize our results to multiscale evolution equations and obtain asymptotic properties of hierarchical minimization and selection of viscosity solutions. Illustrations are given in the field of coupled systems and partial differential equations.

1. Introduction and notation

Throughout the paper, H is a real Hilbert space which is endowed with the scalar product $\langle \cdot, \cdot \rangle$ and the norm $\| \cdot \|$ defined by $\|x\| = \sqrt{\langle x, x \rangle}$ for any $x \in H$. We study the asymptotic behavior of the NonAutonomous Monotone Inclusion

$$\text{(NAMI)} \qquad \dot{x}(t) + A_t(x(t)) \ni 0, \qquad t \geq 0,$$

where for every $t \geq 0$, $A_t : H \rightrightarrows H$ denotes a maximal monotone operator. Following Brézis [17, Definition 3.1], we say that $x : [0, +\infty[\to H$ is a strong global solution of (NAMI) if $x(\cdot)$ is locally absolutely continuous on $[0, +\infty[$ and if (NAMI) holds for almost all $t > 0$. We take for granted the existence of strong solutions to (NAMI). The existence of solutions of nonautonomous differential inclusions governed by time-dependent maximal monotone operators is a nontrivial topic. This issue has been studied extensively in the years 1970-1980; see Brézis [17], Attouch and Damlamian [9], Kenmochi [27], and references therein.

We prove the ergodic weak convergence of the trajectories of (NAMI) under some general condition involving the Brézis-Haraux function associated to the operator A_t. The Brézis-Haraux function $G_M : H \times H \to \mathbb{R} \cup \{+\infty\}$ associated to

Received by the editors July 30, 2015 and, in revised form, April 18, 2016.
2010 *Mathematics Subject Classification.* Primary 34G25, 37N40, 46N10, 47H05.
Key words and phrases. Nonautonomous monotone inclusion, subgradient inclusion, multiscale gradient system, hierarchical minimization, asymptotic behavior, Brézis-Haraux function, Fitzpatrick function.
The first author's effort was sponsored by the Air Force Office of Scientific Research, Air Force Material Command, USAF, under grant No. FA9550-14-1-0056.
The authors were supported by ECOS grant No. C13E03.

the maximal monotone operator M was introduced in [19]. It is defined by

$$G_M(x,u) = \sup_{(y,v)\in \operatorname{gph} M} \langle x-y, v-u\rangle,$$

where $\operatorname{gph} M$ denotes the graph of M. The function G_M is nonnegative and takes the zero value on the graph of M. The function G_M is connected with the Fitzpatrick function F_M via the formula $G_M(x,u) = F_M(x,u) - \langle x,u\rangle$, for every $(x,u) \in H\times H$. If there exists a maximal monotone operator $A_\infty : H \rightrightarrows H$ such that $S_\infty = A_\infty^{-1}(0) \neq \emptyset$ and if

$$\forall (z,p) \in \operatorname{gph} A_\infty, \qquad \int_0^{+\infty} G_{A_t}(z,p)\,dt < +\infty,$$

then we show that every strong global solution of (NAMI) converges weakly in average toward an element of S_∞, as $t\to+\infty$. As a by-product, we recover the Baillon-Brézis theorem [13] in the case of an autonomous evolution inclusion. The above integral condition is well suited for structured problems of the form $A_t = A + \beta(t)B$, with A, $B : H \rightrightarrows H$ maximal monotone operators, and $\beta(t)$ a time-dependent parameter. In this framework, we recover as a particular case a condition due to Bot-Csetnek [15, section 2] that guarantees the weak ergodic convergence of a forward-backward penalty scheme. The Bot-Csetnek condition formulated by means of the Fitzpatrick function is itself a generalization of a former condition given by Attouch-Czarnecki [6]; see also [7, 8].

The second important part of the paper concerns the study of the asymptotic behavior of the NonAutonomous subGradient Inclusion

$$\text{(NAGI)} \qquad \dot{x}(t) + \partial\varphi_t(x(t)) \ni 0, \qquad t\geq 0,$$

where for every $t\geq 0$, $\varphi_t : H \to \mathbb{R}\cup\{+\infty\}$ is a proper closed convex function. Such an evolution inclusion falls into the framework of (NAMI) since the operator $\partial\varphi_t : H \rightrightarrows H$ is maximal monotone. In the context of subdifferential operators, we can obtain convergence of the trajectories instead of ergodic convergence. If we assume that the filtered family $(\varphi_t)_{t\geq 0}$ is nonincreasing with respect to t, the energy function $t\mapsto \varphi_t(x(t))$ decreases as $t\to+\infty$ toward the infimum of the limit function $\varphi_\infty = \operatorname{cl}(\inf_{t\geq 0}\varphi_t)$. By using the Opial lemma along with a suitable summability condition, we deduce the weak convergence of every trajectory toward a point of the set $\operatorname{argmin}\varphi_\infty$; see Theorem 2.2. When no monotonicity assumption is made on the family $(\varphi_t)_{t\geq 0}$, it may be tricky to prove that $\lim_{t\to+\infty}\varphi_t(x(t))$ exists. The reader is referred to [26], where ad hoc conditions are given in order to control the variations in time of the family $(\varphi_t)_{t\geq 0}$. Weak convergence of the trajectories is then obtained via energetical arguments. In the present paper, we propose an alternative approach, based on the study of the distance from the trajectory to the optimal set $\operatorname{argmin}\varphi_\infty$ (where φ_∞ is the limit of φ_t as $t\to+\infty$ in a variational sense). The argument follows from an extension of a result due to Baillon-Cominetti [12] in a finite dimensional framework. Under a suitable summability assumption, we derive the weak convergence of every trajectory of (NAGI) toward a minimizer of φ_∞; see Theorem 2.3.

Particular attention is devoted to the case $\varphi_t = \Phi + \beta(t)\Psi$, where Φ, $\Psi : H \to \mathbb{R}\cup\{+\infty\}$ are proper closed convex functions and $\beta(t)$ is a positive time-dependent parameter. This corresponds to the situation of coupled (sub)gradients with multiscale aspects. If $\beta(t)\to+\infty$ and if the set $C = \operatorname{argmin}\Psi$ is nonempty, the orbits of the Multiscale Asymptotic Gradient dynamics, studied in [6],

$$\text{(MAG)} \qquad \dot{x}(t) + \partial\big(\Phi + \beta(t)\Psi\big)(x(t)) \ni 0$$

tend to minimize the function Φ over the set $\operatorname{argmin}\Psi$, thus leading to a hierarchical minimization process. The problem of convergence as $t \to +\infty$ depends on the behavior as $\varepsilon \to 0$ of the quantity $\omega(\varepsilon)$ defined by

$$\omega(\varepsilon) = \inf_H \left(\left(\Psi - \inf_H \Psi \right) + \varepsilon \left(\Phi - \inf_C \Phi \right) \right).$$

The key condition that implies weak convergence of the trajectories of (MAG) is the following:

$$\int_0^{+\infty} \beta(t)\, |\omega(1/\beta(t))|\, dt < +\infty.$$

The map $\omega(\cdot)$ was introduced by Cabot [22] in the framework of a diagonal proximal point algorithm involving multiscale aspects. The behavior of the map $\omega(\cdot)$ was used later by Alvarez-Cabot [1] to find asymptotic selection properties of viscosity equilibria for semilinear evolution equations. By resorting to the duality theory, we show that the quantity $|\omega(\varepsilon)|$ is majorized by an expression depending only on the function Ψ. More precisely, there exists $p \in H$ in the range of the normal cone operator $N_C : H \rightrightarrows H$, such that[1]

$$|\omega(\varepsilon)| \leq \Psi^*(\varepsilon p) + \min_H \Psi - \sigma_C(\varepsilon p),$$

for every $\varepsilon \geq 0$. Assuming that $\min_H \Psi = 0$, we deduce that the above summability condition is satisfied as soon as

$$\int_0^{+\infty} \beta(t) \left[\Psi^* \left(\frac{p}{\beta(t)} \right) - \sigma_C \left(\frac{p}{\beta(t)} \right) \right] dt < +\infty,$$

for every vector p in the range of N_C. This is precisely the condition due to Attouch-Czarnecki [6] in order to ensure weak convergence of the trajectories of (MAG). When the function Ψ satisfies the quadratic conditioning property $\Psi \geq a\, d^2(\cdot, C)$ for some $a > 0$, the above assumption is fulfilled if $\int_0^{+\infty} (1/\beta(t))\, dt < +\infty$.

Each of the above mentioned convergence results relies on a summability condition with respect to some suitable quantity. The summability condition expresses that the integrand tends to zero sufficiently fast. Therefore the conditions stated above quantify the fact that the operators A_t (resp. functions φ_t) tend sufficiently *fast* toward their limit A_∞ (resp. φ_∞).

The problem of trajectory convergence toward a particular viscosity solution naturally arises when the operators A_t (resp. functions φ_t) *slowly* tend toward their limit. We give an answer to this important issue in two cases:

i) A first answer is given for a family $(\varphi_t)_{t\geq 0}$ of proper closed convex functions by using a technique of central path. For every $t \geq 0$, we assume that the function φ_t has a strong minimum $\xi(t) \in H$, i.e., for all $x \in H$,

$$\varphi_t(x) \geq \varphi_t(\xi(t)) + \alpha(t)\, \|x - \xi(t)\|^2, \quad \text{for some } \alpha(t) > 0.$$

Under the slow condition $\int_0^{+\infty} \alpha(t)\, dt = +\infty$, we show that any solution $x(\cdot)$ of (NAGI) satisfies $\lim_{t\to+\infty} \|x(t) - \xi(t)\| = 0$; thus it is attracted toward the optimal path $\xi(\cdot)$. It ensues that the trajectory $x(.)$ strongly converges if and only if the

[1]The functions Ψ^* and σ_C denote respectively the Fenchel conjugate of Ψ and the support function of C.

optimal path has a limit as $t \to +\infty$, and in this case the limits are equal. The phenomenon of attraction toward the central path was brought to light in [5], under a strong convexity property.

ii) A second answer is given in the case of the multiscaled evolution system

$$(\mathrm{MAG}_\varepsilon) \qquad \dot{x}(t) + \partial\big(\Phi + \varepsilon(t)\Psi\big)(x(t)) \ni 0,$$

where $\varepsilon : \mathbb{R}_+ \to \mathbb{R}_+^*$ is a slowly vanishing viscosity coefficient, i.e., $\lim_{t\to+\infty} \varepsilon(t) = 0$ and $\int_0^{+\infty} \varepsilon(t)\, dt = +\infty$. By reversing the roles of the functions Φ and Ψ, and by using a suitable time rescaling, which allows us to pass from $\beta(t) \to +\infty$ to $\varepsilon(t) \to 0$, we show the convergence of the trajectories of $(\mathrm{MAG}_\varepsilon)$ to particular solutions. As an important special case, if the set $\operatorname{argmin}_C \Psi$ is a singleton $\{\overline{x}\}$ for some $\overline{x} \in H$ (where $C = \operatorname{argmin} \Phi$), then for any strong global solution $x(\cdot)$ of $(\mathrm{MAG}_\varepsilon)$, we have $x(t) \to \overline{x}$ strongly in H as $t \to +\infty$. In the case of the Tikhonov approximation $\Psi(x) = \|x\|^2$, we obtain strong convergence to the element of minimal norm. Note that we do not assume $\varepsilon(\cdot)$ to be nonincreasing. Under such a general assumption, this asymptotic selection result for the Tikhonov approximation was first obtained by Cominetti-Peypouquet-Sorin [23].

The paper is organized as follows. In section 2, we study the asymptotic behavior of the strong global solutions of (NAMI). The main result of subsection 2.2 gives the ergodic weak convergence of the trajectories under some general condition involving the Brézis-Haraux function. Subsection 2.3 is devoted to results concerning the case $A_t = \partial\varphi_t$ for a family $(\varphi_t)_{t\geq 0}$ of proper closed convex functions. In this context, we show several results of convergence of trajectories. In section 3, special attention is dedicated to the case of structured problems of the form $\varphi_t = \Phi + \beta(t)\Psi$, where Φ, $\Psi : H \to \mathbb{R} \cup \{+\infty\}$ are proper closed convex functions and $\beta(t)$ is a parameter tending to infinity as $t \to +\infty$. We show asymptotic hierarchical minimization results. A key ingredient consists of the study of the infimum value associated to the viscosity minimization problem $\inf_H(\Psi + \varepsilon\Phi)$. Section 4 is devoted to this question, with new results obtained by using duality arguments. Symmetrically, in section 5 we consider the case $\varphi_t = \Phi + \varepsilon(t)\Psi$, where $\varepsilon : \mathbb{R}_+ \to \mathbb{R}_+^*$ is a vanishing viscosity coefficient. We complete this study by examining in section 6 two other classes of nonautonomous subgradient inclusions, corresponding respectively to the quasi-autonomous case and the sweeping process. Illustrations of our results in the case of coupled gradient systems with multiscale aspects are given in section 7.

Notation. For a function $f : H \to \mathbb{R} \cup \{+\infty\}$, the set $\operatorname{dom} f = \{x \in H : f(x) < +\infty\}$ is called the domain of f. We call f a proper function if $\operatorname{dom} f$ is a nonempty set. Let $f : H \to \mathbb{R} \cup \{+\infty\}$ be a proper convex function. The subdifferential of f at $x \in \operatorname{dom} f$ is defined by

$$\partial f(x) = \{p \in H : f(y) \geq f(x) + \langle p, y - x\rangle \quad \forall y \in H\},$$

while $\partial f(x) = \emptyset$ if $x \notin \operatorname{dom} f$. If the function f is proper, closed[2] and convex, the multivalued operator $\partial f : H \rightrightarrows H$ is maximal monotone. For a nonempty convex set $C \subset H$, the normal cone to C at $x \in C$ is given by

$$N_C(x) = \{p \in H : \langle p, y - x\rangle \leq 0 \quad \forall y \in C\}.$$

[2]Notice that for a proper convex function, the closedness property coincides with the lower semicontinuity.

It coincides with the set $\partial\delta_C(x)$, where δ_C is the indicator function of C, taking the value 0 on C, and $+\infty$ elsewhere. The Fenchel conjugate of a function $f : H \to \mathbb{R}\cup\{+\infty\}$ is defined by $f^*(p) = \sup_{x\in H}\{\langle p, x\rangle - f(x)\}$ for every $p \in H$. The support function of the set $C \subset H$ is given by $\sigma_C(p) = \delta_C^*(p) = \sup_{x\in C}\langle p, x\rangle$ for every $p \in H$. Given two functions f, $g : H \to \mathbb{R}\cup\{+\infty\}$, we define the inf-convolution of f and g as follows: for every $x \in H$,

$$(f \bigtriangledown g)(x) = \inf_{y\in H}\{f(y) + g(x-y)\}.$$

Recall that the equality $(f \bigtriangledown g)^* = f^* + g^*$ is always true, while the equality $(f+g)^* = f^* \bigtriangledown g^*$ holds true if f, g are convex and if there exists $x_0 \in \operatorname{dom} f \cap \operatorname{dom} g$ such that g is continuous at x_0. This last condition is known as the Moreau-Rockafellar condition. In the Hilbert space setting (even in Banach spaces), when f and g are closed and convex, it is enough to ask for $0 \in \operatorname{int}(\operatorname{dom} f - \operatorname{dom} g)$ in order to obtain the formula for the conjugate of the sum. Let us also mention that these qualification conditions do not only guarantee that $(f+g)^* = f^* \bigtriangledown g^*$ but also that the infimum in the definition of the infimal convolution is achieved. For classical facts on convex analysis, see for example [4, 10, 11, 24, 34, 35].

2. Nonautonomous monotone inclusion—general case

In our approach, the Brézis-Haraux and the Fitzpatrick functions will play a crucial role in order to capture the asymptotic behavior of the filtered sequence of maximal monotone operators $(A_t)_{t\to+\infty}$.

2.1. Graph convergence and convergence of the Brézis-Haraux functions.

A set-valued mapping M from H to H assigns to each $x \in H$ a set $M(x) \subset H$; hence it is a mapping from H to 2^H. Every set-valued mapping $M : H \to 2^H$ can be identified with its graph defined by

$$\operatorname{gph} M = \{(x, u) \in H \times H :\ u \in Mx\}.$$

To emphasize this, we speak of M as a multivalued operator (or multifunction, or correspondence) and we write $M : H \rightrightarrows H$. The domain and range of $M : H \rightrightarrows H$ are taken to be the sets

$$\operatorname{dom} M = \{x \in H :\ \exists u \in H \text{ with } u \in Mx\},$$

$$\operatorname{ran} M = \{u \in H :\ \exists x \in H \text{ with } u \in Mx\}.$$

The inverse mapping $M^{-1} : H \rightrightarrows H$ is defined by $M^{-1}(u) = \{x \in H :\ u \in Mx\}$ for every $u \in H$. The set $M^{-1}(0)$ of the zeros of M is denoted by $\operatorname{zer} M$. An operator $M : H \rightrightarrows H$ is said to be monotone if for any (x, u), $(y, v) \in \operatorname{gph} M$, one has $\langle y-x, v-u\rangle \geq 0$. It is maximal monotone if there exists no monotone operator whose graph strictly contains $\operatorname{gph} M$. For classical facts on maximal monotone operators in Hilbert spaces, see for example [11, 35]. Given a maximal monotone operator M, the Brézis-Haraux function $G_M : H \times H \to \mathbb{R}\cup\{+\infty\}$, introduced in [19], is defined by

$$G_M(x, u) = \sup_{(y,v)\in \operatorname{gph} M}\langle x - y, v - u\rangle.$$

Let us show that G_M is an exterior penalty function with respect to the graph of M. By Minty's theorem, we have the following characterization of $(x,u) \in \operatorname{gph} M$:

$$\begin{aligned} u \in Mx &\Leftrightarrow x + u \in x + Mx \\ &\Leftrightarrow x = (I+M)^{-1}(x+u) \\ &\Leftrightarrow x - (I+M)^{-1}(x+u) = 0. \end{aligned}$$

Thus, the function

$$P_M(x,u) := \|x - (I+M)^{-1}(x+u)\|^2$$

is a penalty function with respect to the graph of M. It is nonnegative, Lipschitz continuous on bounded sets, and $P_M(x,u) = 0 \Leftrightarrow (x,u) \in \operatorname{gph} M$. But P_M is difficult to handle practically because, in general, the computation of the resolvent is a difficult task. Let us show that the Brézis-Haraux function solves some of these difficulties. Given arbitrary $(x,u) \in H \times H$, by Minty's theorem, there exists a unique $y \in H$ such that

$$y + My \ni x + u,$$

which is $y = (I+M)^{-1}(x+u)$. Set $v = x + u - y$. We have $v \in My$, and $v - u = x - y$. Thus

$$\begin{aligned} G_M(x,u) &= \sup_{(\xi,\eta)\in \operatorname{gph} M} \langle x - \xi, \eta - u\rangle \\ &\geq \langle x - y, v - u\rangle \\ &= \|x-y\|^2 \\ &= \|x - (I+M)^{-1}(x+u)\|^2 = P_M(x,u). \end{aligned} \tag{1}$$

On the other hand, by monotonicity of M, we immediately have that G_M is less than or equal to zero on the graph of M. Thus G_M is an exterior penalty function with respect to the graph of M; see also [25, Corollary 3.9]. A major advantage of G_M is that it is more flexible than P_M for the practical computation, as we will show later. Another interesting feature of G_M is its close relationship with the convex analysis.

The Fitzpatrick function $F_M : H \times H \to \mathbb{R} \cup \{+\infty\}$ is defined by

$$\begin{aligned} F_M(x,u) &= \sup_{(y,v)\in \operatorname{gph} M} \{\langle x, v\rangle + \langle y, u\rangle - \langle y, v\rangle\} \\ &= G_M(x,u) + \langle x, u\rangle. \end{aligned}$$

The function F_M was introduced by Fitzpatrick in [25]. As a supremum of continuous affine functions, F_M is convex and lower semicontinuous with respect to the couple (x,u). This property makes it an effective tool to address the problems governed by maximal monotone operators, using methods of convex analysis. It is the subject of active research; see for example [14, 21, 29, 30, 33, 37, 38]. There are exact calculus formulas for the Fitzpatrick function of sums or compositions of maximal monotone operators, provided that suitable qualification conditions are verified. This is an argument in favor of opting for the calculation of the Fitzpatrick function F_M (resp. Brézis-Haraux function G_M) rather than the function P_M introduced above.

The convergence of nets of maximal monotone operators can be formulated in terms of the Brézis-Haraux function.

Proposition 2.1. *Let $\{A_t : H \rightrightarrows H, t \geq 0\}$ be a family of maximal monotone operators. Assume that there exists a maximal monotone operator $A_\infty : H \rightrightarrows H$ such that*

$$\forall (z,p) \in \operatorname{gph} A_\infty, \qquad \lim_{t\to+\infty} G_{A_t}(z,p) = 0.$$

Then, (A_t) converges in the resolvent sense to A_∞. Equivalently, the net (A_t) graph converges to A_∞.

Proof. Take arbitrary $y \in H$. By Minty's theorem there exists a unique $z \in H$ such that $z + A_\infty z \ni y$. Setting $p = y - z$, we have $p \in A_\infty z$, and $z = (I + A_\infty)^{-1} y$. By (1)

$$\begin{aligned} G_{A_t}(z,p) &\geq \|z - (I + A_t)^{-1}(z+p)\|^2 \\ &= \|(I + A_\infty)^{-1} y - (I + A_t)^{-1} y\|^2. \end{aligned} \tag{2}$$

By assumption, $\lim_{t\to+\infty} G_{A_t}(z,p) = 0$, which, by (2), implies the convergence of the resolvents. Recall that, for a sequence of maximal monotone operators, the convergence of the resolvents is equivalent to the graph convergence [2, Proposition 3.60]. □

Remark 2.1. The main ingredient in the previous result is the inequality $G_M \geq P_M$, which already appears in a paper by Penot and Zalinescu; see [33, Lemma 2.3]. By using the same inequality, it is shown in [33, Proposition 3.1] that if G_{A_t} converges to G_{A_∞} in the bounded-Hausdorff sense, then $A_t \to A_\infty$ for the bounded-Hausdorff convergence.

The following example shows that the convergence of the Brézis-Haraux functions (equivalently, of the Fitzpatrick functions) is a stronger notion of convergence than the graph convergence.

Take A a general maximal monotone operator, and $\varepsilon : \mathbb{R}_+ \to H$ a map such that $\lim_{t\to+\infty} \varepsilon(t) = 0$. Set $A_t(x) = A(x) + \varepsilon(t)$, with $\operatorname{dom} A_t = \operatorname{dom} A$. It is immediate to verify that A_t is maximal monotone and A_t graph-converges to A as $t \to +\infty$. An elementary computation gives, for any $(x,u) \in H \times H$,

$$G_{A_t}(x,u) = G_A(x, u - \varepsilon(t)).$$

Therefore, to obtain the convergence of graphs without convergence of the Brézis-Haraux functions, it is sufficient to produce a maximal monotone operator A such that

$$u \mapsto G_A(x,u)$$

is not continuous at a point $(x,u) \in \operatorname{gph} A$. Since they differ by a continuous bilinear term ($G_A(x,u) = F_A(x,u) - \langle x,u\rangle$), it is equivalent to prove the result for the mapping $u \mapsto F_A(x,u)$. Let us specialize $A \in \mathcal{B}(H)$ to be a bounded linear monotone self-adjoint operator. Let $q_A : H \to \mathbb{R}$, $q_A(x) = \frac{1}{2}\langle x, Ax\rangle$ be the quadratic form associated to A. By a straight computation using the Fenchel conjugate (see [11, Example 20.45]),

$$F_A(x,u) = 2(q_A)^* \left(\frac{1}{2}u + \frac{1}{2}Ax\right).$$

As a consequence, it is sufficient to consider A such that $(q_A)^*$ is not continuous. This means that A is not invertible (it is only positive semi-definite). For example, when $A = 0$, then F_A is the indicator function of $H \times \{0\}$, an extreme situation where the continuity property of $u \mapsto F_A(x,u)$ fails to be satisfied. Remark that if

$A \in \mathcal{B}(H)$ is strongly monotone, then $(q_A)^*$ is continuous, and the two notions of convergence coincide (in that particular case).

2.2. **Nonautonomous monotone inclusion: Ergodic convergence.** In this section, we study the asymptotic behavior of the trajectories of

$$\text{(NAMI)} \qquad \dot{x}(t) + A_t(x(t)) \ni 0, \qquad t \geq 0.$$

The trajectory $x(\cdot)$ is a strong global solution of (NAMI) in the sense of Brézis [17, Definition 3.1] if the map $x : [0, +\infty[\to H$ is absolutely continuous on any bounded interval $[0, T]$ and (NAMI) holds for almost every $t > 0$.

Recall that an absolutely continuous function is differentiable almost everywhere and that one can recover the function from its derivative by the usual integration formula. Uniqueness of the solution for a given Cauchy data is an immediate consequence of the monotonicity of the operators A_t. In the sequel, we take for granted the existence of strong solutions to (NAMI).

2.2.1. *Statement of the ergodic convergence result.*

Theorem 2.1. *Let $\{A_t : H \rightrightarrows H, t \geq 0\}$ be a family of maximal monotone operators. Assume that there exists a maximal monotone operator $A_\infty : H \rightrightarrows H$ such that* $\operatorname{zer} A_\infty \neq \emptyset$ *and*

$$(\Sigma 1) \qquad \forall (z, p) \in \operatorname{gph} A_\infty, \qquad \int_0^{+\infty} G_{A_t}(z, p)\, dt < +\infty.$$

Then every strong global solution $x(\cdot)$ of (NAMI) *converges weakly in average to some $x_\infty \in \operatorname{zer} A_\infty$; i.e., as $t \to +\infty$,*

$$\frac{1}{t} \int_0^t x(s)\, ds \rightharpoonup x_\infty.$$

Remark 2.2. From (2), we deduce that condition $(\Sigma 1)$ implies that

$$(3) \qquad \forall y \in H, \quad \int_0^{+\infty} \|(I + A_t)^{-1} y - (I + A_\infty)^{-1} y\|^2\, dt < +\infty.$$

Hence, for all $y \in H$,

$$(4) \qquad \liminf \operatorname{ess}_{t \to +\infty} \|(I + A_t)^{-1} y - (I + A_\infty)^{-1} y\| = 0,$$

a property which is directly related to the graph convergence of A_t to A_∞, as $t \to +\infty$ (recall that the graph convergence of a filtered sequence of maximal monotone operators is equivalent to the pointwise convergence of the resolvents). The detailed study of this relationship is an interesting subject for further research. Let us just say that when H is separable, a thorough inspection of properties (3) and (4), combined with the nonexpansive property of the resolvents, is likely to provide (up to a negligeable set) the graph convergence of A_t to A_∞.

Indeed, it is not necessary to deepen this topological analysis, as for our purpose, the integral form $(\Sigma 1)$, which is used throughout this paper, is a more convenient way to express the convergence of A_t to A_∞. It carries more information than the topological one: it expresses that, in the sense of the Brézis-Haraux functions, the excess of $\operatorname{gph} A_\infty$ over $\operatorname{gph} A_t$ tends to 0 fast enough as $t \to +\infty$.

As a special case of Theorem 2.1, we recover the Baillon-Brézis theorem [13].

Corollary 2.1 ([13])**.** *Let $A : H \rightrightarrows H$ be a maximal monotone operator such that* $\operatorname{zer} A \neq \emptyset$*. Let $x(\cdot)$ be a strong global solution of*

$$\dot{x}(t) + A(x(t)) \ni 0.$$

Then there exists $x_\infty \in \operatorname{zer} A$ such that $\frac{1}{t}\int_0^t x(s)\,ds \rightharpoonup x_\infty$ weakly in H, as $t \to +\infty$.

Proof. Take $A_t = A$ for every $t \geq 0$, and $A_\infty = A$. Since $G_A(z,p) = 0$ for every $(z,p) \in \operatorname{gph} A$, condition ($\Sigma$1) is verified, and therefore Theorem 2.1 applies. $\square$

Remark 2.3. Given a map $f : \mathbb{R}_+ \to H$ such that $\lim_{t\to+\infty} f(t) = 0$, let x be a strong solution of the quasi-autonomous dissipative system

$$\dot{x}(t) + A(x(t)) \ni f(t). \tag{5}$$

The example after Remark 2.1 suggests that it would be illusory to use Theorem 2.1 in order to address the asymptotic behavior of (5) in its full generality. Assuming that $\sup_{t\geq 0} \|x(t)\| < +\infty$ and that $f \in L^1((0,+\infty); H)$, it is shown in [36, Theorem 4.5] that there exists $x_\infty \in \operatorname{zer} A$ such that $\frac{1}{t}\int_0^t x(s)\,ds \rightharpoonup x_\infty$ weakly in H, as $t \to +\infty$. The proof relies on the notion of almost non-expansive curve, along with Opial-like techniques; see below.

2.2.2. *Proof of Theorem* 2.1. Let us recall the Opial lemma [31], along with an ergodic version named the Opial-Passty lemma.

Lemma 2.1 (Opial)**.** *Let H be a Hilbert space and $x : [0,+\infty[\to H$ be a function such that there exists a nonempty set $S \subset H$ which verifies that*

(*i*) *$\forall z \in S$, $\lim_{t\to+\infty} \|x(t) - z\|$ exists.*
(*ii*) *$\forall t_n \to +\infty$ with $x(t_n) \rightharpoonup x_\infty$ weakly in H, we have $x_\infty \in S$.*

Then, $x(t)$ converges weakly as $t \to +\infty$ to some element x_∞ of S.

For the following ergodic variant of the Opial lemma, the reader is referred to [32].

Lemma 2.2 (Opial-Passty)**.** *Let H be a Hilbert space, let S be a nonempty subset of H and let $x : [0,+\infty[\to H$ be a function. For any $t > 0$ set $X(t) = \frac{1}{t}\int_0^t x(s)\,ds$, and assume that*

(*i*) *$\forall z \in S$, $\lim_{t\to+\infty} \|x(t) - z\|$ exists.*
(*ii*) *$\forall t_n \to +\infty$ with $X(t_n) \rightharpoonup X_\infty$ weakly in H, we have $X_\infty \in S$.*

Then, $X(t)$ converges weakly as $t \to +\infty$ to some element X_∞ of S.

The proof of Theorem 2.1 relies on the Opial-Passty lemma applied with $S = \operatorname{zer} A_\infty$. Let us first show that for every $z \in \operatorname{zer} A_\infty$, $\lim_{t\to+\infty} \|x(t) - z\|$ exists. Fix $z \in \operatorname{zer} A_\infty$ and set $h(t) = \frac{1}{2}\|x(t) - z\|^2$. Since $-\dot{x}(t) \in A_t(x(t))$ for a.e. $t \in \mathbb{R}_+$, we have

$$\dot{h}(t) = \langle x(t) - z, \dot{x}(t)\rangle \leq G_{A_t}(z,0) \quad \text{a.e. on } \mathbb{R}_+.$$

From this inequality and assumption (Σ1) at the point $(z,0)$, it follows that $\dot{h}_+ \in L^1(0,+\infty)$, where $\dot{h}_+ := \max\{\dot{h}, 0\}$ denotes the nonnegative part of the function $\dot{h}$. From a classical lemma, this implies that $\lim_{t\to+\infty} h(t)$ exists in $\mathbb{R}$. Let us now show that every sequential weak cluster point of $X(t) = \frac{1}{t}\int_0^t x(s)\,ds$ belongs to $\operatorname{zer} A_\infty$. Let $(z,p) \in \operatorname{gph} A_\infty$, and consider again the function h defined by $h(t) = \frac{1}{2}\|x(t) - z\|^2$. Since $-\dot{x}(t) \in A_t(x(t))$ for a.e. $t \in \mathbb{R}_+$, we obtain

$$\dot{h}(t) + \langle x(t) - z, p\rangle = \langle x(t) - z, p + \dot{x}(t)\rangle \leq G_{A_t}(z,p) \quad \text{a.e. on } \mathbb{R}_+.$$

By integrating on $[0,t]$, we find that

$$h(t) + \left\langle \int_0^t x(s)\,ds - tz, p \right\rangle \leq h(0) + \int_0^t G_{A_s}(z,p)\,ds.$$

After division by t and taking into account $h(t) \geq 0$, we have

$$\begin{aligned} \langle X(t) - z, p \rangle &\leq \frac{1}{t} h(0) + \frac{1}{t} \int_0^t G_{A_s}(z,p)\,ds \\ &\leq \frac{c}{t} \quad \text{with } c = h(0) + \int_0^{+\infty} G_{A_s}(z,p)\,ds. \end{aligned}$$

Suppose now that $X(t_n) \rightharpoonup X_\infty$ as $n \to +\infty$ for a sequence $t_n \to +\infty$. Taking the limit as $n \to +\infty$ in $\langle X(t_n) - z, p \rangle \leq c/t_n$, we immediately obtain $\langle X_\infty - z, p \rangle \leq 0$. Hence we have proved that for every $(z,p) \in \operatorname{gph} A_\infty$,

$$\langle X_\infty - z, 0 - p \rangle \geq 0.$$

The maximal monotonicity of A_∞ allows us to infer that $0 \in A_\infty(X_\infty)$, that is, $X_\infty \in \operatorname{zer} A_\infty$. By Lemma 2.2, we conclude to the weak ergodic convergence of the trajectories of (NAMI). $\square$

2.3. **Nonautonomous subgradient inclusion.** Let us consider the nonautonomous subgradient inclusion

$$\text{(NAGI)} \qquad \dot{x}(t) + \partial\varphi_t(x(t)) \ni 0, \qquad t \geq 0,$$

where for every $t \geq 0$, $\varphi_t : H \to \mathbb{R} \cup \{+\infty\}$ is a proper closed convex function. As in section 2.2, a map $x : [0,+\infty[\to H$ is said to be a strong global solution of (NAGI) if it is absolutely continuous on any bounded interval $[0,T]$ and if (NAGI) holds for almost every $t > 0$. Equation (NAGI) is a particular case of (NAMI), since the operator $A_t = \partial\varphi_t$ is maximal monotone for every $t \geq 0$. In the framework of subdifferential operators, we can make precise the convergence results of section 2.2 and show the convergence (instead of the ergodic convergence) of the trajectories.

In the autonomous case, $\varphi_t \equiv \varphi$ for every $t \geq 0$, and (NAGI) reduces to the steepest descent system

$$\text{(SD)} \qquad \dot{x}(t) + \partial\varphi(x(t)) \ni 0, \qquad t \geq 0.$$

Bruck [20, Theorem 4] gives the weak convergence of the trajectories of (SD) when $\operatorname{argmin}\varphi \neq \emptyset$. It can be derived directly from the Baillon-Brézis theorem [13]. The proof relies on a global estimate of the time derivative (see [18, Theorem 5]) by using the equality

$$x(t) - \frac{1}{t}\int_0^t x(s)\,ds = \frac{1}{t}\int_0^t \dot{x}(s)\,s\,ds.$$

If one obtains the same estimate $\lim_{t\to+\infty} t\dot{x}(t) = 0$ in the present case, the weak convergence of the trajectories of (NAGI) is a direct consequence of the weak ergodic convergence of the trajectories of (NAMI). However, the extension of the energetical argument to the nonautonomous case, leading to the estimate, remains an open question in our general setting. So we provide specific results and proofs in the subgradient case.

2.3.1. *Case of a nonincreasing family $(\varphi_t)_{t\geq 0}$: Energetical approach.* In this subsection, we assume a monotonicity property on the filtered family $(\varphi_t)_{t\geq 0}$. This allows us to use energetical arguments in order to derive convergence of the trajectories of (NAGI).

Theorem 2.2. *Let $\{\varphi_t;\, t\geq 0\}$ be a family of proper closed convex functions from H to $\mathbb{R}\cup\{+\infty\}$. Assume that $\varphi_t \leq \varphi_s$ for every $s,\, t\geq 0$ such that $s\leq t$. Let us set $\varphi_\infty = \mathrm{cl}(\inf_{t\geq 0}\varphi_t)$. Let $x(.)$ be a strong global solution of* (NAGI) *such that the function $t\mapsto \varphi_t(x(t))$ is locally absolutely continuous. Then we have*

(*i*) *The function $t\mapsto\varphi_t(x(t))$ is nonincreasing, and $\lim\limits_{t\to+\infty}\varphi_t(x(t)) = \inf\limits_H \varphi_\infty$.*

Additionally assume that $\inf\limits_H\varphi_\infty > -\infty$. Then

(*ii*) $\displaystyle\int_0^{+\infty} \|\dot{x}(t)\|^2\, dt < +\infty.$

Assume moreover that $\operatorname{argmin}\varphi_\infty \neq \emptyset$ *and that*

$$(\Sigma 2)\qquad \forall z\in \operatorname{argmin}\varphi_\infty,\quad \int_0^{+\infty} G_{\partial\varphi_t}(z,0)\, dt < +\infty.$$

Then

(*iii*) *there exists $x_\infty\in\operatorname{argmin}\varphi_\infty$ such that $w-\lim_{t\to+\infty}x(t) = x_\infty$.*

Proof. (*i*) Let $t>0$ be such that the derivatives $\dot{x}(t)$ and $\frac{d}{dt}\varphi_t(x(t))$ exist at t and such that the inclusion $-\dot{x}(t)\in\partial\varphi_t(x(t))$ holds true. The subdifferential inequality yields for every $\tau\in]0,t[$,

$$\varphi_t(x(t-\tau)) \geq \varphi_t(x(t)) + \langle -\dot{x}(t), x(t-\tau)-x(t)\rangle.$$

Recalling that the family $\{\varphi_t;\, t\geq 0\}$ is nonincreasing, we have $\varphi_{t-\tau}(x(t-\tau)) \geq \varphi_t(x(t-\tau))$, thus implying that

$$\varphi_{t-\tau}(x(t-\tau)) - \varphi_t(x(t)) \geq \langle -\dot{x}(t), x(t-\tau)-x(t)\rangle.$$

Dividing by τ and taking the limit as $\tau\to 0$, we find that

$$(6)\qquad -\frac{d}{dt}\varphi_t(x(t)) \geq \|\dot{x}(t)\|^2 \geq 0.$$

Since this is true for almost every $t>0$, the map $t\mapsto\varphi_t(x(t))$ is nonincreasing and hence converges toward some $l\in\mathbb{R}\cup\{-\infty\}$. Using that $\varphi_t\geq\varphi_\infty$ for every $t\geq 0$, we obtain

$$(7)\qquad l = \lim_{t\to+\infty}\varphi_t(x(t)) \geq \inf_H \varphi_\infty.$$

Let us now fix $z\in H$ and define the auxiliary function $h:\mathbb{R}_+\to\mathbb{R}_+$ by $h(t) = \frac{1}{2}\|x(t)-z\|^2$. By differentiating and using the subdifferential inequality, we find for almost every $t\geq 0$,

$$(8)\qquad \begin{aligned} \dot{h}(t) &= \langle x(t)-z, \dot{x}(t)\rangle \\ &\leq \varphi_t(z) - \varphi_t(x(t)). \end{aligned}$$

Integrating this inequality, we get

$$\int_0^t [\varphi_s(z) - \varphi_s(x(s))]\, ds \geq h(t) - h(0) \geq -h(0).$$

We immediately deduce that $\lim_{s\to+\infty}\varphi_s(z) \geq \lim_{s\to+\infty}\varphi_s(x(s)) = l$. Since this is true for every $z\in H$, the function $\inf_{s\geq 0}\varphi_s = \lim_{s\to+\infty}\varphi_s$ is minorized by l. It ensues that the function $\varphi_\infty = \mathrm{cl}\,(\inf_{s\geq 0}\varphi_s)$ is also minorized by l. In view of (7), we conclude that $l = \inf_H \varphi_\infty$.

(ii) Integrating the first inequality of (6), we find for every $t \geq 0$,

$$\int_0^t \|\dot{x}(s)\|^2\, ds \leq \varphi_0(x(0)) - \varphi_t(x(t)).$$

Taking the limit as $t \to +\infty$, we deduce from (i) that

$$\begin{aligned}\int_0^{+\infty} \|\dot{x}(s)\|^2\, ds &\leq \varphi_0(x(0)) - \inf_H \varphi_\infty \\ &< +\infty \quad \text{since } \inf_H \varphi_\infty > -\infty \text{ by assumption.}\end{aligned}$$

(iii) The proof of the weak convergence $x(t) \rightharpoonup x_\infty$ is based on the Opial lemma. Fix $z \in \operatorname{argmin}\varphi_\infty$ and consider the function h defined above by $h(t) = \frac{1}{2}\|x(t)-z\|^2$. Coming back to equality (8) and recalling that $-\dot{x}(t) \in \partial\varphi_t(x(t))$ for almost every $t \geq 0$, we find that

$$\dot{h}(t) \leq G_{\partial\varphi_t}(z,0) \quad \text{a.e. on } \mathbb{R}_+,$$

where $G_{\partial\varphi_t}$ is the Brézis-Haraux function associated to the operator $\partial\varphi_t$. It follows from this inequality and assumption (Σ2) that $\dot{h}_+ \in L^1(0,+\infty)$. From a classical lemma, this implies that $\lim_{t\to+\infty} h(t)$ exists in $\mathbb{R}$. It suffices now to prove that every sequential weak cluster point of $x(\cdot)$ belongs to $\operatorname{argmin}\varphi_\infty$. Let $x_\infty \in H$ and let $t_n \to +\infty$ be a sequence such that $x(t_n) \rightharpoonup x_\infty$ as $n \to +\infty$. Since the family $(\varphi_t)_{t\geq 0}$ is nonincreasing, it Mosco converges toward $\varphi_\infty = \mathrm{cl}\,(\inf_{s\geq 0}\varphi_s)$; see [2, Theorem 3.20]. It ensues that

$$\begin{aligned}\varphi_\infty(x_\infty) &\leq \liminf_{n\to+\infty} \varphi_{t_n}(x(t_n)) \\ &= \liminf_{t\to+\infty} \varphi_t(x(t)) = \min_H \varphi_\infty \quad \text{in view of } (i).\end{aligned}$$

We conclude that $x_\infty \in \operatorname{argmin}\varphi_\infty$. It suffices then to apply the Opial lemma. $\square$

Remark 2.4. Condition (Σ2) is nothing other than condition (Σ1) applied with $A_t = \partial\varphi_t$ and $p = 0$. Recalling that

$$G_{\partial\varphi_t}(z,0) = F_{\partial\varphi_t}(z,0) \leq \varphi_t(z) + \varphi_t^*(0) = \varphi_t(z) - \inf_H \varphi_t,$$

we deduce that assumption (Σ2) is implied by

$$(\Sigma 3) \qquad \forall z \in \operatorname{argmin}\varphi_\infty, \quad \int_0^{+\infty} \left[\varphi_t(z) - \inf_H \varphi_t\right] dt < +\infty.$$

Remark 2.5. Assumptions (Σ2) and (Σ3) seem to be new in the study of the asymptotic behavior of the dynamical system (NAGI). Furuya, Miyashiba, and Kenmochi obtained the weak convergence of the trajectories of (NAGI) under an alternative condition; see [26, Theorem 2]. Their condition also requires some quantity to be summable, but it differs significantly from (Σ2) and (Σ3). In the framework of the diagonal proximal point method, Lemaire used a discrete anologue of (Σ3) to derive the weak convergence of the iterates; see [28, Section 4].

2.3.2. *A general result of convergence relying on the study of the distance to the optimal set* $\operatorname{argmin}\varphi_\infty$. As in the previous subsection, $x(\cdot)$ denotes a strong global solution of the evolution inclusion (NAGI). We now study the distance of the solution $x(t)$ to the set $\operatorname{argmin}\varphi_\infty$, and we show that it vanishes as $t \to +\infty$. This is in fact an extension of a result due to Baillon-Cominetti [12] in a finite dimensional framework. To obtain such an extension in a general Hilbert space, one has to assume some inf-compactness property on the functions φ_t. Let us recall that a function $f : H \to \mathbb{R}\cup\{-\infty,+\infty\}$ is said to be inf-compact if, for every $l \in \mathbb{R}$, the lower level set $\{x \in H : f(x) \le l\}$ is relatively compact in H. A weaker notion consists of requiring that the function $f + \delta_{\overline{B}(0,R)}$ is inf-compact[3] for every $R > 0$. Here $\overline{B}(0,R)$ denotes the closed ball of radius R centered at 0. This condition amounts to assuming that for every $R > 0$ and $l \in \mathbb{R}$ the lower level set

$$\{x \in H : \|x\| \le R,\ f(x) \le l\} \text{ is relatively compact in } H. \tag{9}$$

If H is finite dimensional, the ball $\{x \in H : \|x\| \le R\}$ is compact, and the inf-compactness property above is satisfied for every function $f : H \to \mathbb{R}\cup\{-\infty,+\infty\}$.

Theorem 2.3. *Let $\{\varphi_t;\, t \ge 0\}$ be a family of proper closed convex functions from H to $\mathbb{R}\cup\{+\infty\}$. Assume that*[4]

- (H1) *There exists a proper closed convex function $\varphi_\infty : H \to \mathbb{R}\cup\{+\infty\}$ such that the set* $\operatorname{argmin}\varphi_\infty$ *is nonempty and bounded.*
- (H2) $\varphi_\infty(x_\infty) \le \liminf_{k\to+\infty} \varphi_{t_k}(x_k)$ *for all convergent sequences $x_k \to x_\infty$ and $t_k \to +\infty$.*
- (H3) $\lim_{t\to+\infty} v_\infty(t) = \min_H \varphi_\infty$, *where* $v_\infty(t) = \sup\limits_{z\in\operatorname{argmin}\varphi_\infty} \varphi_t(z)$.
- (H4) *For t large enough, all functions φ_t are uniformly minorized by a function $f : H \to \mathbb{R}\cup\{-\infty,+\infty\}$ satisfying*[5] *the inf-compactness property* (9).

Let $x(\cdot)$ be a strong global solution of (NAGI). *Then we have*

- (*i*) $\lim_{t\to+\infty} d(x(t), \operatorname{argmin}\varphi_\infty) = 0$.
- (*ii*) *If we assume moreover that*

$$\forall z \in \operatorname{argmin}\varphi_\infty, \quad \int_0^{+\infty} G_{\partial\varphi_t}(z,0)\,dt < +\infty, \tag{$\Sigma 2$}$$

then there exists $x_\infty \in \operatorname{argmin}\varphi_\infty$ such that $x(t) \rightharpoonup x_\infty$ weakly in H as $t \to +\infty$.

Recall that assumption ($\Sigma 2$) is satisfied under the stronger condition

$$\forall z \in \operatorname{argmin}\varphi_\infty, \quad \int_0^{+\infty} \left[\varphi_t(z) - \inf_H \varphi_t\right] dt < +\infty; \tag{$\Sigma 3$}$$

see Remark 2.4.

Proof. (*i*) In a finite dimensional space, Baillon and Cominetti proved that

$$\lim_{t\to+\infty} d(x(t), \operatorname{argmin}\varphi_\infty) = 0$$

under (H1)-(H2)-(H3); see [12, Theorem 2.1]. An immediate adaptation of their arguments shows that this property still holds true in a Hilbert space under the additional assumption (H4).

[3] We use here the convention $(-\infty) + (+\infty) = +\infty$.

[4] For the convenience of the reader and coherence with the literature, we keep the name of the assumptions (H1)-(H2)-(H3) as in [12].

[5] If $H = \mathbb{R}^n$, assumption (H4) is automatically satisfied (take $f \equiv -\infty$).

(ii) The proof of the weak convergence $x(t) \rightharpoonup x_\infty$ is based on the Opial lemma. To show that $\lim_{t\to+\infty} \|x(t)-z\|$ exists for every $z \in \operatorname{argmin} \varphi_\infty$, we use the map h defined by $h(t) = \frac{1}{2}\|x(t)-z\|^2$ and we proceed as in the proof of Theorem 2.2(iii). The second point consists of proving that every weak limit point of $x(\cdot)$ belongs to $\operatorname{argmin} \varphi_\infty$. In fact, this is an immediate consequence of (i) and of the weak lower semicontinuity of the convex continuous function $d(., \operatorname{argmin} \varphi_\infty)$. ☐

2.3.3. *Slow case and strong attraction of the optimal path.* In this subsection, we assume that for every $t \geq 0$, there exist $\xi(t) \in H$ and $\alpha(t) > 0$ such that

$$\forall x \in H, \quad \varphi_t(x) \geq \varphi_t(\xi(t)) + \alpha(t)\, \|x - \xi(t)\|^2.$$

It implies that $\xi(t)$ is a strong minimum of the function φ_t.

Remark 2.6. Fix $z \in \operatorname{argmin} \varphi_\infty$. We deduce from the above condition that

$$\alpha(t)\, \|z - \xi(t)\|^2 \leq v_\infty(t) - \min_H \varphi_t.$$

If $\xi^* = \lim_{t\to+\infty} \xi(t)$ exists and is not equal to z, there exists $m > 0$ such that $\|z - \xi(t)\| \geq m$ for t large enough. It ensues that

$$\alpha(t) \leq \frac{1}{m^2}(v_\infty(t) - \min_H \varphi_t) \quad \text{ for } t \text{ large enough.}$$

We assume that the function α is measurable and satisfies

$$\int_0^{+\infty} \alpha(t)\, dt = +\infty,$$

which corresponds to a slow decay condition. Let us first consider the case of an optimal trajectory having a finite length. The following result is a variant of [5, Theorem 3.2], up to a slight modification of the arguments.[6]

Theorem 2.4. *Let $\{\varphi_t,\, t \geq 0\}$ be a family of proper closed convex functions from H to $\mathbb{R} \cup \{+\infty\}$. Assume that*

(i) $\forall x \in H, \quad \varphi_t(x) \geq \varphi_t(\xi(t)) + \alpha(t)\, \|x - \xi(t)\|^2$;
(ii) $\int_0^{+\infty} \alpha(t)\, dt = +\infty$;
(iii) *the optimal path $\xi(\cdot)$ is locally absolutely continuous on $\mathbb{R}_+$ and satisfies* $\int_0^{+\infty} \|\dot{\xi}(t)\|\, dt < +\infty$.

If $x(\cdot)$ is a strong global solution of (NAGI)*, then $\lim_{t\to+\infty} \|x(t) - \xi(t)\| = 0$, and hence $\lim_{t\to+\infty} x(t) = \xi^*$ strongly in H, where ξ^* is the limit of the optimal path $\xi(t)$ as $t \to +\infty$.*

Proof. Consider the function k defined by $k(t) = \frac{1}{2}\|x(t) - \xi(t)\|^2$. This function is absolutely continuous, and for almost every $t \in]0, +\infty[$ we have

$$\begin{aligned} \dot{k}(t) &= \langle \dot{x}(t) - \dot{\xi}(t), x(t) - \xi(t)\rangle \\ &\leq \langle \dot{x}(t), x(t) - \xi(t)\rangle + \|\dot{\xi}(t)\|\|x(t) - \xi(t)\|. \end{aligned}$$

Since $-\dot{x}(t) \in \partial\varphi_t(x(t))$, we deduce from the subdifferential inequality that

$$\dot{k}(t) + \varphi_t(x(t)) - \varphi_t(\xi(t)) \leq \|\dot{\xi}(t)\|\|x(t) - \xi(t)\|.$$

Invoking assumption (i), we get

$$\dot{k}(t) + \alpha(t)\, \|x(t) - \xi(t)\|^2 \leq \|\dot{\xi}(t)\|\|x(t) - \xi(t)\|,$$

[6] A strong convexity property is required in the statement of [5, Theorem 3.2]. The strong convexity property is relaxed and replaced here with the strong minimum property (i).

or equivalently

$$\dot{k}(t) + 2\alpha(t)\, k(t) \leq \sqrt{2}\|\dot{\xi}(t)\|\sqrt{k(t)}.$$

The rest of the proof is analogous to that of [5, Theorem 3.2]. ☐

Let us now consider the case of an optimal trajectory satisfying $\|\dot{\xi}(t)\| = o(\alpha(t))$ as $t \to +\infty$; see [5, Theorem 3.3].

Theorem 2.5. *Under the assumptions* (*i*) *and* (*ii*) *of Theorem* 2.4*, assume moreover that the optimal path* $\xi(\cdot)$ *is locally absolutely continuous on* $\mathbb{R}_+$ *and that* $\lim_{t\to+\infty} \|\dot{\xi}(t)\|/\alpha(t) = 0$*. Let* $x(\cdot)$ *be a strong global solution of* (NAGI)*. Then* $\lim_{t\to+\infty} \|x(t) - \xi(t)\| = 0$*; therefore it converges strongly in* H *if and only if the optimal path* $\xi(t)$ *has a limit as* $t \to +\infty$*.*

For the proof of this result, the reader is referred to [5, Theorem 3.3].

3. Coupling with multiscale aspects: $A_t = A + \beta(t)B$ with $\beta(t) \to +\infty$

In this section, we specify our general ergodic convergence result to the case of a structured operator of the form $A_t = A + \beta(t)B$. The parameter $\beta(t)$ is assumed to tend to $+\infty$, thus leading to a two-scale problem.

3.1. **Case of general maximal monotone operators.**

Theorem 3.1. *Let* A*,* $B : H \rightrightarrows H$ *be two maximal monotone operators such that* $\operatorname{zer} B \neq \emptyset$ *and* $\operatorname{zer}(A + N_{\operatorname{zer} B}) \neq \emptyset$*. Assume that the operator* $A + N_{\operatorname{zer} B}$ *is maximal monotone. Given a map* $\beta : \mathbb{R}_+ \to \mathbb{R}_+^*$*, assume that the operator* $A + \beta(t)B$ *is maximal monotone for every* $t \geq 0$*. Suppose additionally that*

$$(\Sigma 4) \qquad \forall z \in \operatorname{zer} B, \quad \forall q \in N_{\operatorname{zer} B}(z), \qquad \int_0^{+\infty} \beta(t) G_B\left(z, \frac{q}{\beta(t)}\right) dt < +\infty.$$

Then every strong global solution $x(\cdot)$ *of the Multiscale Asymptotic Monotone Inclusion*

$$(\text{MAMI}) \qquad \dot{x}(t) + A(x(t)) + \beta(t)\, B(x(t)) \ni 0$$

converges weakly in average to some $x_\infty \in \operatorname{zer}(A + N_{\operatorname{zer} B})$*, i.e., as* $t \to +\infty$*,*

$$\frac{1}{t}\int_0^t x(s)\, ds \rightharpoonup x_\infty.$$

Remark 3.1. A particularly (new) interesting situation covered by the above theorem is the case $\beta(t) \to +\infty$. Indeed, let us assume that there exists $m > 0$ such that $\beta(t) \leq m$ for every $t \geq 0$. Fix $z \in \operatorname{zer} B$ and $q \in N_{\operatorname{zer} B}(z)$. Since $0 \in B(z)$, we have $G_B(z, 0) = 0$. The convexity of the function $p \mapsto G_B(z,p)$ then implies that

$$m\, G_B\left(z, \frac{q}{m}\right) \leq \beta(t) G_B\left(z, \frac{q}{\beta(t)}\right).$$

From formula (Σ4), we deduce that $G_B\left(z, \frac{q}{m}\right) = 0$, hence $\frac{q}{m} \in B(z)$. Since this is true for every $z \in \operatorname{zer} B$ and $q \in N_{\operatorname{zer} B}(z)$, we infer that the graph of $N_{\operatorname{zer} B}$ is included in the graph of B. By using the maximal monotonicity of the operator $N_{\operatorname{zer} B}$, we conclude that $B = N_{\operatorname{zer} B}$, a situation where the classical ergodic convergence theorem of Baillon-Brézis can be applied.

Remark 3.2. Denoting by F_B the Fitzpatrick function associated to the operator B, we have for every $q \in N_{\operatorname{zer} B}(z)$,

$$\begin{aligned} G_B\left(z, \frac{q}{\beta(t)}\right) &= F_B\left(z, \frac{q}{\beta(t)}\right) - \left\langle z, \frac{q}{\beta(t)}\right\rangle \\ &= F_B\left(z, \frac{q}{\beta(t)}\right) - \sigma_{\operatorname{zer} B}\left(\frac{q}{\beta(t)}\right). \end{aligned}$$

The last equality is an immediate consequence of the Fenchel extremality relation $\delta_{\operatorname{zer} B}(z) + \sigma_{\operatorname{zer} B}(q) = \langle z, q\rangle$. It ensues that condition (Σ4) can be equivalently rewritten as

(Σ5)

$$\forall z \in \operatorname{zer} B,\ \forall q \in N_{\operatorname{zer} B}(z), \quad \int_0^{+\infty} \beta(t)\left[F_B\left(z, \frac{q}{\beta(t)}\right) - \sigma_{\operatorname{zer} B}\left(\frac{q}{\beta(t)}\right)\right] dt < +\infty.$$

This last condition was recently used by Bot and Csetnek [16] as a generalization of condition (Σ6) below. The discrete version of this condition was introduced for the first time in [15].

As a consequence of Theorem 3.1, we recover the ergodic convergence result of Attouch and Czarnecki [6].

Corollary 3.1 ([6, Theorem 2.1(i)]). *Let $A : H \rightrightarrows H$ be a maximal monotone operator, let $\Psi : H \to \mathbb{R}_+ \cup \{+\infty\}$ be a proper closed convex function such that $C = \operatorname{argmin} \Psi = \Psi^{-1}(0) \neq \emptyset$, and let $\beta : \mathbb{R}_+ \to \mathbb{R}_+^*$ be a measurable function. Assume that $A + N_C$ is a maximal monotone operator with $\operatorname{zer}(A + N_C) \neq \emptyset$ and*

$$(\Sigma 6) \qquad \forall p \in \operatorname{ran}(N_C), \quad \int_0^{+\infty} \beta(t)\left[\Psi^*\left(\frac{p}{\beta(t)}\right) - \sigma_C\left(\frac{p}{\beta(t)}\right)\right] dt < +\infty.$$

Then, for every strong global solution trajectory $x(\cdot)$ of the differential inclusion

$$\dot{x}(t) + A(x(t)) + \beta(t)\,\partial\Psi(x(t)) \ni 0$$

there exists $x_\infty \in \operatorname{zer}(A + N_C)$ such that

$$w - \lim_{t \to +\infty} \frac{1}{t}\int_0^t x(s)ds = x_\infty.$$

Indeed, apply Theorem 3.1 with $B = \partial\Psi$. Recalling that

$$F_{\partial\Psi}\left(z, \frac{q}{\beta(t)}\right) \le \Psi(z) + \Psi^*\left(\frac{q}{\beta(t)}\right) = \Psi^*\left(\frac{q}{\beta(t)}\right),$$

condition (Σ6) implies condition (Σ5), which is in turn equivalent to (Σ4). Hence all the assumptions of Theorem 3.1 are fulfilled.

3.1.1. *Proof of Theorem* 3.1. Let us start with the following preliminary result.

Lemma 3.1. *Let A, $B : H \rightrightarrows H$ be two monotone operators. Then the following properties hold:*

(*i*) *For every $(z, p) \in H \times H$,*

$$G_{A+B}(z, p) \le \inf_{q \in H} G_A(z, q) + G_B(z, p - q).$$

(*ii*) *For every $(z, p) \in H \times H$ and every $\lambda > 0$, $G_{\lambda A}(z, p) = \lambda\, G_A(z, p/\lambda)$.*

(*iii*) *For every $z \in \overline{\operatorname{dom} A}$ and $p \in N_{\overline{\operatorname{dom} A}}(z)$, $G_A(z, p) \le G_A(z, 0)$.*

Proof. (i) Given $(z,p) \in H \times H$, the following inequality holds true:

$$F_{A+B}(z,p) \leq \inf_{q \in H} \{F_A(z,q) + F_B(z,p-q)\};$$

see for example [14, Proposition 4.2]. By subtracting $\langle z,p\rangle$ from each member, we immediately find the announced inequality.

(ii) Let $(z,p) \in H \times H$ and $\lambda > 0$. From the definition of $G_{\lambda A}(z,p)$, we have

$$\begin{aligned} G_{\lambda A}(z,p) &= \sup_{(y,q) \in \operatorname{gph}(\lambda A)} \langle z-y, q-p\rangle \\ &= \lambda \sup_{(y,q') \in \operatorname{gph} A} \langle z-y, q'-p/\lambda\rangle = \lambda\, G_A(z,p/\lambda). \end{aligned}$$

(iii) Fix $z \in \overline{\operatorname{dom} A}$ and $p \in N_{\overline{\operatorname{dom} A}}(z)$. For every $(y,q) \in \operatorname{gph} A$, we have

$$\begin{aligned} \langle z-y, q-p\rangle &= \langle z-y, q\rangle + \langle y-z, p\rangle \\ &\leq \langle z-y, q\rangle \quad \text{since } p \in N_{\overline{\operatorname{dom} A}}(z) \text{ and } y \in \operatorname{dom} A \\ &\leq G_A(z,0). \end{aligned}$$

Taking the supremum over $(y,q) \in \operatorname{gph} A$, we deduce that $G_A(z,p) \leq G_A(z,0)$. ☐

Let us now come back to the proof of Theorem 3.1. The main point consists in checking that the assumption $(\Sigma 1)$ of Theorem 2.1 is verified with $A_t = A + \beta(t)\, B$ and $A_\infty = A + N_{\operatorname{zer} B}$. Let $(z,p) \in \operatorname{gph}(A + N_{\operatorname{zer} B})$. Since $p \in Az + N_{\operatorname{zer} B}(z)$, there exists $q \in N_{\operatorname{zer} B}(z)$ such that $p - q \in Az$. Observe that

$$\begin{aligned} G_{A+\beta(t)B}(z,p) &\leq G_A(z,p-q) + G_{\beta(t)B}(z,q) && \text{in view of Lemma 3.1}(i), \\ &= G_{\beta(t)B}(z,q) && \text{since } (z,p-q) \in \operatorname{gph} A, \\ &= \beta(t)\, G_B(z,q/\beta(t)) && \text{in view of Lemma 3.1}(ii). \end{aligned}$$

The assumption $\int_0^{+\infty} \beta(t)\, G_B(z,q/\beta(t))\, dt < +\infty$ then implies that

$$\int_0^{+\infty} G_{A+\beta(t)B}(z,p)\, dt < +\infty.$$

It suffices now to apply Theorem 2.1.

3.2. **Coupled gradients with multiscale aspects.** Let us now consider the case $\varphi_t = \Phi + \beta(t)\Psi$, where the functions Φ, $\Psi : H \to \mathbb{R} \cup \{+\infty\}$ are proper closed convex and the parameter $\beta(t)$ tends to $+\infty$. The corresponding multiscale gradient dynamics reads as

$$\text{(MAG)} \qquad \dot{x}(t) + \partial\big(\Phi + \beta(t)\Psi\big)(x(t)) \ni 0.$$

Let us observe that $\partial\Phi + \beta(t)\partial\Psi \subset \partial\big(\Phi + \beta(t)\Psi\big)$ and that equality holds under some general qualification condition. Therefore, each trajectory of

$$\text{(MAG')} \qquad \dot{x}(t) + \partial\Phi(x(t)) + \beta(t)\,\partial\Psi(x(t)) \ni 0$$

satisfies (MAG). The solutions of (MAG) tend to minimize the function Φ over the set $C = \operatorname{argmin} \Psi$. If the parameter $\beta(t)$ tends *rather fast* to $+\infty$, then any trajectory converges weakly to a point of $\operatorname{argmin}_C \Phi$. Following [22], let us define the map $\omega : \mathbb{R}_+ \to \overline{\mathbb{R}}$ by

$$\omega(\varepsilon) = \inf_H (\Psi + \varepsilon\Phi), \tag{10}$$

for every $\varepsilon \geq 0$. Theorem 3.2 below shows that the map ω plays a crucial role in the asymptotic study of the dynamical system (MAG). A detailed study of the map ω will be carried out in section 4.

Theorem 3.2. *Assume that*

($\mathcal{H}_\Psi$) $\Psi : H \to \mathbb{R} \cup \{+\infty\}$ *is a proper closed convex function such that* $\inf_H \Psi = 0$ *and* $C =$ argmin $\Psi \neq \emptyset$.

($\mathcal{H}_\Phi$) $\Phi : H \to \mathbb{R} \cup \{+\infty\}$ *is a proper closed convex function such that* $\inf_C \Phi = 0$ *and* $\operatorname{argmin}_C \Phi \neq \emptyset$.

Assume that the set $\operatorname{argmin}_C \Phi$ *is bounded and that the function* $\Psi + \Phi$ *satisfies the inf-compactness property* (9). *Let* $\beta : \mathbb{R}_+ \to \mathbb{R}_+^*$ *be a map such that* $\lim_{t\to+\infty} \beta(t) = +\infty$. *Let* $x(\cdot)$ *be a strong global solution of* (MAG). *Then we have*

(*i*) $\lim_{t\to+\infty} d(x(t), \operatorname{argmin}_C \Phi) = 0$. *In particular, if the set* $\operatorname{argmin}_C \Phi$ *is a singleton* $\{\overline{x}\}$ *for some* $\overline{x} \in H$, *then* $x(t) \to \overline{x}$ *strongly in* H *as* $t \to +\infty$.

Additionally assume that

$$(\Sigma 7) \qquad \int_0^{+\infty} \beta(t)\, |\omega(1/\beta(t))|\, dt < +\infty.$$

Then

(*ii*) *there exists* $x_\infty \in \operatorname{argmin}_C \Phi$ *such that* $w - \lim_{t\to+\infty} x(t) = x_\infty$.

Proof. Let us check that the hypotheses of Theorem 2.3 are satisfied by $\varphi_t = \Phi + \beta(t)\Psi$. Taking $\varphi_\infty = \Phi + \delta_C$, we have $\operatorname{argmin} \varphi_\infty = \operatorname{argmin}_C \Phi$ and assumption (H1) is fulfilled. Now let $(x_k) \subset H$ and $(t_k) \subset \mathbb{R}_+$ be such that $x_k \to x_\infty$ and $t_k \to +\infty$ as $k \to +\infty$. Let us fix $m > 0$. Since $\lim_{k\to+\infty} \beta(t_k) = +\infty$, we have $\beta(t_k) \geq m$ for k large enough and hence

$$\liminf_{k\to+\infty} (\Phi(x_k) + \beta(t_k)\Psi(x_k)) \geq \liminf_{k\to+\infty} (\Phi(x_k) + m\Psi(x_k)).$$

Recalling that $x_k \to x_\infty$ and that the functions Φ and Ψ are closed, we deduce that

$$\liminf_{k\to+\infty} (\Phi(x_k) + \beta(t_k)\Psi(x_k)) \geq \Phi(x_\infty) + m\Psi(x_\infty).$$

Letting $m \to +\infty$, we infer that

$$\liminf_{k\to+\infty} (\Phi(x_k) + \beta(t_k)\Psi(x_k)) \geq \Phi(x_\infty) + \delta_C(x_\infty);$$

hence (H2) is fulfilled. For every $z \in \operatorname{argmin}_C \Phi$, we have $\varphi_t(z) = 0$; therefore $v_\infty(t) = 0$ for every $t \geq 0$, and (H3) is trivially satisfied. Since $\beta(t) \to +\infty$ we have $\Phi + \Psi \leq \Phi + \beta(t)\Psi = \varphi_t$ for t large enough, and assumption (H4) is satisfied with $f = \Phi + \Psi$. Now observe that for every $t \geq 0$ and $z \in \operatorname{argmin}_C \Phi$,

$$\begin{aligned}
\varphi_t(z) - \inf_H \varphi_t &= -\inf_H (\Phi + \beta(t)\Psi) \quad \text{since } \Phi(z) = \Psi(z) = 0 \\
&= -\beta(t)\, \omega(1/\beta(t)) \quad \text{by definition of the map } \omega \\
&= \beta(t)\, |\omega(1/\beta(t))| \quad \text{because } \omega \leq 0.
\end{aligned}$$

In view of condition (Σ7), condition (Σ3) is clearly satisfied, thus implying (Σ2). Conclusions (*i*)–(*ii*) then follow from Theorem 2.3. ∎

Remark 3.3. In the context of the previous theorem, one can easily show that[7]

$$\lim_{t\to+\infty} \Psi(x(t)) = 0; \tag{11}$$

see for example [6, Lemma 3.3]. Hence there exists $t_0 \geq 0$ such that $\Psi(x(t)) \leq 1$ for every $t \geq t_0$. Since the trajectory $x(\cdot)$ is bounded, there exists $R > 0$ such that $\|x(t)\| \leq R$ for every $t \geq 0$. If Ψ satisfies the inf-compactness property (9), we deduce that the set $\{x(t),\, t \geq t_0\}$ is relatively compact for the strong topology of H. Recalling from Theorem 3.2(*ii*) that the trajectory $x(\cdot)$ weakly converges to x_∞, we immediately deduce that it converges strongly to x_∞.

Remark 3.4. Assume that the function Ψ satisfies the quadratic conditioning property

$$\Psi \geq a\, d^2(\cdot, C) \quad \text{for some } a > 0.$$

Under this condition, there exists $c > 0$ such that $|\omega(\varepsilon)| \leq c\,\varepsilon^2$ for every $\varepsilon \geq 0$; see section 4. Hence, in this case, assumption $(\Sigma 7)$ is fulfilled if $\int_0^{+\infty} (1/\beta(t))\, dt < +\infty$.

Corollary 3.2. *Under hypotheses* $(\mathcal{H}_\Psi)$-$(\mathcal{H}_\Phi)$, *assume that the set* $S = \operatorname{argmin}\Psi \cap \operatorname{argmin}\Phi$ *is nonempty and bounded. Suppose that the function* $\Psi + \Phi$ *satisfies the inf-compactness property* (9). *Let* $\beta : \mathbb{R}_+ \to \mathbb{R}_+^*$ *be a map that satisfies* $\lim_{t\to+\infty} \beta(t) = +\infty$. *Let* $x(\cdot)$ *be a strong global solution of* (MAG). *Then there exists* $x_\infty \in S$ *such that* $x(t) \rightharpoonup x_\infty$ *weakly in* H *as* $t \to +\infty$.

Proof. If $\operatorname{argmin}\Psi \cap \operatorname{argmin}\Phi \neq \emptyset$, the infimum in the definition of $\omega(\varepsilon)$ is attained at every $x \in \operatorname{argmin}\Psi \cap \operatorname{argmin}\Phi$, and it equals 0. It ensues that $\omega(\varepsilon) = 0$ for every $\varepsilon \geq 0$. Therefore condition $(\Sigma 7)$ of Theorem 3.2 is automatically satisfied. $\square$

As a consequence of Theorem 3.2, we recover the convergence result of the trajectories of (MAG) from [6].

Corollary 3.3 ([6, Theorem 5.1])**.** *Let* $\Psi, \Phi : H \to \mathbb{R} \cup \{+\infty\}$ *be functions satisfying hypotheses* $(\mathcal{H}_\Psi)$–$(\mathcal{H}_\Phi)$, *together with the following qualification condition:*

(QC) *there exists* $x_0 \in C \cap \operatorname{dom}\Phi$ *such that* Φ *is continuous at* x_0.

Assume that the set $\operatorname{argmin}_C\Phi$ *is bounded and that the function* $\Psi + \Phi$ *satisfies the inf-compactness property* (9). *Let* $\beta : \mathbb{R}_+ \to \mathbb{R}_+^*$ *be a map such that* $\lim_{t\to+\infty} \beta(t) = +\infty$. *Assume moreover that*

$$\forall p \in \operatorname{ran}(N_C), \quad \int_0^{+\infty} \beta(t) \left[\Psi^*\left(\frac{p}{\beta(t)}\right) - \sigma_C\left(\frac{p}{\beta(t)}\right)\right] dt < +\infty. \tag{$\Sigma 6$}$$

Let $x(\cdot)$ *be a strong global solution*[8] *of* (MAG). *Then there exists* $x_\infty \in \operatorname{argmin}_C\Phi$ *such that* $x(t) \rightharpoonup x_\infty$ *weakly in* H *as* $t \to +\infty$.

Proof. It relies on the study of the map ω that we carry out in section 4. Precisely, it is a consequence of the forthcoming Proposition 4.2(*d*). $\square$

[7] Equality (11) is a basic result which requires neither condition $(\Sigma 7)$ nor the inf-compactness of $\Phi + \Psi$.

[8] In view of the qualification condition (QC), the equality $\partial\Phi + \beta(t)\partial\Psi = \partial\big(\Phi + \beta(t)\Psi\big)$ holds for every $t \geq 0$; hence the differential inclusions (MAG) and (MAG') coincide.

4. Infimum value associated to the viscosity problem $\inf_H(\Psi + \varepsilon\Phi)$

As we have already pointed out, the map $\varepsilon \mapsto \omega(\varepsilon) = \inf_H(\Psi + \varepsilon\Phi)$ plays a crucial role in the asymptotic study of the dynamic system (MAG). We now make a systematic study of this function. Throughout this section, we assume $(\mathcal{H}_\Psi)$ and $(\mathcal{H}_\Phi)$, i.e.,

$(\mathcal{H}_\Psi)$ $\Psi : H \to \mathbb{R} \cup \{+\infty\}$ is a proper closed convex function such that $\inf_H \Psi = 0$ and $C = \operatorname{argmin} \Psi \neq \emptyset$.

$(\mathcal{H}_\Phi)$ $\Phi : H \to \mathbb{R} \cup \{+\infty\}$ is a proper closed convex function such that $\inf_C \Phi = 0$ and $S = \operatorname{argmin}_C \Phi \neq \emptyset$.

For every $\varepsilon \geq 0$, we denote by $(\mathcal{P}_\varepsilon)$ the minimization problem

$$(\mathcal{P}_\varepsilon) \qquad \inf_{x\in H} \{\Psi(x) + \varepsilon\,\Phi(x)\},$$

so we have $\omega(\varepsilon) = \inf \mathcal{P}_\varepsilon$.

Remark 4.1. Assumption $(\mathcal{H}_\Phi)$ implies that the domain of Φ intersects the set C of minimizers of Ψ. This corresponds to a regular perturbation situation, where we can expect a simple asymptotic development for $\omega(\varepsilon)$ as ε goes to zero, as well as the convergence of the filtered sequence of solutions of $(\mathcal{P}_\varepsilon)$ to a solution of the hierarchical minimization problem $\min_C \Phi$. That is the situation we consider. In contrast, when the domain of Φ does not intersect the set $C = \operatorname{argmin} \Psi$, we are faced with a singular perturbation. This is a more involved situation, which one encounters for example in phase transition, when considering the Van der Waals-Cahn-Hilliard viscous approximation of the Gibbs free energy. In this case, we must appeal to Γ-convergence methods for rescaled energy functions; see [3], [4, Chap. 12.5], [39].

4.1. General properties of ω.

The following proposition gathers properties of the map ω.

Proposition 4.1. *Assume hypotheses* $(\mathcal{H}_\Psi)$–$(\mathcal{H}_\Phi)$.

(a) *The map* $\varepsilon \mapsto \omega(\varepsilon)$ *is nonpositive, nonincreasing and concave on* $\mathbb{R}_+$.

Assume moreover that the function $\Psi + \Phi$ *is coercive.*[9] *Then*

(b) *for every* $\varepsilon \in [0, 1]$, *we have* $\omega(\varepsilon) > -\infty$, *and the infimum is attained in the definition of* $\omega(\varepsilon)$.

(c) $\lim_{\varepsilon\to 0^+} \omega(\varepsilon)/\varepsilon = 0$. *In other words, the following asymptotic expansion holds*[10] *as* $\varepsilon \to 0$:

$$\min_H(\Psi + \varepsilon\Phi) = \min_H \Psi + \varepsilon \min_C \Phi + o(\varepsilon). \tag{12}$$

Proof. (a) Given $z \in S$, we have

$$\omega(\varepsilon) \leq \Psi(z) + \varepsilon\Phi(z) = 0,$$

hence $\omega(\varepsilon) \leq 0$ for every $\varepsilon \geq 0$. Observe that the map $\varepsilon \mapsto \Psi(x) + \varepsilon\Phi(x)$ is affine; hence the map $\varepsilon \mapsto \omega(\varepsilon)$ is concave as an infimum of affine functions. Since the

---

[9] The coercivity of $\Psi + \Phi$ implies that of $\delta_C + \Phi$, and we deduce classically that $\operatorname{argmin}_C \Phi = \operatorname{argmin}(\delta_C + \Phi) \neq \emptyset$.

[10] For simplicity, we assumed $\min_H \Psi = \min_C \Phi = 0$. The statement remains valid without any assumption on the (finite) values of $\min_H \Psi$ and $\min_C \Phi$. The asymptotic expansion (12) can be found in [3, Theorem 2.5].

function $\omega : \mathbb{R}_+ \to \mathbb{R} \cup \{-\infty\}$ is concave, it admits a right (resp. left) derivative at every $t \geq 0$ (resp. $t > 0$). In particular, we have

$$\omega'_+(0) = \lim_{\varepsilon \to 0^+} \frac{1}{\varepsilon}(\omega(\varepsilon) - \omega(0)) \leq 0,$$

since $\omega(0) = 0$, and $\omega(\varepsilon) \leq 0$ for every $\varepsilon > 0$. The concavity of ω implies that $\omega'_+(\varepsilon) \leq 0$ (resp. $\omega'_-(\varepsilon) \leq 0$) for every $\varepsilon > 0$. We deduce that the function ω is nonincreasing on $\mathbb{R}_+$.

(b) First observe that the conclusion is immediate for $\varepsilon = 0$. Now assume that $\varepsilon \in]0, 1]$. Since $\Psi(x) \geq 0$, we have

$$\Psi(x) + \varepsilon\, \Phi(x) \geq \varepsilon\, (\Psi(x) + \Phi(x)).$$

From the coercivity of $\Psi + \Phi$, we deduce that the lower semicontinuous convex function $x \mapsto \Psi(x) + \varepsilon\, \Phi(x)$ is coercive. It ensues classically that the minimization problem $(\mathcal{P}_\varepsilon)$ has at least one solution and that $\omega(\varepsilon) = \inf \mathcal{P}_\varepsilon > -\infty$.

(c) Let us argue by contradiction and assume that there exist $\eta > 0$ and a sequence (ε_n) tending toward 0 such that $\omega(\varepsilon_n)/\varepsilon_n \leq -\eta$. From the definition of $\omega(\varepsilon_n)$, there exists a sequence (x_n) in H such that

$$\forall n \in \mathbb{N}, \quad \Psi(x_n) + \varepsilon_n\, \Phi(x_n) \leq -\frac{\eta}{2}\, \varepsilon_n. \tag{13}$$

Since $\lim_{n\to+\infty} \varepsilon_n = 0$ and $\Psi(x_n) \geq 0$, we have $\varepsilon_n\, \Psi(x_n) \leq \Psi(x_n)$ for n large enough, say $n \geq n_0$. In view of (13), this implies that for every $n \geq n_0$,

$$\Psi(x_n) + \Phi(x_n) \leq -\frac{\eta}{2}, \tag{14}$$

or equivalently

$$x_n \in \left[\Psi + \Phi \leq -\frac{\eta}{2}\right].$$

Recalling that the function $\Psi + \Phi$ is coercive by assumption, we deduce that the sequence (x_n) is bounded in H. Therefore there exist $x_\infty \in H$ and a subsequence of (x_n), still denoted by (x_n), that converges weakly to x_∞ in H. Since Φ is closed and convex, it has a continuous affine minorant. Hence there exist $a \in \mathbb{R}$ and $p \in H$ such that $\Phi(x) \geq a + \langle p, x\rangle$ for every $x \in H$. By using inequality (13), we infer that

$$\Psi(x_n) \leq -\varepsilon_n \left[\frac{\eta}{2} + a + \langle p, x_n\rangle\right].$$

Taking the upper limit when $n \to +\infty$, we find that

$$\limsup_{n\to+\infty} \Psi(x_n) \leq 0. \tag{15}$$

On the other hand, since $\Psi(x_n) \geq 0$, we infer from (14) that

$$\limsup_{n\to+\infty} \Phi(x_n) \leq -\frac{\eta}{2}. \tag{16}$$

From the closedness of Ψ (resp. Φ) with respect to the weak topology in H and inequality (15) (resp. (16)), we deduce respectively that

$$\Psi(x_\infty) \leq \liminf_{n\to+\infty} \Psi(x_n) \leq \limsup_{n\to+\infty} \Psi(x_n) \leq 0,$$

$$\Phi(x_\infty) \leq \liminf_{n\to+\infty} \Phi(x_n) \leq \limsup_{n\to+\infty} \Phi(x_n) \leq -\frac{\eta}{2}.$$

The first inequality implies that $x_\infty \in C$, and the second one gives the contradiction. $\square$

By using the duality theory, we are going to prove that the behavior of the map $\varepsilon \mapsto \omega(\varepsilon)$ can be interpreted with the conjugates of Ψ and Φ. Let us first recall the following general theorem; see for example [24, Theorem 4.1 p. 58].

Theorem 4.1. *Given two normed spaces V and Y, let $F : V \to \mathbb{R} \cup \{+\infty\}$ and $G : Y \to \mathbb{R} \cup \{+\infty\}$ be proper closed convex functions, and let $L \in \mathcal{L}(V, Y)$. Consider the primal problem*

$$(\mathcal{P}) \qquad \inf_{u \in V} \{F(u) + G(Lu)\}$$

and the dual problem

$$(\mathcal{P}^*) \qquad \sup_{p^* \in Y^*} \{-F^*(L^* p^*) - G^*(-p^*)\}.$$

Then we have $\sup \mathcal{P}^* \leq \inf \mathcal{P}$. *If moreover* $\inf \mathcal{P}$ *is finite and if there exists* $u_0 \in \operatorname{dom} F$ *such that G is continuous at* $Lu_0 \in \operatorname{dom} G$, *then* $\inf \mathcal{P} = \sup \mathcal{P}^*$, *and* $(\mathcal{P}^*)$ *has at least one solution.*

Proposition 4.2. *Assume hypotheses* $(\mathcal{H}_\Psi)$–$(\mathcal{H}_\Phi)$.

(a) *For every $\varepsilon \geq 0$, we have*

$$|\omega(\varepsilon)| \leq \inf_{p \in H} \{\Psi^*(\varepsilon\, p) + \varepsilon\, \Phi^*(-p)\}. \tag{17}$$

(b) *Letting $\varepsilon \geq 0$, assume $\omega(\varepsilon) > -\infty$ and the following qualification condition:*

(QC') *there exists $x_0 \in \operatorname{dom} \Psi \cap \operatorname{dom} \Phi$ such that Φ is continuous at x_0.*

Then we have

$$|\omega(\varepsilon)| = \min_{p \in H} \{\Psi^*(\varepsilon\, p) + \varepsilon\, \Phi^*(-p)\}. \tag{18}$$

(c) *Assume the qualification condition* (QC):[11]

(QC) *there exists $x_0 \in C \cap \operatorname{dom} \Phi$ such that Φ is continuous at x_0.*

Then there exists $p \in \operatorname{ran}(N_C)$ such that, for every $\varepsilon \geq 0$,

$$|\omega(\varepsilon)| \leq \Psi^*(\varepsilon\, p) - \sigma_C(\varepsilon\, p). \tag{19}$$

(d) *Assume* (QC). *Then condition* $(\Sigma 6)$ *implies condition* $(\Sigma 7)$.

Proof. (a) Let us apply Theorem 4.1 with $V = Y = H$, $F = \Psi$, $G = \varepsilon\, \Phi$ and $L = \mathrm{Id}_H$. The primal minimization problem $(\mathcal{P}_\varepsilon)$ reads as

$$(\mathcal{P}_\varepsilon) \qquad \inf_{x \in H} \{\Psi(x) + \varepsilon\, \Phi(x)\}.$$

For every $\varepsilon > 0$, the dual problem is

$$(\mathcal{P}^*_\varepsilon) \qquad \sup_{p \in H} \{-\Psi^*(p) - \varepsilon\, \Phi^*(-p/\varepsilon)\}.$$

From the general relation $\sup \mathcal{P}^*_\varepsilon \leq \inf \mathcal{P}_\varepsilon$, we deduce that

$$|\omega(\varepsilon)| = -\omega(\varepsilon) \leq \inf_{p \in H} \{\Psi^*(p) + \varepsilon\, \Phi^*(-p/\varepsilon)\}.$$

Replacing p with $\varepsilon\, p$, we immediately obtain inequality (17). This inequality trivially holds true for $\varepsilon = 0$; hence it is valid for every $\varepsilon \geq 0$.

[11] Notice that (QC) is slightly stronger than (QC').

(b) Since condition (QC') is satisfied, Theorem 4.1 shows that $\inf \mathcal{P}_\varepsilon = \sup \mathcal{P}_\varepsilon^*$ and that $(\mathcal{P}_\varepsilon^*)$ has at least one solution. This implies that

$$|\omega(\varepsilon)| = \min_{p\in H} \{\Psi^*(p) + \varepsilon\, \Phi^*(-p/\varepsilon)\}\,.$$

Equality (18) follows immediately.

(c) Given $\overline{x} \in S = \operatorname{argmin}_C \Phi$, we have $0 \in \partial(\Phi + \delta_C)(\overline{x})$. The qualification condition (QC) implies $\partial(\Phi + \delta_C)(\overline{x}) = \partial\Phi(\overline{x}) + N_C(\overline{x})$. We deduce that $0 \in \partial\Phi(\overline{x}) + N_C(\overline{x})$, whence the existence of $p \in N_C(\overline{x}) \cap (-\partial\Phi(\overline{x}))$. For every $\varepsilon \geq 0$, let us write that

$$\begin{aligned}\Psi^*(\varepsilon p) + \varepsilon\, \Phi^*(-p) \quad &= \quad [\Psi^*(\varepsilon p) - \sigma_C(\varepsilon p)] + \varepsilon\, [\sigma_C(p) + \delta_C(\overline{x}) - \langle p, \overline{x}\rangle] \\ &\qquad +\varepsilon\, [\Phi^*(-p) + \Phi(\overline{x}) + \langle p, \overline{x}\rangle]\,.\end{aligned}$$

Since $p \in N_C(\overline{x})$ and $-p \in \partial\Phi(\overline{x})$, the Fenchel extremality relation shows that the second and third brackets are equal to zero. This implies that, for every $\varepsilon \geq 0$,

$$\Psi^*(\varepsilon p) + \varepsilon\, \Phi^*(-p) = \Psi^*(\varepsilon p) - \sigma_C(\varepsilon p).$$

Inequality (19) then immediately follows from (17).

(d) It follows from (c) and the statement of conditions $(\Sigma 6)$ and $(\Sigma 7)$. □

Remark 4.2. The qualification condition (QC) may be slightly weakened in the statement of Proposition 4.2, items (c)–(d). It suffices to assume that the operator $\partial\Phi + N_C$ is maximal monotone. The same remark applies to the statement of Corollary 3.3, as was observed in [6, Theorem 5.1].

4.2. **Illustrating examples.** We now review several examples for which we are able to majorize explicitly the function $\Psi^* - \sigma_C$. This yields sufficient conditions for $(\Sigma 6)$, and hence for $(\Sigma 7)$ in view of Proposition 4.2(d).

Example 4.1. Let $\Psi : H \to \mathbb{R} \cup \{+\infty\}$ be a proper closed convex function such that $C = \operatorname{argmin} \Psi \neq \emptyset$. Suppose that for every $x \in H$,

$$\Psi(x) \geq \theta(d(x, C)),$$

where the closed convex function $\theta : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ is even[12] and such that $\theta(0) = 0$. Then we have for every $\varepsilon \geq 0$ and $p \in H$,

$$\Psi^*(\varepsilon\, p) - \sigma_C(\varepsilon\, p) \leq \theta^*(\varepsilon\, \|p\|). \tag{20}$$

Proof. From a classical result, the conjugate of the function $\theta(d(., C))$ is the function $\theta^*(\|\,.\,\|) + \sigma_C$; see for example [10, Exercise IV.17]. It ensues that $\Psi^* \leq \theta^*(\|\,.\,\|) + \sigma_C$, and the conclusion follows immediately. □

Under the assumptions of Example 4.1, the key condition $(\Sigma 6)$ of Corollary 3.3 is satisfied if for every $p \in H$,

$$\int_0^{+\infty} \beta(t)\, \theta^*(\|p\|/\beta(t))\, dt < +\infty.$$

[12]The assumptions on θ automatically imply that $0 \in \operatorname{argmin} \theta$.

Remark 4.3. Assume that there exists $a > 0$ such that $\Psi(x) \geq a\, d(x, C)$ for every $x \in H$. By applying the above proposition with $\theta(t) = a\,|t|$, we find $\Psi^*(\varepsilon\, p) - \sigma_C(\varepsilon\, p) \leq \delta_{[-a,a]}(\varepsilon\,|p|)$, and hence $\Psi^*(\varepsilon\, p) - \sigma_C(\varepsilon\, p) = 0$ for ε small enough. In this case, condition $(\Sigma 6)$ is automatically satisfied.

Remark 4.4. Assume that there exist $a > 0$ and $r > 1$ such that

$$\Psi(x) \geq a\, d^r(x, C), \tag{21}$$

for every $x \in H$. Let us apply the above proposition with the function $\theta : \mathbb{R} \to \mathbb{R}$ defined by $\theta(t) = a\,|t|^r$. Since $(|\,.\,|^r/r)^* = (|\,.\,|^{r^*}/r^*)$, where r^* is the conjugate exponent of r, *i.e.*, $r^* = 1/(1 - 1/r)$, we easily obtain

$$\theta^*(t) = \frac{(ar)^{1-r^*}}{r^*}\,|t|^{r^*}.$$

In view of (20), we infer that $\Psi^*(\varepsilon\, p) - \sigma_C(\varepsilon\, p) \leq \frac{(ar)^{1-r^*}}{r^*}\,(\varepsilon\,\|p\|)^{r^*}$. In this case, condition $(\Sigma 6)$ is satisfied as soon as

$$\int_0^{+\infty} (1/\beta(t))^{r^*-1}\, dt < +\infty.$$

Example 4.2. Let $L \in \mathcal{L}(H)$ and let $\Psi : H \to \mathbb{R} \cup \{+\infty\}$ be a proper closed convex function such that $C = \operatorname{argmin} \Psi = \ker L$. Suppose that $\Psi(x) \geq \frac{1}{2}\|Lx\|^2$ for all $x \in H$. Then we have for every $\varepsilon \geq 0$ and $p \in \operatorname{ran}(L^*)$,

$$\Psi^*(\varepsilon\, p) - \sigma_C(\varepsilon\, p) \leq \frac{\varepsilon^2}{2}\, d^2\left(0, (L^*)^{-1}(p)\right).$$

Proof. By applying Theorem 4.1, we can show[13] that the conjugate of the function $x \mapsto \frac{1}{2}\|Lx\|^2$ is given by

$$p \mapsto \begin{cases} \frac{1}{2} d^2\left(0, (L^*)^{-1}(p)\right) & \text{if } p \in \operatorname{ran}(L^*), \\ +\infty & \text{if } p \notin \operatorname{ran}(L^*). \end{cases}$$

It ensues that for every $p \in \operatorname{ran}(L^*)$,

$$\Psi^*(p) \leq \frac{1}{2} d^2\left(0, (L^*)^{-1}(p)\right). \tag{22}$$

On the other hand, since the set $\ker L$ is a subspace of H, we have

$$\sigma_{\ker L} = (\delta_{\ker L})^* = \delta_{(\ker L)^\perp}. \tag{23}$$

Recalling that $\operatorname{ran}(L^*) \subset (\ker L)^\perp$, we deduce from (22) and (23) that for every $\varepsilon \geq 0$ and $p \in \operatorname{ran}(L^*)$,

$$\Psi^*(\varepsilon\, p) - \sigma_{\ker L}(\varepsilon\, p) \leq \frac{\varepsilon^2}{2} d^2\left(0, (L^*)^{-1}(p)\right). \qquad \square$$

[13] The details are left to the reader.

Under the assumptions of Example 4.2, the key condition ($\Sigma 6$) of Corollary 3.3 is satisfied if

$$\int_0^{+\infty} 1/\beta(t)\, dt < +\infty.$$

5. Coupling with multiscale aspects: $A_t = A + \varepsilon(t)B$ with $\varepsilon(t) \to 0$

5.1. **Case of general maximal monotone operators.** By reversing the roles of the operators A and B and by using a suitable time rescaling, we obtain the following version of Theorem 3.1.

Theorem 5.1. *Let A, $B : H \rightrightarrows H$ be two maximal monotone operators such that $\operatorname{zer} A \neq \emptyset$ and $\operatorname{zer}(B + N_{\operatorname{zer} A}) \neq \emptyset$. Assume that the operator $B + N_{\operatorname{zer} A}$ is maximal monotone. Given a map $\varepsilon : \mathbb{R}_+ \to \mathbb{R}_+^*$, assume that the operator $A + \varepsilon(t)B$ is maximal monotone for every $t \geq 0$. Suppose additionally that $\int_0^{+\infty} \varepsilon(t)\, dt = +\infty$ and that*

$$\forall z \in \operatorname{zer} A, \quad \forall q \in N_{\operatorname{zer} A}(z), \qquad \int_0^{+\infty} G_A\left(z, \varepsilon(t)\, q\right)\, dt < +\infty. \tag{24}$$

Then for every strong global solution $x(\cdot)$ of

$$\dot{x}(t) + A(x(t)) + \varepsilon(t)\, B(x(t)) \ni 0, \tag{MAMIε}$$

there exists $x_\infty \in \operatorname{zer}(B + N_{\operatorname{zer} A})$ such that

$$\frac{1}{t}\int_0^t x(s)\, ds \rightharpoonup x_\infty \text{ weakly in } H, \text{ as } t \to +\infty.$$

Proof. It is done by a time rescaling, following [6]. Let us rewrite the dynamical system (MAMIε) as

$$\frac{1}{\varepsilon(t)}\dot{x}(t) + B(x(t)) + \frac{1}{\varepsilon(t)}A(x(t)) \ni 0.$$

Then use the time rescaling $s = \sigma(t) = \int_0^t \varepsilon(u)\, du$. Define $y(\cdot)$ and $\alpha(\cdot)$ by $y(s) = x(\sigma^{-1}(s))$ and $\alpha(s) = 1/\varepsilon(\sigma^{-1}(s))$. We then have $\dot{y}(s) = \dot{x}(t)/\varepsilon(t)$, so that $y(\cdot)$ satisfies the differential inclusion

$$\dot{y}(s) + B(y(s)) + \alpha(s)A(y(s)) \ni 0.$$

In terms of the variable s, condition (24) can be translated as

$$\forall z \in \operatorname{zer} A, \quad \forall q \in N_{\operatorname{zer} A}(z), \qquad \int_0^{+\infty} \alpha(s)\, G_A\left(z, \frac{q}{\alpha(s)}\right)\, ds < +\infty.$$

The assumptions of Theorem 3.1 are satisfied after reversing the roles of the operators A and B. The conclusion follows immediately. $\square$

Condition $\int_0^{+\infty} \varepsilon(t)\, dt = +\infty$ expresses that $\varepsilon(t)$ does not tend too fast toward zero as $t \to +\infty$. On the other hand, condition (24) prevents the parameter $\varepsilon(t)$ from converging very slowly toward zero. Hence the conditions in Theorem 5.1 imply a moderately slow convergence $\varepsilon(t) \to 0$ as $t \to +\infty$. Let us now analyze the case $\int_0^{+\infty} \varepsilon(t)\, dt < +\infty$ corresponding to a fast decaying parameter.

Corollary 5.1. *Let A, $B : H \rightrightarrows H$ be two maximal monotone operators such that $A + N_{\overline{\mathrm{dom}\, B}}$ is maximal monotone and* $\mathrm{zer}\,(A + N_{\overline{\mathrm{dom}\, B}}) \neq \emptyset$*. Given a map $\varepsilon : \mathbb{R}_+ \to \mathbb{R}_+^*$, assume that the operator $A + \varepsilon(t)B$ is maximal monotone for every $t \geq 0$. Suppose additionally that $\int_0^{+\infty} \varepsilon(t)\, dt < +\infty$ and that $G_B(z, 0) < +\infty$ for every $z \in \mathrm{dom}\, A \cap \overline{\mathrm{dom}\, B}$. Then for every strong global solution $x(\cdot)$ of*

$$\text{(MAMI}\varepsilon) \qquad \dot{x}(t) + A(x(t)) + \varepsilon(t)\, B(x(t)) \ni 0,$$

there exists $x_\infty \in \mathrm{zer}\,(A + N_{\overline{\mathrm{dom}\, B}})$ such that $\frac{1}{t}\int_0^t x(s)\, ds \rightharpoonup x_\infty$ weakly in H, as $t \to +\infty$.

Proof. The main point consists of checking that the assumption ($\Sigma 1$) of Theorem 2.1 is verified with $A_t = A + \varepsilon(t)\, B$ and $A_\infty = A + N_{\overline{\mathrm{dom}\, B}}$. The details are left to the reader. □

5.2. **Case of subdifferential operators.** Let us start with a fast vanishing coefficient $\varepsilon(t) \to 0$.

Corollary 5.2. *Let Ψ, $\Phi : H \to \mathbb{R} \cup \{+\infty\}$ be closed convex functions such that $\mathrm{dom}\, \Psi \cap \mathrm{dom}\, \Phi \neq \emptyset$. Assume that Ψ is nonnegative. Let $\varepsilon : \mathbb{R}_+ \to \mathbb{R}_+$ be a nonincreasing map such that $\lim_{t \to +\infty} \varepsilon(t) = 0$. Let $x(\cdot)$ be a strong global solution of*

$$(\mathrm{MAG}_\varepsilon) \qquad \dot{x}(t) + \partial\big(\Phi + \varepsilon(t)\Psi\big)(x(t)) \ni 0,$$

such that the function $t \mapsto \Phi(x(t)) + \varepsilon(t)\Psi(x(t))$ is locally absolutely continuous. Then we have:

- (*i*) *The function $t \mapsto \Phi(x(t)) + \varepsilon(t)\Psi(x(t))$ is nonincreasing and tends toward $\inf_H(\Phi + \delta_{\mathrm{dom}\, \Psi})$ as $t \to +\infty$.*

Additionally assume that $\inf_H(\Phi + \delta_{\mathrm{dom}\, \Psi}) > -\infty$; then

- (*ii*) $\int_0^{+\infty} \|\dot{x}(t)\|^2\, dt < +\infty$.

Assume moreover that the set $S_\infty = \mathrm{argmin}\,(\mathrm{cl}(\Phi + \delta_{\mathrm{dom}\, \Psi}))$ is not empty and included in $\mathrm{dom}\, \Psi$. If $\int_0^{+\infty} \varepsilon(t)\, dt < +\infty$, then

- (*iii*) *there exists $x_\infty \in S_\infty$ such that $w - \lim_{t \to +\infty} x(t) = x_\infty$.*

Proof. It suffices to check that the assumptions of Theorem 2.2 are satisfied with $\varphi_t = \Phi + \varepsilon(t)\Psi$ and $\varphi_\infty = \mathrm{cl}(\Phi + \delta_{\mathrm{dom}\, \Psi})$. The proof is easy and therefore omitted. □

Let us now consider the case of a slowly vanishing coefficient $\varepsilon(t)$ satisfying $\int_0^{+\infty} \varepsilon(t)\, dt = +\infty$. The following result is obtained by reversing the roles of the functions Φ and Ψ in Theorem 3.2 and by using a suitable time rescaling, which allows us to pass from $\beta(t) \to +\infty$ to $\varepsilon(t) \to 0$.

Corollary 5.3. *Let $\Phi : H \to \mathbb{R} \cup \{+\infty\}$ be a proper closed convex function such that $C = \mathrm{argmin}\, \Phi \neq \emptyset$ and $\min_H \Phi = 0$. Let $\Psi : H \to \mathbb{R} \cup \{+\infty\}$ be a proper closed convex function such that $\mathrm{argmin}_C \Psi \neq \emptyset$ and $\min_C \Psi = 0$. Assume that the set $\mathrm{argmin}_C \Psi$ is bounded and that the function $\Phi + \Psi$ satisfies the inf-compactness property* (9)*. Let $\varepsilon : \mathbb{R}_+ \to \mathbb{R}_+^*$ be a map such that $\lim_{t \to +\infty} \varepsilon(t) = 0$ and $\int_0^{+\infty} \varepsilon(t)\, dt = +\infty$. Let $x(\cdot)$ be a strong global solution of*

$$(\mathrm{MAG}_\varepsilon) \qquad \dot{x}(t) + \partial\big(\Phi + \varepsilon(t)\Psi\big)(x(t)) \ni 0.$$

Then we have

(i) $\lim_{t\to+\infty} d(x(t), \operatorname{argmin}_C \Psi) = 0$. *In particular, if the set* $\operatorname{argmin}_C \Psi$ *is a singleton* $\{\overline{x}\}$ *for some* $\overline{x} \in H$, *then* $x(t) \to \overline{x}$ *strongly in* H *as* $t \to +\infty$.

Additionally assume that

$$\int_0^{+\infty} |\omega(\varepsilon(t))|\, dt < +\infty,$$

where the map $\omega : \mathbb{R}_+ \to \mathbb{R}$ *is defined by* $\omega(\varepsilon) = \inf_H(\Phi + \varepsilon\Psi)$. *Then*

(ii) *there exists* $x_\infty \in \operatorname{argmin}_C \Psi$ *such that* $w - \lim_{t\to+\infty} x(t) = x_\infty$.

The proof is based on the time rescaling $s = \sigma(t) = \int_0^t \varepsilon(u)\, du$; see the proof of Theorem 5.1.

Remark 5.1. Cominetti-Peypouquet-Sorin [23] pay special attention to the following steepest descent system with vanishing Tikhonov regularization:

$$(\mathrm{SD}_\varepsilon) \qquad \dot{x}(t) + \partial\Phi(x(t)) + \varepsilon(t)\, x(t) \ni 0.$$

If $\int_0^{+\infty} \varepsilon(t)\, dt = +\infty$, it is proved in [23] that any solution $x(\cdot)$ of (SD_ε) strongly converges as $t \to +\infty$ toward the least-norm minimizer of Φ. With essentially the same arguments, if the function Ψ is strongly convex and if $\int_0^{+\infty} \varepsilon(t)\, dt = +\infty$, then any solution of ($\mathrm{MAG}_\varepsilon$) strongly converges as $t \to +\infty$ toward the unique minimizer of Ψ over the set $C = \operatorname{argmin} \Phi$. This convergence result can be recovered from Corollary 5.3(i). Notice that in this framework, the inf-compactness property required by Corollary 5.3 appears to be superfluous.

6. Further examples of nonautonomous subgradient inclusions

The following examples illustrate the versatility of our approach and its limits.

6.1. **Quasi-autonomous case.** Let us consider the quasi-autonomous subgradient inclusion $\dot{x}(t) + \partial\Phi(x(t)) \ni f(t)$, where $\Phi : H \to \mathbb{R} \cup \{+\infty\}$ is a proper closed convex function and the map $f : \mathbb{R}_+ \to H$ tends to $f_\infty \in H$ as $t \to +\infty$. This differential inclusion falls into the setting of Theorem 2.3, thus leading to the following result.

Corollary 6.1. *Let* $f : \mathbb{R}_+ \to H$ *be a map such that* $\lim_{t\to+\infty} f(t) = f_\infty \in H$. *Let* $\Phi : H \to \mathbb{R} \cup \{+\infty\}$ *be a proper closed convex function such that* $S = \operatorname{argmin}(\Phi - \langle f_\infty, \cdot\rangle)$ *is nonempty*[14] *and bounded. Suppose that the function* Φ *satisfies the inf-compactness property* (9). *Let* $x(\cdot)$ *be a strong global solution of*

$$(25) \qquad \dot{x}(t) + \partial\Phi(x(t)) \ni f(t).$$

Then we have

(i) $\lim_{t\to+\infty} d(x(t), S) = 0$.

(ii) *Assume moreover that*

$$(26) \qquad \forall z \in S, \quad \int_0^{+\infty} G_{\partial\Phi}(z, f(t))\, dt < +\infty.$$

Then there exists $x_\infty \in S$ *such that* $x(t) \rightharpoonup x_\infty$ *weakly in* H *as* $t \to +\infty$.

Proof. One can easily verify that the hypotheses of Theorem 2.3 are fulfilled with $\varphi_t = \Phi - \langle f(t), .\rangle$ and $\varphi_\infty = \Phi - \langle f_\infty, .\rangle$. □

[14] By writing the optimality condition for the elements of S, we immediately see that $S = (\partial\Phi)^{-1}(f_\infty)$. It ensues that the nonvacuity of S is equivalent to the condition $f_\infty \in \operatorname{ran}(\partial\Phi)$.

Assumption (26) is satisfied under the following stronger condition:

$$\forall z \in S, \quad \int_0^{+\infty} [\Phi^*(f(t)) + \Phi(z) - \langle f(t), z\rangle]\, dt < +\infty. \tag{27}$$

Indeed, it suffices to observe that $G_{\partial\Phi}(z, f(t)) \leq \Phi^*(f(t)) + \Phi(z) - \langle f(t), z\rangle$. The next proposition gives sufficient conditions which guarantee that assumption (27) is satisfied.

Proposition 6.1. *Let $f : \mathbb{R}_+ \to H$ be a map such that $\lim_{t\to+\infty} f(t) = f_\infty \in H$. Let $\Phi : H \to \mathbb{R} \cup \{+\infty\}$ be a proper closed convex function such that the set $S = \operatorname{argmin}\big(\Phi - \langle f_\infty, \cdot\rangle\big)$ is nonempty and bounded. The following hold true:*

- (i) *If the function $\Phi - \langle f_\infty, \cdot\rangle$ is coercive and if $\int_0^{+\infty} \|f(t) - f_\infty\|\, dt < +\infty$, then condition* (27) *is satisfied.*
- (ii) *Assume that*

$$\Phi - \langle f_\infty, \cdot\rangle - \min_H \big(\Phi - \langle f_\infty, \cdot\rangle\big) \geq a\, d^2(., S) \tag{28}$$

for some $a > 0$ and that

$$\int_0^{+\infty} \|\Pi_F(f(t) - f_\infty)\|\, dt < +\infty \quad \textit{and} \quad \int_0^{+\infty} \|\Pi_{F^\perp}(f(t) - f_\infty)\|^2\, dt < +\infty,$$

where Π_F (resp. $\Pi_{F^\perp}$) denotes the orthogonal projection on the linear space $F = \operatorname{cl}[\mathbb{R}_+(S - S)]$ (resp. $F^\perp$). Then condition (27) *is satisfied.*

The proof of Proposition 6.1 is left to the reader. By combining Corollary 6.1 and Proposition 6.1(i), we derive that if $\int_0^{+\infty} \|f(t) - f_\infty\|\, dt < +\infty$, then any trajectory of (25) converges weakly toward some point of $S = (\partial\Phi)^{-1}(f_\infty)$. This result can be recovered directly by using the Opial lemma and the fact that the energy function $t \mapsto \Phi(x(t)) - \langle f_\infty, x(t)\rangle$ tends toward its minimum as $t \to +\infty$. The inf-compactness assumption on Φ appears to be useless; hence the result obtained as a consequence of Corollary 6.1 and Proposition 6.1(i) is not optimal. The original part of Proposition 6.1 lies in point (ii), which brings to light that the L^1-type condition on the function $f - f_\infty$ may be relaxed. If we assume the quadratic conditioning property (28), Proposition 6.1(ii) shows that it is enough to require an L^2-type condition for the part of $f - f_\infty$ that is projected on $F^\perp$.

6.2. **Sweeping process.** The sweeping process was originally considered by J. J. Moreau in the study of evolution problems from unilateral mechanics.

Given $t \mapsto C(t)$ a time-dependent closed convex set in H (the moving constraint) and $\Phi : H \to \mathbb{R}$ a convex differentiable function (the driving force). The sweeping process consists of studying the following differential inclusion:

$$\dot{x}(t) + N_{C(t)}(x(t)) + \nabla\Phi(x(t)) \ni 0, \qquad t \geq 0, \tag{SW}$$

where $N_{C(t)}(x)$ stands for the normal cone to $C(t)$ at $x \in C(t)$. Since then, its range of applications has been extended to various domains, like economical and social sciences and control theory. An abundant literature has been devoted to its study, but curiously only few results concern its asymptotical behavior.

The differential inclusion (SW) falls in the setting of Theorems 2.2 and 2.3 by taking

$$\varphi_t = \delta_{C(t)}(\cdot) + \Phi.$$

The monotonicity assumption required by Theorem 2.2 amounts to saying that the family $\{C(t);\, t \geq\, 0\}$ is nondecreasing for the set inclusion. On the other hand, it is easy to check that assumptions (H2)–(H3) of Theorem 2.3 imply that the set $C(t)$ tends toward C_∞ as $t \to +\infty$ in the Painlevé-Kuratowski sense and that $C_\infty \subset C(t)$ for t large enough. These assumptions on the family $\{C(t);\, t \geq 0\}$ are clearly quite stringent, and it is better to work directly with inclusion (SW) without resorting to the general results mentioned above.

For simplicity, we assume in the sequel that $\Phi = 0$. Most of the existence results concerning (SW) rely on energy estimates. Thus we take for granted that the trajectories have finite energy, i.e., $\int_0^{+\infty} \|\dot{x}(t)\|^2 dt < +\infty$. The result stated below is an illustration of the energetical methods.

Theorem 6.1. *Let $\{C(t);\, t \geq 0\}$ be a family of closed convex sets in H. Assume that $C(t)$ converges to some nonempty set C_∞ in the Mosco sense and that*

$$\forall z \in C_\infty,\ \exists z(t) \to z \text{ such that } z(t) \in C(t) \text{ and } \int_0^{+\infty} \|z(t) - z\|^2 dt < +\infty. \tag{29}$$

Let $x(\cdot)$ be a strong global solution of (SW) *which has a finite energy, i.e.,*

$$\int_0^{+\infty} \|\dot{x}(t)\|^2 dt < +\infty. \tag{30}$$

Then, there exists $x_\infty \in C_\infty$ such that $x(t) \rightharpoonup x_\infty$ weakly in H as $t \to +\infty$.

Proof. The proof consists of applying the Opial lemma to $x(\cdot)$ and $S = C_\infty$. It can be easily obtained by standard energetic arguments and is left to the reader. □

7. Examples of coupled gradient systems with multiscale aspects

7.1. **A two-dimensional example.** Take $H = \mathbb{R}^2$ and fix $a > 0$. Consider the function $\Psi : \mathbb{R}^2 \to \mathbb{R} \cup \{+\infty\}$ defined by

$$\Psi(x,y) = \begin{cases} \frac{y^2}{2(a^2-x^2)} & \text{if} \quad (x,y) \in]-a,a[\times\mathbb{R}, \\ 0 & \text{if} \quad (x,y) \in \{(-a,0),(a,0)\}, \\ +\infty & \text{elsewhere.} \end{cases}$$

It is easy to check that $\Psi(x,y) = \frac{1}{2a}(\sigma_D(a+x,y) + \sigma_D(a-x,y))$, where σ_D is the support function of the set D defined by

$$D = \{(x,y) \in \mathbb{R}^2, \quad 2x + y^2 \leq 0\};$$

see for example [35, Example 2.38]. The function Ψ is closed, convex and satisfies $C = \operatorname{argmin} \Psi = [-a,a] \times \{0\}$. Let us now fix $b \in\,]0,a[$ and define the function $\Phi : \mathbb{R}^2 \to \mathbb{R}$ by

$$\Phi(x,y) = y + \frac{1}{2}[x-b]_+^2 + \frac{1}{2}[x+b]_-^2,$$

for every $(x,y) \in \mathbb{R}^2$. The function Φ is convex and differentiable on $\mathbb{R}^2$. It can easily be seen that $\min_C \Phi = 0$ and that $S = \operatorname{argmin}_C \Phi = [-b,b] \times \{0\}$. Given a nondecreasing map $\beta : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\lim_{t\to+\infty} \beta(t) = +\infty$, we are interested in the asymptotic behavior as $t \to +\infty$ of the following dynamical system:

$$\dot{X}(t) + \partial\Phi(X(t)) + \beta(t)\partial\Psi(X(t)) \ni 0, \qquad \text{with } X(t) = (x(t),y(t)). \tag{31}$$

From Theorem 3.2(i), we obtain that $\lim_{t\to+\infty} d(X(t), \operatorname{argmin}_C \Phi) = 0$. We let the reader check that for every $\varepsilon > 0$, $(0, -a^2\varepsilon)$ is the unique minimum point of the function $\Psi + \varepsilon\,\Phi$ over $\mathbb{R}^2$. The corresponding minimal value equals $\omega(\varepsilon) = (\Psi + \varepsilon\,\Phi)(0, -a^2\varepsilon) = -a^2\varepsilon^2/2$. Condition ($\Sigma$7) of Theorem 3.2 amounts to

$$\int_0^{+\infty} 1/\beta(t)\, dt < +\infty.$$

Under this condition, Theorem 3.2(ii) shows that $\lim_{t\to+\infty}(x(t), y(t)) = (x_\infty, 0)$, for some $x_\infty \in [-b, b]$. For every $(x, y) \in]-a, a[\times\mathbb{R}$, we have

$$\begin{aligned} (\Psi + \varepsilon\,\Phi)(x, y) \quad &- \quad (\Psi + \varepsilon\,\Phi)(0, -a^2\varepsilon) \\ &= \frac{y^2}{2(a^2 - x^2)} + \varepsilon\, y + \frac{\varepsilon}{2}[x - b]_+^2 + \frac{\varepsilon}{2}[x + b]_-^2 + \frac{1}{2}a^2\varepsilon^2 \\ &\geq \frac{y^2}{2(a^2 - x^2)} + \varepsilon\, y + \frac{1}{2}a^2\varepsilon^2. \end{aligned} \tag{32}$$

Observe that

$$\frac{y^2}{2(a^2 - x^2)} + \varepsilon\, y + \frac{1}{2}a^2\varepsilon^2 \geq \frac{y^2}{2a^2} + \varepsilon\, y + \frac{1}{2}a^2\varepsilon^2 = \frac{1}{2a^2}(y + a^2\varepsilon)^2. \tag{33}$$

On the other hand, we have

$$\frac{y^2}{2(a^2 - x^2)} + \varepsilon\, y + \frac{1}{2}a^2\varepsilon^2 = \frac{1}{2}\varepsilon^2 x^2 + \frac{(y + \varepsilon(a^2 - x^2))^2}{2(a^2 - x^2)} \geq \frac{1}{2}\varepsilon^2 x^2. \tag{34}$$

By combining (32), (33) and (34), we find for every $(x, y) \in]-a, a[\times\mathbb{R}$,

$$(\Psi + \varepsilon\,\Phi)(x, y) - (\Psi + \varepsilon\,\Phi)(0, -a^2\varepsilon) \geq \frac{1}{4}\varepsilon^2 x^2 + \frac{1}{4a^2}(y + a^2\varepsilon)^2.$$

This inequality trivially holds true if $(x, y) \notin \operatorname{dom}\Psi$ or if $(x, y) \in \{(-a, 0), (a, 0)\}$. We infer that for every $(x, y) \in \mathbb{R}^2$ and every $\varepsilon \leq 1/a$,

$$\begin{aligned} (\Psi + \varepsilon\,\Phi)(x, y) - (\Psi + \varepsilon\,\Phi)(0, -a^2\varepsilon) \quad &\geq \quad \frac{\varepsilon^2}{4}\left(x^2 + (y + a^2\varepsilon)^2\right) \\ &= \quad \frac{\varepsilon^2}{4}\left\|(x, y) - (0, -a^2\varepsilon)\right\|^2. \end{aligned}$$

Dividing by ε and replacing ε with $1/\beta(t)$, we obtain that for every $X = (x, y) \in \mathbb{R}^2$ and every t large enough,

$$(\beta(t)\Psi + \Phi)(X) - (\beta(t)\Psi + \Phi)(\xi(t)) \geq \frac{1}{4\,\beta(t)}\,\|X - \xi(t)\|^2,$$

with $\xi(t) = (0, -a^2/\beta(t))$. This shows that assumption (i) of Theorem 2.4 is satisfied. The optimal path $t \mapsto \xi(t)$ converges toward $(0, 0)$ as $t \to +\infty$. The finite length assumption of Theorem 2.4 is fulfilled because the map $t \mapsto 1/\beta(t)$ tends non-increasingly toward 0. Assumption (ii) of Theorem 2.4 amounts to $\int_0^{+\infty} 1/\beta(t)\,dt = +\infty$. Under this last condition, Theorem 2.4 shows that $\lim_{t\to+\infty}(x(t), y(t)) = (0,0)$. To summarize, we have proved that

- if $1/\beta \in L^1(0,+\infty)$, then $\lim_{t\to+\infty}(x(t), y(t)) = (x_\infty, 0)$, for some $x_\infty \in [-b, b]$;
- if $1/\beta \notin L^1(0,+\infty)$, then $\lim_{t\to+\infty}(x(t), y(t)) = (0,0)$.

7.2. **An example in PDE theory.** Let $\Omega \subset \mathbb{R}^N$ be a bounded domain with $\mathcal{C}^1$ boundary. Let us consider the space $H = L^2(\Omega)$ endowed with the scalar product $\langle u, v\rangle_H = \int_\Omega uv$ and the corresponding norm. Let $h \in L^2(\Omega)$ be a given function satisfying $\int_\Omega h = 0$, and let a, $b \in \mathbb{R}$ be such that $a \le b$. Take

- $\Psi : L^2(\Omega) \to \mathbb{R}\cup\{+\infty\}$ defined by $\Psi(u) = \frac{1}{2}\int_\Omega \|\nabla u\|^2 - \int_\Omega hu$ if $u \in H^1(\Omega)$ and $\Psi(u) = +\infty$ otherwise.
- $\Phi : L^2(\Omega) \to \mathbb{R}$ defined by $\Phi(u) = \frac{1}{2}\int_\Omega \left\{[u(x)-b]_+^2 + [a-u(x)]_+^2\right\} dx$ for every $u \in L^2(\Omega)$.

The function Ψ is closed and convex. It is immediate to check that the variational formulation of $\xi \in \partial\Psi(u)$ is given by

$$\forall v \in H^1(\Omega), \qquad \int_\Omega \xi\, v = \int_\Omega \nabla u.\nabla v - \int_\Omega h\, v. \tag{35}$$

The function Φ is convex, differentiable and satisfies $\nabla\Phi(u) = [u-b]_+ - [a-u]_+$ for every $u \in L^2(\Omega)$. Given a map $\beta : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\lim_{t\to+\infty}\beta(t) = +\infty$, we are interested in the asymptotic behavior as $t \to +\infty$ of the dynamical system

$$\dot u(t) + \partial\Phi(u(t)) + \beta(t)\partial\Psi(u(t)) \ni 0.$$

If $u(\cdot)$ is a solution of the above differential inclusion, then for almost every $t \ge 0$, there exists $\xi(t) \in \partial\Psi(u(t))$ such that

$$\dot u(t) + [u(t)-b]_+ - [a-u(t)]_+ + \beta(t)\xi(t) = 0.$$

Taking the scalar product with $v \in H^1(\Omega)$, we obtain in view of (35)

$$\int_\Omega \dot u(t)\, v + \int_\Omega ([u(t)-b]_+ - [a-u(t)]_+)\, v + \beta(t)\left[\int_\Omega \nabla u(t).\nabla v - \int_\Omega h\, v\right] = 0.$$

By using Green's formula, we find that for every $v \in H^1(\Omega)$,

$$\begin{aligned}\int_\Omega \dot u(t)\, v \quad &+ \quad \int_\Omega ([u(t)-b]_+ - [a-u(t)]_+)\, v \\ &+ \quad \beta(t)\left[-\int_\Omega \Delta u(t)\, v + \int_{\partial\Omega} \frac{\partial u(t)}{\partial n}\, v - \int_\Omega h\, v\right] = 0.\end{aligned}$$

This yields

$$\begin{cases} \dot u(t) + [u(t)-b]_+ - [a-u(t)]_+ + \beta(t)\left[-\Delta u(t) - h\right] & = 0 \quad \text{on } \Omega, \\ \frac{\partial u(t)}{\partial n} & = 0 \quad \text{on } \partial\Omega. \end{cases}$$

The elements of $C = \operatorname{argmin} \Psi$ are solutions of the minimization problem

$$\inf \left\{ \frac{1}{2} \int_\Omega \|\nabla u\|^2 - \int_\Omega h\, u : \quad u \in H^1(\Omega) \right\}.$$

The corresponding weak variational formulation is given by

$$\forall v \in H^1(\Omega), \quad \int_\Omega \nabla u . \nabla v = \int_\Omega h\, v. \tag{36}$$

Since $\int_\Omega h = 0$, it is well-known that such solutions exist, and they satisfy the following Neumann boundary value problem:

$$\begin{cases} -\Delta u - h & = 0 \quad \text{on } \Omega, \\ \quad\frac{\partial u}{\partial n} & = 0 \quad \text{on } \partial\Omega. \end{cases}$$

Denoting by $\widehat{u}$ a particular solution, the set $C = \operatorname{argmin} \Psi$ is the straight line $C = \{\widehat{u} + m, m \in \mathbb{R}\}$. Let us now check that the function Ψ satisfies the inf-compactness property (9). Given $R > 0$ and $l \in \mathbb{R}$, let $u \in L^2(\Omega)$ be in the lower level set

$$\Lambda_{R,l} = \{u \in L^2(\Omega), \quad \|u\|_{L^2} \leq R, \ \Psi(u) \leq l\}.$$

From the definition of Ψ, we have $u \in H^1(\Omega)$ and

$$\begin{aligned} \int_\Omega \|\nabla u\|^2 &\leq 2l + 2 \int_\Omega h\, u \\ &\leq 2l + 2\, \|h\|_{L^2} \|u\|_{L^2} \leq 2\, l + 2\, R\, \|h\|_{L^2}. \end{aligned}$$

We immediately deduce that

$$\|u\|^2_{H^1} = \int_\Omega u^2 + \int_\Omega \|\nabla u\|^2 \leq R^2 + 2\, l + 2\, R\, \|h\|_{L^2},$$

which shows that the set $\Lambda_{R,l}$ is bounded for the $H^1(\Omega)$-norm. Since Ω is bounded with $\mathcal{C}^1$ boundary, by the Rellich-Kondrachov theorem, the injection $H^1(\Omega) \hookrightarrow L^2(\Omega)$ is compact. We conclude that $\Lambda_{R,l}$ is relatively compact for the $L^2(\Omega)$-norm; hence the function Ψ satisfies the inf-compactness property (9).

Let us now determine the set $S = \operatorname{argmin}_C \Phi$. Since the function Φ is continuous, convex and coercive, the set $\operatorname{argmin}_C \Phi$ is a nonempty segment included in C. Recall that $u \in \operatorname{argmin}_C \Phi$ if and only if it satisfies the optimality condition $-\nabla\Phi(u) \in N_C(u)$. Since the set C is a straight line directed by the space of constant functions, it is clear that

$$N_C(u) = \big\{p \in L^2(\Omega), \quad \langle p, 1\rangle_{L^2(\Omega)} = 0\big\} = \left\{p \in L^2(\Omega), \quad \int_\Omega p = 0\right\}.$$

Finally, we obtain the equivalences

$$\begin{aligned} u \in \operatorname{argmin}_C \Phi &\iff \int_\Omega \nabla\Phi(u)(x)\, dx = 0 \\ &\iff \int_\Omega \Big([u(x) - b]_+ - [a - u(x)]_+\Big)\, dx = 0. \end{aligned} \tag{37}$$

Assuming that $\widehat{u} \in \operatorname{argmin}_C \Phi$, let us denote by $\inf_\Omega \widehat{u}$ (resp. $\sup_\Omega \widehat{u}$) the essential infimum (resp. supremum) of $\widehat{u}$ over the set Ω. We distinguish the cases $\sup_\Omega \widehat{u} - \inf_\Omega \widehat{u} > b-a$ and $\sup_\Omega \widehat{u} - \inf_\Omega \widehat{u} \leq b-a$.

Case 1. $\sup_\Omega \widehat{u} - \inf_\Omega \widehat{u} > b-a$. In view of condition (37), we deduce that the sets

$$\Omega_+ = \{x \in \Omega, \quad \widehat{u}(x) > b\} \qquad \text{and} \qquad \Omega_- = \{x \in \Omega, \quad \widehat{u}(x) < a\}$$

have positive measures. For $m \in \mathbb{R}$, let us define the quantity $\theta(m)$ by

$$\theta(m) = \int_\Omega \Big([\widehat{u}(x) + m - b]_+ - [a - m - \widehat{u}(x)]_+ \Big)\, dx.$$

Recalling that $\theta(0) = 0$, we have for every $m \geq 0$,

$$\begin{aligned}
\theta(m) \quad &= \quad \int_\Omega \Big([\widehat{u}(x) + m - b]_+ - [\widehat{u}(x) - b]_+ \Big)\, dx \\
&+ \quad \int_\Omega \Big([a - \widehat{u}(x)]_+ - [a - m - \widehat{u}(x)]_+ \Big)\, dx \\
&\geq \quad \int_\Omega \Big([\widehat{u}(x) + m - b]_+ - [\widehat{u}(x) - b]_+ \Big)\, dx \\
&\geq \quad \int_{\Omega_+} \Big([\widehat{u}(x) + m - b]_+ - [\widehat{u}(x) - b]_+ \Big)\, dx \\
&= \quad \int_{\Omega_+} m\, dx = m\, |\Omega_+|.
\end{aligned}$$

In the same way, we obtain $\theta(m) \leq m\, |\Omega_-|$ for every $m \leq 0$. Since $|\Omega_+|$ and $|\Omega_-|$ are positive, this implies that $\theta(m) = 0$ if and only if $m = 0$. In view of (37), we conclude that $\widehat{u}$ is the unique minimum of Φ over the set $C = \{\widehat{u} + m,\ m \in \mathbb{R}\}$. We then infer from Theorem 3.2(i) that $\lim_{t\to+\infty} u(t) = \widehat{u}$ strongly in $L^2(\Omega)$.

Case 2. $\sup_\Omega \widehat{u} - \inf_\Omega \widehat{u} \leq b-a$. In view of condition (37), we deduce that $\widehat{u}(x) \in [a,b]$ for almost every $x \in \Omega$. We then have $\Phi(\widehat{u}) = 0$, hence $\widehat{u} \in \operatorname{argmin} \Phi$. It ensues that

$$\begin{aligned}
S \quad &= \quad \operatorname{argmin} \Psi \cap \operatorname{argmin} \Phi \\
&= \quad \left\{ \widehat{u} + m, \quad m \in \left[a - \inf_\Omega \widehat{u}, b - \sup_\Omega \widehat{u} \right] \right\}.
\end{aligned}$$

By combining Corollary 3.2 and Remark 3.3, we deduce that there exists $\overline{u} \in S$ such that $\lim_{t\to+\infty} u(t) = \overline{u}$ strongly in $L^2(\Omega)$.

In fact the convergence is strong in $H^1(\Omega)$ in each of the above cases. Indeed, observe that

$$\begin{aligned}
\|u(t) - \overline{u}\|_{H^1}^2 \quad &= \quad \int_\Omega \|\nabla u(t) - \nabla \overline{u}\|^2 + \int_\Omega |u(t) - \overline{u}|^2 \\
&= \quad \int_\Omega \|\nabla u(t)\|^2 - 2 \int_\Omega \nabla u(t) \nabla \overline{u} + \int_\Omega \|\nabla \overline{u}\|^2 + \int_\Omega |u(t) - \overline{u}|^2.
\end{aligned}$$

By using the weak variational formulation (36), we obtain that $\int_\Omega \nabla u(t)\nabla \overline{u} = \int_\Omega h\, u(t)$ and that $\int_\Omega \|\nabla \overline{u}\|^2 = \int_\Omega h\, \overline{u}$. We immediately deduce from the above equality that

$$\|u(t)-\overline{u}\|_{H^1}^2 = 2\left(\Psi(u(t)) - \Psi(\overline{u})\right) + \int_\Omega |u(t)-\overline{u}|^2.$$

Since $\lim_{t\to+\infty} \Psi(u(t)) = \min_H \Psi$ (see Remark 3.3) and $\lim_{t\to+\infty} \|u(t)-\overline{u}\|_{L^2} = 0$, we conclude that $\lim_{t\to+\infty} \|u(t)-\overline{u}\|_{H^1} = 0$.

Acknowledgment

The authors express their gratitude to the anonymous referees for their careful reading of the paper. Their valuable suggestions and critical comments made numerous improvements throughout.

References

[1] F. Alvarez and A. Cabot, *Asymptotic selection of viscosity equilibria of semilinear evolution equations by the introduction of a slowly vanishing term*, Discrete Contin. Dyn. Syst. **15** (2006), no. 3, 921–938, DOI 10.3934/dcds.2006.15.921. MR2220756

[2] H. Attouch, *Variational convergence for functions and operators*, Applicable Mathematics Series, Pitman (Advanced Publishing Program), Boston, MA, 1984. MR773850

[3] H. Attouch, *Viscosity solutions of minimization problems*, SIAM J. Optim. **6** (1996), no. 3, 769–806, DOI 10.1137/S1052623493259616. MR1402205

[4] H. Attouch, G. Buttazzo, and G. Michaille, *Variational analysis in Sobolev and BV spaces: Applications to PDEs and optimization*, 2nd ed., MOS-SIAM Series on Optimization, vol. 17, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA; Mathematical Optimization Society, Philadelphia, PA, 2014. MR3288271

[5] H. Attouch and R. Cominetti, *A dynamical approach to convex minimization coupling approximation with the steepest descent method*, J. Differential Equations **128** (1996), no. 2, 519–540, DOI 10.1006/jdeq.1996.0104. MR1398330

[6] H. Attouch and M.-O. Czarnecki, *Asymptotic behavior of coupled dynamical systems with multiscale aspects*, J. Differential Equations **248** (2010), no. 6, 1315–1344, DOI 10.1016/j.jde.2009.06.014. MR2593044

[7] H. Attouch, M.-O. Czarnecki, and J. Peypouquet, *Prox-penalization and splitting methods for constrained variational problems*, SIAM J. Optim. **21** (2011), no. 1, 149–173, DOI 10.1137/100789464. MR2765493

[8] H. Attouch, M.-O. Czarnecki, and J. Peypouquet, *Coupling forward-backward with penalty schemes and parallel splitting for constrained variational inequalities*, SIAM J. Optim. **21** (2011), no. 4, 1251–1274, DOI 10.1137/110820300. MR2854582

[9] H. Attouch and A. Damlamian, *Strong solutions for parabolic variational inequalities*, Nonlinear Anal. **2** (1978), no. 3, 329–353, DOI 10.1016/0362-546X(78)90021-4. MR512663

[10] D. Azé, *Eléments d'analyse convexe et variationnelle*, Ellipses, Paris, 1997.

[11] H. H. Bauschke and P. L. Combettes, *Convex analysis and monotone operator theory in Hilbert spaces*, CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, Springer, New York, 2011. MR2798533

[12] J. B. Baillon and R. Cominetti, *A convergence result for nonautonomous subgradient evolution equations and its application to the steepest descent exponential penalty trajectory in linear programming*, J. Funct. Anal. **187** (2001), no. 2, 263–273, DOI 10.1006/jfan.2001.3828. MR1875148

[13] J. B. Baillon and H. Brezis, *Une remarque sur le comportement asymptotique des semigroupes non linéaires* (French), Houston J. Math. **2** (1976), no. 1, 5–7. MR0394328

[14] H. H. Bauschke, D. A. McLaren, and H. S. Sendov, *Fitzpatrick functions: inequalities, examples, and remarks on a problem by S. Fitzpatrick*, J. Convex Anal. **13** (2006), no. 3-4, 499–523. MR2291550

[15] R. I. Boţ and E. R. Csetnek, *Forward-backward and Tseng's type penalty schemes for monotone inclusion problems*, Set-Valued Var. Anal. **22** (2014), no. 2, 313–331, DOI 10.1007/s11228-014-0274-7. MR3207742

[16] R. I. Boţ and E. R. Csetnek, *Approaching the solving of constrained variational inequalities via penalty term-based dynamical systems*, J. Math. Anal. Appl. **435** (2016), no. 2, 1688–1700, DOI 10.1016/j.jmaa.2015.11.032. MR3429667

[17] H. Brézis, *Opérateurs maximaux monotones dans les espaces de Hilbert et équations d'évolution*, Lecture Notes, vol. 5, North-Holland, 1972.

[18] H. Brézis, *Asymptotic behavior of some evolution systems*, Nonlinear evolution equations (Proc. Sympos., Univ. Wisconsin, Madison, Wis., 1977), Publ. Math. Res. Center Univ. Wisconsin, vol. 40, Academic Press, New York-London, 1978, pp. 141–154. MR513816

[19] H. Brezis and A. Haraux, *Image d'une somme d'opérateurs monotones et applications*, Israel J. Math. **23** (1976), no. 2, 165–186, DOI 10.1007/BF02756796. MR0399965

[20] R. E. Bruck Jr., *Asymptotic convergence of nonlinear contraction semigroups in Hilbert space*, J. Funct. Anal. **18** (1975), 15–26, DOI 10.1016/0022-1236(75)90027-0. MR0377609

[21] R. S. Burachik and B. F. Svaiter, *Maximal monotone operators, convex functions and a special family of enlargements*, Set-Valued Anal. **10** (2002), no. 4, 297–316, DOI 10.1023/A:1020639314056. MR1934748

[22] A. Cabot, *Proximal point algorithm controlled by a slowly vanishing term: applications to hierarchical minimization*, SIAM J. Optim. **15** (2004/05), no. 2, 555–572, DOI 10.1137/S105262340343467X. MR2144181

[23] R. Cominetti, J. Peypouquet, and S. Sorin, *Strong asymptotic convergence of evolution equations governed by maximal monotone operators with Tikhonov regularization*, J. Differential Equations **245** (2008), no. 12, 3753–3763, DOI 10.1016/j.jde.2008.08.007. MR2462703

[24] I. Ekeland and R. Témam, *Convex analysis and variational problems*, corrected reprint of the 1976 English edition, Classics in Applied Mathematics, vol. 28, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1999. MR1727362

[25] S. Fitzpatrick, *Representing monotone operators by convex functions*, Workshop/Miniconference on Functional Analysis and Optimization (Canberra, 1988), Proc. Centre Math. Anal. Austral. Nat. Univ., vol. 20, Austral. Nat. Univ., Canberra, 1988, pp. 59–65. MR1009594

[26] H. Furuya, K. Miyashiba, and N. Kenmochi, *Asymptotic behavior of solutions to a class of nonlinear evolution equations*, J. Differential Equations **62** (1986), no. 1, 73–94, DOI 10.1016/0022-0396(86)90106-3. MR830048

[27] N. Kenmochi, *Solvability of nonlinear equations with time-dependent constraints and applications*, Bull. Fac. Educ. Chiba Univ. **30** (1981), 1–87.

[28] B. Lemaire, *On the convergence of some iterative methods for convex minimization*, Recent developments in optimization (Dijon, 1994), Lecture Notes in Econom. and Math. Systems, vol. 429, Springer, Berlin, 1995, pp. 252–268, DOI 10.1007/978-3-642-46823-0_20. MR1358403

[29] J.-E. Martínez-Legaz and B. F. Svaiter, *Monotone operators representable by l.s.c. convex functions*, Set-Valued Anal. **13** (2005), no. 1, 21–46, DOI 10.1007/s11228-004-4170-4. MR2128696

[30] J.-E. Martinez-Legaz and M. Théra, *A convex representation of maximal monotone operators*, J. Nonlinear Convex Anal. **2** (2001), no. 2, 243–247. MR1848704

[31] Z. Opial, *Weak convergence of the sequence of successive approximations for nonexpansive mappings*, Bull. Amer. Math. Soc. **73** (1967), 591–597, DOI 10.1090/S0002-9904-1967-11761-0. MR0211301

[32] G. B. Passty, *Ergodic convergence to a zero of the sum of monotone operators in Hilbert space*, J. Math. Anal. Appl. **72** (1979), no. 2, 383–390, DOI 10.1016/0022-247X(79)90234-8. MR559375

[33] J.-P. Penot and C. Zălinescu, *On the convergence of maximal monotone operators*, Proc. Amer. Math. Soc. **134** (2006), no. 7, 1937–1946, DOI 10.1090/S0002-9939-05-08275-4. MR2215762

[34] R. T. Rockafellar, *Convex analysis*, Princeton Mathematical Series, No. 28, Princeton University Press, Princeton, N.J., 1970. MR0274683

[35] R. T. Rockafellar and R. J.-B. Wets, *Variational analysis*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 317, Springer-Verlag, Berlin, 1998. MR1491362

[36] B. Djafari Rouhani, *Asymptotic behaviour of quasi-autonomous dissipative systems in Hilbert spaces*, J. Math. Anal. Appl. **147** (1990), no. 2, 465–476, DOI 10.1016/0022-247X(90)90361-I. MR1050218

[37] S. Simons and C. Zălinescu, *A new proof for Rockafellar's characterization of maximal monotone operators*, Proc. Amer. Math. Soc. **132** (2004), no. 10, 2969–2972, DOI 10.1090/S0002-9939-04-07462-3. MR2063117

[38] S. Simons and C. Zălinescu, *Fenchel duality, Fitzpatrick functions and maximal monotonicity*, J. Nonlinear Convex Anal. **6** (2005), no. 1, 1–22. MR2138099

[39] D. Torralba, *Développements asymptotiques pour les méthodes d'approximation par viscosité* (French, with English and French summaries), C. R. Acad. Sci. Paris Sér. I Math. **322** (1996), no. 2, 123–128. MR1373747

Institut Montpelliérain Alexander Grothendieck, UMR 5149 CNRS, Université Montpellier 2, place Eugène Bataillon, 34095 Montpellier cedex 5, France
E-mail address: hedy.attouch@univ-montp2.fr

Institut de Mathématiques de Bourgogne, UMR 5584, CNRS, Université Bourgogne Franche-Comté, 21000 Dijon, France
E-mail address: alexandre.cabot@u-bourgogne.fr

Institut Montpelliérain Alexander Grothendieck, UMR 5149 CNRS, Université Montpellier 2, place Eugène Bataillon, 34095 Montpellier cedex 5, France
E-mail address: marco@math.univ-montp2.fr