Comportement asymptotique de systèmes dynamiques discrets et continus en Optimisation et EDP: algorithmes de minimisation proximale alternée et dynamique du deuxième ordre à dissipation évanescente

Pierre Frankel

UNIVERSITÉ MONTPELLIER II

- SCIENCES ET TECHNIQUES DU LANGUEDOC -

THÈSE

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITE MONTPELLIER II

Ecole Doctorale: Information, Structures et Systèmes
Formation Doctorale: Mathématiques
Spécialité : Mathématiques Appliquées

PIERRE FRANKEL

Comportement asymptotique de systèmes dynamiques discrets et continus en Optimisation et EDP: algorithmes de minimisation proximale alternée et dynamique du deuxième ordre à dissipation évanescente.

Soutenance le 27 septembre 2011 devant le jury composé de:

F. Alvarez	Universidad de Chile	Rapporteur
H. Attouch	Université Montpellier II	Co-directeur
A. Cabot	Université Montpellier II	Co-directeur
P.-L. Combettes	Université Pierre et Marie Curie - Paris 6	Rapporteur
A. Haraux	Université Pierre et Marie Curie - Paris 6	Examinateur
G. Michaille	Université Montpellier II	Examinateur
A. Seeger	Université d'Avignon	Examinateur
B. Svaiter	Instituto Nacional de Matemática Pura e Aplicada (IMPA)	Examinateur

Remerciements

Ce travail est l'aboutissement de trois années d'études à l'Université Montpellier II menées en parallèle avec un enseignement au lycée Dhuoda de Nîmes ainsi que des vacations à l'Université des Carmes de Nîmes. Ce parcours n'aurait pas été possible sans la contribution indispensable de mes parents. Je remercie également mes directeurs de thèse, Alexandre Cabot et Hedy Attouch, pour la trés grande qualité de leur encadrement. Je remercie aussi Gérard Michaille, Professeur à l'Université de Montpellier II, pour m'avoir permis de donner des vacations et pour nos nombreuses discussions Mathématiques. Je remercie Felipe Alvarez, Professeur à l'Université du Chili à Santiago, et Juan Peypouquet, Professeur à l'Université Federico Santa Maria de Valparaiso, pour leur accueil lors de mon séjour de recherche au Chili et pour leur collaboration. Enfin, je remercie Frédéric Terraza, Professeur au lycée Dhuoda, pour ses explications concernant l'utilisation du Logiciel LateX.

Contents

INTRODUCTION

La première partie de cette thèse (articles I et II) est consacrée à l'étude du comportement asymptotique des solutions de dynamiques du second ordre avec dissipation évanescente.

Dans l'article I, nous nous intéressons à une équation hyperbolique semi-linéaire amortie. Soit V et H deux espaces de Hilbert réels. Soit $a : V \times V \to \mathbb{R}$ une forme bilinéaire, continue, symétrique, positive et semi-coercive (c'est-à-dire $\exists \lambda \geq 0, \mu > 0$ tel que $\forall u \in V, \quad a(u,u) + \lambda |u|_H^2 \geq \mu \|u\|_V^2$). Nous associons à $a(.,.)$ l'opérateur linéaire continu $A : V \to V'$ défini par $\langle Au, v \rangle_{V',V} = a(u,v)$ pour tout $u, v \in V$. Etant donnée une fonction $f : V \to H$, nous considérons l'équation d'évolution semi-linéaire du second ordre

$$(E) \qquad \frac{d^2u}{dt^2}(t) + \gamma(t)\frac{du}{dt}(t) + Au(t) + f(u(t)) = 0, \quad t \geq 0,$$

où $\gamma : \mathbb{R}_+ \to \mathbb{R}_+$ avec $\lim_{t \to +\infty} \gamma(t) = 0$. Cette équation modélise par exemple des phénomènes de propagation d'ondes ou de vibrations soumis à une force extérieure $-f(u)$ et à une force de frottement ou d'amortissement évanescente $-\gamma \frac{du}{dt}$. Dans un cadre fonctionnel différent, Cabot, Engler et Gadat [10, 11] ont étudié le comportement asymptotique des solutions de l'équation différentielle du second ordre plus générale suivante

$$(\mathcal{S}_1) \qquad \ddot{x}(t) + \gamma(t)\,\dot{x}(t) + \nabla\Phi(x(t)) = 0, \quad t \geq 0,$$

où H est un espace de Hilbert et $\Phi : H \to \mathbb{R}$ est une fonction de classe C^1 et convexe. L'analyse repose sur l'utilisation de la fonction énergie définie par

$$\mathcal{E}(t) = \frac{1}{2}|\dot{x}(t)|^2 + \Phi(x(t)),$$

qui est l'énergie mécanique du point matériel. Lorsque $\gamma(t) \equiv \gamma > 0$, l'équation $(\mathcal{S}_1)$ est dénommée problème de la boule pesante avec frottement et a été etudiée par Alvarez [2]. Les premiers résultats obtenus par Cabot, Engler et Gadat [10, 11] sur la sommabilité et la convergence de la fonction énergie $\mathcal{E}$ du système $(\mathcal{S}_1)$ sont les suivants:

Proposition 0.1 *Supposons $\gamma : \mathbb{R}_+ \to \mathbb{R}_+$ dérivable et décroissante, $\Phi : H \to \mathbb{R}$ C^1 et convexe, alors toute solution x de $(\mathcal{S}_1)$ vérifie*

(i) $\dot{\mathcal{E}}(t) = -\gamma(t)|\dot{x}(t)|^2$.

Si de plus Argmin $\Phi \neq \varnothing$, alors toute solution x bornée dans H vérifie

(ii) $\int_0^{+\infty} \gamma(t)(\mathcal{E}(t) - \min \Phi)dt < \infty$,

(iii) si de plus $\gamma \notin L^1(0, +\infty)$, alors $\lim_{t\to+\infty} \mathcal{E}(t) = \min \Phi$ *et donc* $\lim_{t\to+\infty} |\dot{x}(t)| = 0$ *et* $\lim_{t\to+\infty} \Phi(t) = \min \Phi$.

Nous obtenons des résultats similaires dans le cas de l'équation (E) en utilisant une fonction énergie appropriée. Dans le cas d'un amortissement constant $\gamma(t) \equiv \gamma > 0$, la convergence des solutions de (E) a été obtenue par Alvarez et Attouch [3]. La fonction $f : V \to H$ est supposée conservatrice

$$\exists F \in C^1(V, \mathbb{R}) / \quad \forall u, v \in V, \quad \langle F'(u), v\rangle_{V',V} = (f(u), v)_H,$$

et monotone

$$\forall u, v \in V, \quad \langle f(u) - f(v), u - v\rangle_V \geq 0.$$

Les auteurs ont obtenu le théorème suivant:

Théorème 0.1 *Supposons $\gamma(t) \equiv \gamma > 0$. Soit $a : V \times V \to \mathbb{R}$ une forme bilinéaire continue, symétrique, positive et semi-coercive et soit $f : V \to H$ conservatrice et monotone. Supposons que $S = \{v \in V; \quad Av + f(v) = 0\} \neq \varnothing$. Alors toute solution u de (E) converge faiblement dans V quand $t \to +\infty$ vers un point de S.*

Nous généralisons ces résultats de convergence à l'équation (E) dans le cas d'un amortissement évanescent. Sous les mêmes hypothèses sur la forme bilinéaire $a(.,.)$ et sur la fonction f, nous obtenons la convergence faible dans V vers un point de S de toute solution bornée dans H si l'application γ tend *lentement* vers 0 quand $t \to +\infty$ (par exemple, s'il existe $\alpha \in]0; 1[$ tel que $\gamma(t) \sim \frac{1}{t^\alpha}$ quand $t \to +\infty$).

Dans l'article II, nous nous intéressons à l'algorithme proximal inertiel suivant

$$(\mathcal{A}) \qquad x_{n+1} - x_n - \alpha_n(x_n - x_{n-1}) + \beta_n \partial\Phi(x_{n+1}) \ni 0,$$

où H est un espace de Hilbert, $\Phi : H \to \mathbb{R} \cup \{+\infty\}$ est une fonction convexe propre s.c.i., (α_n) et (β_n) sont des suites strictement positives. Nous pouvons réécrire l'algorithme $(\mathcal{A})$ de la façon suivante:

$$\frac{x_{n+1} - 2x_n + x_{n-1}}{\beta_n} + \frac{1 - \alpha_n}{\beta_n}(x_n - x_{n-1}) + \partial\Phi(x_{n+1}) \ni 0.$$

L'algorithme $(\mathcal{A})$ apparaît donc comme une discrétisation implicite du système continu $(\mathcal{S}_1)$ avec un pas de temps égal à $\sqrt{\beta_n}$, tandis que $\frac{1-\alpha_n}{\sqrt{\beta_n}}$ correspond à la valeur de γ au temps $t_n = \sum_{k=0}^n \sqrt{\beta_k}$. L'algorithme $(\mathcal{A})$ a été étudié par Alvarez [2] qui a obtenu le résultat suivant:

Théorème 0.2 *Supposons que $\Phi : H \to \mathbb{R} \cup \{+\infty\}$ est s.c.i., convexe et propre. Soit (α_n) et (β_n) des suites de réels strictement positifs telles que $\alpha_n \le \overline{\alpha} < 1$, (β_n) majorée et $(\frac{\alpha_n}{\beta_n})$ décroissante. Si Argmin $\Phi \neq \varnothing$, alors toute suite (x_n) générée par $(\mathcal{A})$ converge faiblement dans H vers un minimiseur de Φ.*

L'étude repose sur l'utilisation d'une fonction énergie et du lemme d'Opial où l'hypothèse $\alpha_n \le \overline{\alpha} < 1$ joue un rôle crucial. Nous étudions la convergence de $(\mathcal{A})$ sous l'hypothése plus générale $0 < \alpha_n \le 1$ et examinons le cas $\lim_{n\to+\infty} \alpha_n = 1$. Dans un premier temps, nous étudions la convergence de la fonction énergie

$$\mathcal{E}_n = \frac{1}{2\beta_{n-1}}|x_n - x_{n-1}|^2 + \Phi(x_n),$$

où (x_n) est générée par $(\mathcal{A})$. Nous obtenons des résultats de sommabilité et de convergence de la suite $(\mathcal{E}_n)$ similaires aux résultats obtenus à la Proposition 0.1 dans le cas du système continu $(\mathcal{S}_1)$. Dans l'article [10], Cabot, Engler et Gadat ont prouvé que, si $\int_0^{+\infty} e^{-\int_0^t \gamma(s)ds} dt = \infty$, toute solution x de $(\mathcal{S}_1)$ telle que $(x(0), \dot{x}(0)) \notin \text{Argmin}\, \Phi \times \{0\}$ ne converge pas. Nous trouvons des résultats analogues dans le cas discret pour l'algorithme $(\mathcal{A})$.

La deuxième partie de cette thèse (articles III à VI) est consacrée à l'étude de plusieurs algorithmes de type proximal. Nous montrons que ces algorithmes convergent vers des solutions de certains problèmes de minimisation. Dans chaque cas, une application est donnée dans le cadre de la décomposition de domaine pour les EDP.

$\mathcal{X}$, $\mathcal{Y}$, $\mathcal{Z}$ sont des espaces de Hilbert, $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ et $g : \mathcal{Y} \to \mathbb{R} \cup \{+\infty\}$ sont des fonctions s.c.i., convexes et propres, $A : \mathcal{X} \to \mathcal{Z}$ et $B : \mathcal{Y} \to \mathcal{Z}$ sont des opérateurs linéaires continus. Nous considérons la fonction convexe $\Phi_\gamma : \mathcal{X} \times \mathcal{Y} \to \mathbb{R} \cup \{+\infty\}$ définie par

$$\Phi_\gamma(x, y) = f(x) + g(y) + \frac{1}{2\gamma}\|Ax - By\|^2_{\mathcal{Z}},$$

où γ est un paramètre strictement positif. Dans le but de minimiser la fonction Φ_γ, Attouch, Bolte, Redont et Soubeyran [5] ont introduit l'algorithme alterné avec termes de coûts-aux-changements

$$(\mathcal{A}_1) \quad \begin{cases} x_{n+1} = \text{Argmin}\left\{ f(x) + \frac{1}{2\gamma}\|Ax - By_n\|^2_{\mathcal{Z}} + \frac{\alpha}{2}\|x - x_n\|^2_{\mathcal{X}}; \quad x \in \mathcal{X} \right\} \\ y_{n+1} = \text{Argmin}\left\{ g(y) + \frac{1}{2\gamma}\|Ax_{n+1} - By\|^2_{\mathcal{Z}} + \frac{\nu}{2}\|y - y_n\|^2_{\mathcal{Y}}; \quad y \in \mathcal{Y} \right\}, \end{cases}$$

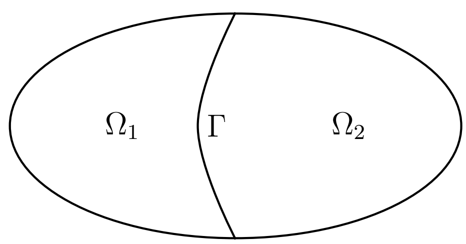
où α, ν sont des paramètres strictement positifs. Les auteurs ont montré que toute suite générée par l'algorithme $(\mathcal{A}_1)$ converge faiblement vers un point solution du problème

$$(\mathcal{P}_1) \qquad \min\{\Phi_\gamma(x,y); \quad (x,y) \in \mathcal{X}\times\mathcal{Y}\}$$
$$= \min\left\{f(x)+g(y)+\frac{1}{2\gamma}\|Ax-By\|^2_{\mathcal{Z}}; \quad (x,y)\in\mathcal{X}\times\mathcal{Y}\right\}.$$

L'algorithme $(\mathcal{A}_1)$ utilise la structure de la fonction objectif Φ_γ pour résoudre le problème initial sur $\mathcal{X}\times\mathcal{Y}$ en résolvant respectivement des problèmes sur $\mathcal{X}$ et $\mathcal{Y}$. Dans un article antérieur, Acker et Prestel [1] avaient étudié le problème fortement couplé ($\mathcal{X}=\mathcal{Y}$, $A=B=\mathcal{I}$ et $\alpha=\nu=0$ dans l'algorithme). Dans l'article III, nous généralisons les méthodes et les résultats de convergence de [1] au problème faiblement couplé $(\mathcal{P}_1)$. Nous retrouvons la convergence faible dans $\mathcal{X}\times\mathcal{Y}$ de la suite (x_n,y_n) générée par l'algorithme $(\mathcal{A}_1)$ vers un point solution de $(\mathcal{P}_1)$ et montrons la convergence forte dans $\mathcal{Z}$ de la suite de variables duales $(-\frac{1}{\gamma}(Ax_n-By_n))$ vers l'unique solution du problème dual[1]

$$\inf\left\{f^*(A^*z^*)+g^*(-B^*z^*)+\frac{\gamma}{2}\|z^*\|^2_{\mathcal{Z}}; \quad z^*\in\mathcal{Z}\right\}.$$

Le cadre d'application à la décomposition de domaine pour les EDP est le suivant: nous considérons un domaine borné $\Omega=\Omega_1\cup\Omega_2\cup\Gamma$ de $\mathbb{R}^N$ suffisamment régulier qui peut se décomposer en deux sous-domaines Ω_1 et Ω_2 avec une interface commune Γ.



Nous choisissons $\mathcal{X}=H^1(\Omega_1)$, $\mathcal{Y}=H^1(\Omega_2)$ et $\mathcal{Z}=L^2(\Gamma)$. Les opérateurs $A:\mathcal{X}\to\mathcal{Z}$ et $B:\mathcal{Y}\to\mathcal{Z}$ sont les opérateurs traces sur Γ. Le terme $[w]=Au-Bv$ correspond au saut de l'application $w=\begin{cases}u \text{ sur } \Omega_1\\ v \text{ sur } \Omega_2\end{cases}$ à travers l'interface Γ. Les fonctions f et g sont définies par

$$f(u)=\frac{1}{2}\int_{\Omega_1}|\nabla u|^2-\int_{\Omega_1}hu \quad \text{et} \quad g(v)=\frac{1}{2}\int_{\Omega_2}|\nabla v|^2-\int_{\Omega_2}hv.$$

La fonction $h\in L^2(\Omega)$ est fixée. Dans ce cas l'algorithme $(\mathcal{A}_1)$ permet de résoudre par décomposition le problème de minimisation suivant

[1] $f^*:\mathcal{Z}\to\mathbb{R}\cup\{+\infty\}$ et $g^*:\mathcal{Z}\to\mathbb{R}\cup\{+\infty\}$ sont les conjuguées de Fenchel des fonctions f et g, $A^*:\mathcal{Z}\to\mathcal{X}$ et $B^*:\mathcal{Z}\to\mathcal{Y}$ sont les opérateurs adjoints de A et B.

$$\min\left\{\frac{1}{2}\int_{\Omega_1}|\nabla u|^2+\frac{1}{2}\int_{\Omega_2}|\nabla v|^2+\frac{1}{2\gamma}\int_{\Gamma}[w]^2-\int_{\Omega}hw;\quad u\in H^1(\Omega_1),v\in H^1(\Omega_2)\right\},$$

qui est la formulation variationnelle du problème avec conditions au bord mixtes de Dirichlet-Neumann suivant sur Ω_1

$$\begin{cases}-\Delta u=h & \text{dans}\ \ \Omega_1\\ \frac{\partial u}{\partial n}=0 & \text{sur}\ \ \partial\Omega_1\cap\partial\Omega\\ \frac{\partial u}{\partial n}=-\frac{1}{\gamma}(u-v) & \text{sur}\ \ \Gamma,\end{cases}$$

et Ω_2

$$\begin{cases}-\Delta v=h & \text{dans}\ \ \Omega_2\\ \frac{\partial v}{\partial n}=0 & \text{sur}\ \ \partial\Omega_2\cap\partial\Omega\\ \frac{\partial v}{\partial n}=\frac{1}{\gamma}(u-v) & \text{sur}\ \ \Gamma.\end{cases}$$

Ce type de problème peut apparaître dans la description de phénomèmes autorisant des discontinuités à travers l'interface Γ.

Le problème de minimisation avec contraintes

$$(\mathcal{P}_2)\qquad\qquad \min\{f(x)+g(y);\quad Ax=By\}$$

correspond formellement à minimiser la fonction Φ_γ avec $\gamma=0$. Dans l'article IV, nous remplaçons dans l'algorithme $(\mathcal{A}_1)$ le paramètre constant γ par une suite strictement positive (γ_n) qui tend vers 0. L'algorithme s'écrit

$$(\mathcal{A}_2)\begin{cases}x_{n+1}=\text{Argmin}\left\{\gamma_{n+1}\,f(x)+\frac{1}{2}\|Ax-By_n\|^2_{\mathcal{Z}}+\frac{\alpha}{2}\|x-x_n\|^2_{\mathcal{X}};\quad x\in\mathcal{X}\right\}\\ y_{n+1}=\text{Argmin}\left\{\gamma_{n+1}\,g(y)+\frac{1}{2}\|Ax_{n+1}-By\|^2_{\mathcal{Z}}+\frac{\nu}{2}\|y-y_n\|^2_{\mathcal{Y}};\quad y\in\mathcal{Y}\right\}.\end{cases}$$

La fonction $\Psi(x,y)=\frac{1}{2}\|Ax-By\|^2_{\mathcal{Z}}$ agit comme une fonction de pénalisation de la contrainte $Ax=By$ et $\frac{1}{\gamma_n}$ apparaît comme un paramètre de pénalisation. Sous des hypothèses adéquates, la suite générée par le nouvel algorithme $(\mathcal{A}_2)$ converge faiblement vers un point solution de $(\mathcal{P}_2)$, c'est-à-dire minimise la fonction $\Phi(x,y)=f(x)+g(y)$ sur $\text{Argmin}\,\Psi=\{(x,y)\in\mathcal{X}\times\mathcal{Y};\quad Ax=By\}$. Ce type de minimisation hierarchisée a été étudié par Cabot [9]. Soit $\psi:\mathbb{R}^m\to\mathbb{R}\cup\{+\infty\}$ une fonction s.c.i., convexe et $\varphi:\mathbb{R}^m\to\mathbb{R}$ une fonction finie et convexe. Avec ces notations, l'algorithme de [9] s'écrit

$$x_{n+1}=\text{Argmin}\left\{\psi(x)+\gamma_{n+1}\varphi(x)+\frac{\alpha_n}{2}\|x-x_n\|^2_{\mathbb{R}^m};\quad x\in\mathbb{R}^m\right\}.$$

La vitesse de convergence de la suite (γ_n) vers zéro joue un rôle primordial dans le processus de minimisation. Soit (w_n) la suite définie par

$$w_n=\inf_{x\in\mathbb{R}^m}\{\psi(x)+\gamma_{n+1}(\varphi(x)-\min\varphi)\},$$

Cabot [9] a obtenu, dans le cadre de la dimension finie, le résultat suivant:

Théorème 0.3 *Supposons que la suite (α_n) vérifie, pour tout $n \in \mathbb{N}$, $0 < \underline{\alpha} \leq \alpha_n \leq \overline{\alpha}$. Supposons que l'ensemble $C = Argmin\,\psi$ soit non vide, que la fonction φ soit minorée et que l'ensemble $Argmin_C\,\varphi$ soit non vide et borné. Si $(\gamma_n) \notin l^1$ alors*

(i) $\lim_{n\to+\infty} \psi(x_n) = \min \psi$ *et* $\lim_{n\to+\infty} \varphi(x_n) = \min_C \varphi$,

(ii) si de plus

$$(h_1) \qquad (w_n^-) \in l^1,$$

alors (x_n) converge vers un élément de $Argmin_C\,\varphi$.

L'hypothèse $(\gamma_n) \notin l^1$ exprime que la suite (γ_n) converge *lentement* vers zéro alors que l'hypothèse (h_1) exprime que la suite (γ_n) ne converge *pas trop lentement* vers zéro (sous des hypothèses adéquates sur les fonctions ψ et φ, (h_1) est réalisée si $(\gamma_n) \in l^2$). Nous utilisons des hypothèses similaires sur la suite (γ_n) dans l'article IV et montrons que la suite générée par l'algorithme $(\mathcal{A}_2)$ converge faiblement vers un point solution du problème $(\mathcal{P}_2)$. L'analyse est aussi étendue au cadre des opérateurs maximaux monotones. Avec le cadre d'application aux EDP précédent, la contrainte force le saut à travers l'interface à être nul et interdit les discontinuités à travers l'interface. L'algorithme permet de résoudre par décomposition le problème de minimisation suivant

$$\min\left\{\frac{1}{2}\int_\Omega |\nabla w|^2 - \int_\Omega hw; \quad w \in H^1(\Omega)\right\},$$

qui correspond à la formulation faible variationnelle du problème de Neumann sur le domaine Ω

$$\begin{cases} -\Delta w = h & \text{dans } \Omega \\ \frac{\partial w}{\partial n} = 0 & \text{sur } \partial\Omega. \end{cases}$$

Dans l'article V, la suite (γ_n) est supposée tendre vers $+\infty$. L'algorithme s'écrit

$$(\mathcal{A}_3)\begin{cases} x_{n+1} = \text{Argmin}\left\{f(x) + \dfrac{1}{2\gamma_{n+1}}\|Ax - By_n\|^2_{\mathcal{Z}} + \dfrac{\alpha}{2}\|x - x_n\|^2_{\mathcal{X}}; \quad x \in \mathcal{X}\right\} \\ y_{n+1} = \text{Argmin}\left\{g(y) + \dfrac{1}{2\gamma_{n+1}}\|Ax_{n+1} - By\|^2_{\mathcal{Z}} + \dfrac{\nu}{2}\|y - y_n\|^2_{\mathcal{Y}}; \quad y \in \mathcal{Y}\right\}. \end{cases}$$

Nous pouvons supposer sans perte de généralité que $\min f = \min g = 0$. Dans ce cas, c'est la fonction $\Phi(x, y) = f(x) + g(y)$ qui agit comme une fonction de pénalisation de la contrainte $\text{Argmin}\, f \times \text{Argmin}\, g$ et γ_n comme un paramètre de pénalisation. De manière symétrique à l'article précédent, nous retrouvons un processus de minimisation hierarchisée et la suite générée par le nouvel algorithme $(\mathcal{A}_3)$ converge faiblement vers un point solution de

$$(\mathcal{P}_3) \qquad \min\left\{\|Ax - By\|^2_{\mathcal{Z}}; \quad (x, y) \in \text{Argmin}\, f \times \text{Argmin}\, g\right\}.$$

Nous utilisons une hypothèse introduite par Attouch et Czarnecki [6]. Les auteurs ont étudié le système dynamique continu suivant

$$(\mathcal{S}_2) \qquad \dot{x}(t) + \partial\varphi(x(t)) + \frac{1}{\gamma(t)}\partial\psi(t) \ni 0,$$

où H est un espace de Hilbert, $\varphi : H \to \mathbb{R}_+ \cup \{+\infty\}$ et $\psi : H \to \mathbb{R} \cup \{+\infty\}$ sont des fonctions s.c.i., convexes et propres et $\gamma : \mathbb{R}_+ \to \mathbb{R}_+\backslash\{0\}$ est une fonction de classe C^1 telle que $\lim_{t\to+\infty} \gamma(t) = +\infty$. Les auteurs ont obtenu le résultat suivant:

Théorème 0.4 *Supposons de plus que $C = Argmin\,\varphi = \varphi^{-1}(0) \neq \varnothing$ et que $Argmin_C\,\psi \neq \varnothing$. Soit $\gamma : \mathbb{R}_+ \to \mathbb{R}_+\backslash\{0\}$ une fonction de classe C^1 croissante telle que $\dot{\gamma}$ soit majorée et*[2]

$$(h_2) \qquad \forall p \in R(N_C), \quad \left(\varphi^*\left(\frac{p}{\gamma(t)}\right) - \sigma_C\left(\frac{p}{\gamma(t)}\right)\right) \in L^1(0,+\infty).$$

Soit x une solution forte[3] *du système $(\mathcal{S}_2)$. Alors*

(i) $\lim_{t\to+\infty} \varphi(t) = 0$ *et* $\lim_{t\to+\infty} \psi(t) = \min_C \psi$,

(ii) x converge faiblement dans H vers un point de $Argmin_C\,\psi$.

La vitesse de croissance de la fonction γ vers l'infini joue là encore un rôle primordial pour assurer la convergence des trajectoires vers l'ensemble d'équilibre. L'hypothèse $\dot{\gamma}$ majorée exprime que la fonction γ tend *lentement* vers l'infini alors que l'hypothèse (h_2) exprime que la fonction γ ne tend *pas trop lentement* vers l'infini (sous des hypothèses adéquates sur la fonction φ, (h_2) est réalisée si $\frac{1}{\gamma} \in L^2(0,+\infty)$). Avec des hypothèses analogues traduites dans le cas discret, nous obtenons la convergence faible de la suite générée par l'algorithme $(\mathcal{A}_3)$ vers un point solution du problème $(\mathcal{P}_3)$. Dans le cadre des EDP, l'algorithme permet de résoudre le problème de minimisation suivant

$$\min\left\{\frac{1}{2}\int_\Gamma [w]^2\right\},$$

où $[w]$ est le saut de w à travers l'interface Γ, $w = \begin{cases} u & \text{sur } \Omega_1 \\ v & \text{sur } \Omega_2 \end{cases}$ et $u \in H^1(\Omega_1)$, $v \in H^1(\Omega_2)$ sont solutions faibles des problèmes avec conditions aux bords de Neumann suivants

---

[2] $N_C(x)$ est le cône normal à C en x,

$$N_C(x) = \{p \in \mathcal{X} : \langle p, \zeta - x\rangle_{\mathcal{X}} \leq 0 \quad \forall \zeta \in \mathcal{X}\}.$$

$R(N_C)$ est l'image de N_C, c'est-à-dire $p \in R(N_C)$ si et seulement s'il existe un $x \in C$ tel que $p \in N_C(x)$. σ_C est la fonction support de C: pour tout $x \in \mathcal{X}$, $\sigma_C(x) = \sup_{\zeta\in\mathcal{X}} \langle x, \zeta\rangle_{\mathcal{X}}$.

[3] Dans le sens de Brezis ([8], définition 3.1). En particulier, x est absolument continue sur tout intervalle $[0; T]$ avec $T < +\infty$.

$$\begin{cases} -\Delta u = h & \text{dans} \quad \Omega_1 \\ \frac{\partial u}{\partial n} = 0 & \text{sur} \quad \partial\Omega_1, \end{cases} \qquad \begin{cases} -\Delta v = h & \text{dans} \quad \Omega_2 \\ \frac{\partial v}{\partial n} = 0 & \text{sur} \quad \partial\Omega_2, \end{cases}$$

et $h \in L^2(\Omega)$ est une fonction donnée.

Enfin, dans le dernier article, nous utilisons des méthodes proximales et lagrangiennes inspirées des articles [12, 7] dans le but de résoudre le problème

$$(\mathcal{P}_4) \qquad \min \{ f(x); \quad Ax \in \mathcal{C} \},$$

où $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ est une fonction s.c.i, convexe et propre, $A : \mathcal{X} \to \mathcal{Y}$ est un opérateur linéaire continu et $\mathcal{C}$ est un ensemble convexe fermé de $\mathcal{Y}$. Dans l'article [12], Chen et Teboulle ont considéré le problème de minimisation avec contraintes linéaires suivant

$$(\mathcal{Q}_4) \qquad \min \{ f(x) + g(y); \quad Ax = y \},$$

où $f : \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$ et $g : \mathbb{R}^p \to \mathbb{R} \cup \{+\infty\}$ sont des fonctions s.c.i, convexes et propres et $A : \mathbb{R}^m \to \mathbb{R}^p$ est un opérateur linéaire. La fonction de Lagrange associée au problème $(\mathcal{Q}_4)$ est la fonction $\mathcal{L} : \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^p \to \mathbb{R} \cup \{+\infty\}$ définie par

$$\mathcal{L}(x, y, \mu) = f(x) + g(y) + \langle \mu, Ax - y \rangle_{\mathbb{R}^p}.$$

Elle est s.c.i, convexe pour les variables primales x et y et concave pour la variable duale μ. Les auteurs ont introduit un algorihme basé sur une minimisation proximale pour les variables x et y et sur une maximisation proximale pour la variable μ. Par linéarité, l'algorithme s'écrit

$$\begin{cases} \tilde{\mu}_{n+1} = \mu_n + \lambda_{n+1}(Ax_n - y_n), \\ x_{n+1} = \text{Argmin} \left\{ \mathcal{L}(x, y_n, \tilde{\mu}_{n+1}) + \frac{1}{2\lambda_{n+1}} \|x - x_n\|^2_{\mathbb{R}^m}; \quad x \in \mathbb{R}^m \right\}, \\ y_{n+1} = \text{Argmin} \left\{ \mathcal{L}(x_n, y, \tilde{\mu}_{n+1}) + \frac{1}{2\lambda_{n+1}} \|y - y_n\|^2_{\mathbb{R}^p}; \quad y \in \mathbb{R}^p \right\}, \\ \mu_{n+1} = \mu_n + \lambda_{n+1}(Ax_{n+1} - y_{n+1}). \end{cases}$$

Le résultat principal obtenu dans [12] est:

Théorème 0.5 *Supposons que l'ensemble des points selles*[4] *de $\mathcal{L}$ soit non vide et que la suite (λ_n) vérifie, pour tout $n \in \mathbb{N}$,*

$$\epsilon \leq \lambda_n \leq \min \left(\frac{1-\epsilon}{2}, \frac{1-\epsilon}{2\|A\|} \right),$$

pour $0 < \epsilon \leq \min \left(\frac{1}{3}, \frac{1}{2\|A\|+1} \right)$. Alors (x_n, y_n, μ_n) converge vers un point selle de $\mathcal{L}$ et donc (x_n, y_n) converge vers un point solution du problème primal $(\mathcal{Q}_4)$.

[4] $(x^*, y^*, \mu^*) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^p$ est un point selle de $\mathcal{L}$ si, pour tout $(x, y, \mu) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^p$,

$$\mathcal{L}(x^*, y^*, \mu) \leq \mathcal{L}(x^*, y^*, \mu^*) \leq \mathcal{L}(x, y, \mu^*).$$

Dans l'article VI, nous étendons ces résultats de convergence à la dimension infinie pour le problème de minimisation $(\mathcal{P}_4)$. Pour cela, nous introduisons une variable de contrainte $\nu \in \mathbb{R}^q$ et une fonction de pénalisation $P : \mathcal{Y} \to (\mathbb{R}^+)^q$ telle que $y \in \mathcal{C}$ si et seulement si $P(y) = 0$. La fonction de Lagrange considérée est la fonction $L : \mathcal{X} \times \mathcal{Y} \times \mathcal{Y} \times \mathbb{R}^q \to \mathbb{R} \cup \{+\infty\}$ définie par $L(x, y, \mu, \nu) = f(x) + \langle \mu, Ax - y \rangle_{\mathcal{Y}} + \langle \nu, P(y) \rangle_{\mathbb{R}^q}$. Nous introduisons un algorithme inspiré de [12]. Nous montrons que, si la fonction P est lipschitzienne et sous des hypothèses adéquates sur la suite (λ_n), la suite (x_n, y_n, μ_n, ν_n) générée par cet algorithme converge faiblement dans $\mathcal{X} \times \mathcal{Y} \times \mathcal{Y} \times \mathbb{R}^q$ vers un point selle de L et donc (x_n, y_n) converge faiblement dans $\mathcal{X} \times \mathcal{Y}$ vers un point solution de $(\mathcal{P}_4)$. L'étude est aussi étendue au cadre des opérateurs maximaux monotones. L'algorithme permet de résoudre le problème de minimisation suivant

$$\min \left\{ \frac{1}{2} \int_{\Omega_1} |\nabla u|^2 - \int_{\Omega_1} hu + \frac{1}{2} \int_{\Omega_2} |\nabla v|^2 - \int_{\Omega_2} hv; \quad (u, v) \in H^1(\Omega_1) \times H^1(\Omega_2) \text{ et } u_{|\Gamma} \geq v_{|\Gamma} \right\}.$$

Ce type de problème peut intervenir dans la description de phénomènes faisant intervenir un matériau semi-conducteur ou un système de valve.

References

[1] F. Acker, M.A. Prestel, Convergence d'un schéma de minimisation alternée, *Annales de la Faculté des Sciences de Toulouse V, série Mathématiques*, 2, (1980), pp. 1-9.

[2] F. Alvarez, On the minimizing property of a second order dissipative system in Hilbert spaces, *SIAM J. on Control and Optimization*, 38 (2000), n° 4, pp. 1102-1119.

[3] F. Alvarez, H. Attouch, Convergence and asymptotic stabilization for some damped hyperbolic equations with non-isolated equilibria, *ESAIM: COCV*, 6 (2001), pp. 539-552.

[4] F. Alvarez, H. Attouch, An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping, *Set-Valued Analysis*, vol. 9 (2001), pp. 3-11.

[5] H. Attouch, J. Bolte, P. Redont, A. Soubeyran, Alternating proximal algorithms for weakly coupled convex minimization problems. Applications to dynamical games and PDE's, *Journal of Convex Analysis*, 15 (2008), No 3, pp. 485-506.

[6] H. Attouch, M.-O. Czarnecki, Asymptotic behavior of coupled dynamical systems with multiscale aspects, *Journal of Differential Equations*, 248 (2010), pp. 1315-1344.

[7] H. Attouch, M. Soueycatt, Augmented Lagrangian and proximal alternating direction methods of multipliers in Hilbert spaces. Applications to games, PDE's and control, *Pacific Journal of Optimization*, 5 (2009), No 1, pp. 17-37.

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

[9] A. Cabot, Proximal point algorithm controlled by a slowly vanishing term: Applications to hierarchical minimization, *SIAM Journal on Optimization*, 15 (2005), No 2, pp. 555-572.

[10] A. Cabot, H. Engler, S. Gadat, On the long time behavior of second order differential equations with asymptotically small dissipation, *Trans. Amer. Math. Soc.*, 361 (2009), no. 11, pp. 5983-6017.

[11] A. Cabot, H. Engler, S. Gadat, Second order differential equations with asymptotically small dissipation and piecewise flat potentials, *Electronic Journal of Differential Equations*, 17 (2009), 33-38.

[12] G. Chen, M. Teboulle, A proximal-based decomposition method for convex minimization problems, *Math. Programming* **64** (1994), no. 1, Ser. A, 81-101.

I. COMPORTEMENT ASYMPTOTIQUE POUR CERTAINES EQUATIONS HYPERBOLIQUES SEMILINEAIRES AVEC TERME D'AMORTISSEMENT NON-AUTONOME

Asymptotics for some semilinear hyperbolic equations with non-autonomous damping

A. Cabot[1] and P. Frankel[2]

[1] Université Montpellier II, Place Eugène Bataillon, 34095 Montpellier Cedex 05, France
acabot@math.univ-montp2.fr
[2] Université Montpellier II, Place Eugène Bataillon, 34095 Montpellier Cedex 05, France
p.frankel30@orange.fr

Accepted in *JDE*

Summary. Let V and H be Hilbert spaces such that $V \subset H \subset V'$ with dense and continuous injections. Consider a linear continuous operator $A : V \to V'$ which is assumed to be symmetric, monotone and semi-coercive. Given a function $f : V \to H$ and a map $\gamma \in W^{1,1}_{loc}(\mathbb{R}_+, \mathbb{R}_+)$ such that $\lim_{t\to+\infty} \gamma(t) = 0$, our purpose is to study the asymptotic behavior of the following semilinear hyperbolic equation

$$(E) \qquad \frac{d^2u}{dt^2}(t) + \gamma(t)\frac{du}{dt}(t) + Au(t) + f(u(t)) = 0, \quad t \geq 0.$$

The nonlinearity f is assumed to be monotone and conservative. Condition $\int_0^{+\infty} \gamma(t)\,dt = +\infty$ guarantees that some suitable energy function tends toward its minimum. The main contribution of this paper is to provide a general result of convergence for the trajectories of (E): if the quantity $\gamma(t)$ behaves as k/t^α, for some $\alpha \in]0, 1[$, $k > 0$ and t large enough, then $u(t)$ weakly converges in V toward an equilibrium as $t \to +\infty$. Strong convergence in V holds true under compactness or symmetry conditions. We also give estimates for the speed of convergence of the energy under some ellipticity-like conditions. The abstract results are applied to particular semilinear evolution problems at the end of the paper.

Key words: Semilinear evolution problem, dissipative hyperbolic equation, non-autonomous damping, asymptotic behavior, rate of convergence.
Subject classification:34G10, 34G20, 35B40, 35L70.

1 Introduction

Throughout this paper, V stands for a real Hilbert space, whose scalar product and norm are respectively denoted by $((\cdot,\cdot))$ and $\|\cdot\|$. Let H be another real Hilbert space with scalar product $(\cdot,\cdot)$ and norm $|\cdot|$. Suppose that V is dense in H with continuous injection. By duality, the topological dual space H' of H is identified with a dense subspace of the topological dual V' of V. Identifying H with H', we obtain $V \subset H \subset V'$, where each space is dense in the next one, each injection being continuous. We denote by $\langle\cdot,\cdot\rangle_{V',V}$ the duality pairing between V' and V. Let $a : V \times V \to \mathbb{R}$ be a continuous bilinear form satisfying

(h_1) $a(.,.)$ is symmetric, positive,

(h_2) $\exists \lambda \geq 0, \mu > 0$ such that $\forall u \in V, \quad a(u,u) + \lambda |u|^2 \geq \mu \|u\|^2.$

This last property is known as the semi-coercivity of the form a. We associate with $a(.,.)$ the continuous operator $A : V \to V'$ defined by $\langle Au, v\rangle_{V',V} = a(u,v)$ for all $u, v \in V$. We denote by $D(A)$ the domain of the operator A, *i.e.* $D(A) = \{v \in V;\ Av \in H\}$. Given a function $f : V \to H$ and a map $\gamma \in W^{1,1}_{loc}(\mathbb{R}_+, \mathbb{R}_+)$, we consider the following semilinear evolution equation of second-order in time

$$(E) \qquad \frac{d^2u}{dt^2}(t) + \gamma(t)\frac{du}{dt}(t) + Au(t) + f(u(t)) = 0, \quad t \geq 0.$$

The nonlinearity f is assumed to be conservative, *i.e.* derives from some potential $F \in \mathcal{C}^1(V, \mathbb{R})$. The main purpose of the paper is to investigate the asymptotic behavior of the trajectories of (E) for a vanishing damping term, *i.e.* $\gamma(t) \to 0$ as $t \to +\infty$. It is clear that the decay properties of the map γ play a central role in the analysis. In particular, if the quantity $\gamma(t)$ tends to 0 too rapidly as $t \to +\infty$, convergence of the trajectories may fail. To motivate our study, let us show how it is connected to other questions of interest.

Case of a constant damping. If $\gamma(t) \equiv \gamma$, existence and uniqueness are well-known in the framework of damped wave equations. More precisely, if the map $f : V \to H$ is Lipschitz continuous on the bounded sets of V and if the map F satisfies suitable growth conditions, then for any $(u_0, v_0) \in D(A) \times V$, there exists a unique solution $u \in W^{1,\infty}_{loc}(\mathbb{R}_+, V) \cap W^{2,\infty}_{loc}(\mathbb{R}_+, H)$ of (E) such that $u(0) = u_0$ and $\frac{du}{dt}(0) = v_0$, see [12, Theorem II.3.2.1] or [20, Ch. IV, Theorem 4.1]. The trajectories of (E) are known to converge toward an equilibrium point $u_\infty \in \{v \in V,\ Av + f(v) = 0\}$ under assumptions like monotonicity or analyticity. In the case of a monotone map f, convergence is obtained for the weak topology of V and the main ingredient of the proof is the Opial lemma, *cf.* [3]. When the nonlinearity is analytic, convergence of the trajectories relies on the Lojasiewicz inequality, see [15, 16] and the pioneering work [19] for parabolic problems.

Averaged heat equation. With the same assumptions as above, consider the abstract heat equation

$$\frac{dv}{ds}(s) + Av(s) = 0, \quad s \geq 0. \tag{1}$$

It may be of interest to examine the case where the velocity $\frac{dv}{ds}(s)$ is proportional, not to the instantaneous vector $Av(s)$, but to some average taken over the interval $[0, s]$. The simplest such equation is

$$\frac{dv}{ds}(s) + \frac{1}{s}\int_0^s Av(\sigma)\, d\sigma = 0, \quad s > 0. \tag{2}$$

After multiplying this equality by s and differentiating, we obtain the following second-order in time equation

$$s\,\frac{d^2v}{ds^2}(s) + \frac{dv}{ds}(s) + Av(s) = 0, \quad s > 0.$$

The change of variable $s = \frac{t^2}{4}$ allows to rewrite the above equation as

$$\frac{d^2u}{dt^2}(t) + \frac{1}{t}\,\frac{du}{dt}(t) + Au(t) = 0, \quad t > 0,$$

where the map u is defined by $u(t) = v\left(\frac{t^2}{4}\right)$ for every $t \geq 0$. This is exactly equation (E) with $\gamma(t) = \frac{1}{t}$ and $f \equiv 0$. Assuming that the injection $V \hookrightarrow H$ is compact, there exists a nondecreasing sequence $(\lambda_i)_{i\geq 1}$ of eigenvalues of A, along with a complete orthonormal basis of H, $(e_i)_{i\geq 1}$ consisting of the corresponding eigenvectors. Let $u(t) = \sum_{i=1}^{+\infty} u_i(t)\, e_i$ be the decomposition of the solution $u(t)$ on the basis of eigenfunctions. Every component u_i satisfies the following equation

$$\ddot{u}_i(t) + \frac{1}{t}\,\dot{u}_i(t) + \lambda_i\, u_i(t) = 0, \quad t > 0.$$

It ensues that each kernel component u_i, $i \in \{1, \dots, \dim(\mathrm{Ker} A)\}$ verifies $u_i(t) = a_i \ln t + b_i$, for some $a_i, b_i \in \mathbb{R}$. In particular, it cannot converge as $t \to +\infty$, unless it is stationary. When the eigenvalue λ_i is positive, we let the reader check that

$$u_i(t) = a_i'\, J_0\left(\sqrt{\lambda_i}\, t\right) + b_i'\, Y_0\left(\sqrt{\lambda_i}\, t\right), \qquad \text{for some } a_i', b_i' \in \mathbb{R},$$

where J_0 and Y_0 denote respectively the zeroth Bessel functions of the first and second kind[3]. Recalling that

$$J_0(t) \sim \sqrt{\frac{2}{\pi\, t}}\,\cos\left(t - \frac{\pi}{4}\right) \quad \text{and} \quad Y_0(t) \sim \sqrt{\frac{2}{\pi\, t}}\,\sin\left(t - \frac{\pi}{4}\right) \quad \text{as } t \to +\infty,$$

we deduce that $u_i(t) \sim \frac{c_i}{\sqrt{t}}\,\cos(\sqrt{\lambda_i}\, t - \varphi_i)$ as $t \to +\infty$, for some c_i, $\varphi_i \in \mathbb{R}$. Coming back to the averaged heat equation (2), we then obtain for each component v_i

$$v_i(s) \sim \frac{c_i}{\sqrt{2}}\, s^{-\frac{1}{4}}\,\cos\left(2\sqrt{\lambda_i\, s} - \varphi_i\right) \qquad \text{as } s \to +\infty.$$

It converges toward zero much more slowly than the corresponding component of the “pure” heat equation, equal to $v_i(0)\, e^{-\lambda_i s}$. The above discussion shows that the global behavior of (2) -or more generally (E)- differs considerably from the one of equation (1).

[3] See [1, 5] for standard references on Bessel equations.

Heavy ball with asymptotically small friction. Given a continuous map $\gamma : \mathbb{R}_+ \to \mathbb{R}_+$ and a potential $\Phi : H \to \mathbb{R}$ of class $\mathcal{C}^1$ with a locally Lipschitz gradient, let us consider the following ordinary differential equation in the Hilbert space H

$$\ddot{x}(t) + \gamma(t)\,\dot{x}(t) + \nabla\Phi(x(t)) = 0, \quad t \geq 0. \tag{3}$$

When $\gamma(t) \equiv \gamma > 0$, the above equation is known under the terminology of "Heavy Ball with Friction" system, (HBF) for short. From a mechanical point of view, (HBF) corresponds to the equation describing the motion of a material point subjected to the conservative force $-\nabla\Phi(x)$ and the viscous friction force $-\gamma\,\dot{x}$. The (HBF) system can be studied in the classical framework of the theory of dissipative dynamical systems, *cf.* [11, 13]. The trajectories of (HBF) are known to converge toward a critical point of Φ under various assumptions (see [2, 4] for convex potentials and [14] for analytic ones). In the recent papers [8, 9], it is considered the case of a vanishing damping $\gamma(t) \to 0$ as $t \to +\infty$. The corresponding equation is typically obtained from a first-order gradient system involving some memory aspects, see [7]. If the function Φ is convex and has a unique minimum $\overline{x}$, condition $\int_0^{+\infty} \gamma(t)\,dt = +\infty$ is sufficient to ensure (weak) convergence of the trajectories of (3) toward $\overline{x}$. When the function Φ has a continuum of equilibria, the more stringent condition $\displaystyle\int_0^{+\infty} e^{-\int_0^t \gamma(s)\,ds} dt < +\infty$ is necessary to obtain convergence of the trajectories. In the one-dimensional case, the slightly stronger condition $\displaystyle\int_0^{+\infty} e^{-\theta \int_0^t \gamma(s)\,ds} dt < +\infty$, for some $\theta \in]0,1[$ is shown to be sufficient. In the higher-dimensional case, the general question of convergence is left open in [8, 9]. The new techniques developed in the present paper allow to address this question and to fill partially the gap between necessary and sufficient conditions for convergence, see comments below.

Let us come back to equation (E) and precise now the framework of the paper. The nonlinearity f is assumed to be monotone and conservative, *i.e.* derives from some convex potential $F \in \mathcal{C}^1(V, \mathbb{R})$. The set of equilibria $S = \{v \in V,\ Av + f(v) = 0\}$ is supposed to be nonempty. It is not our purpose to develop the well-posedness of equation (E) for given initial conditions. Throughout the paper, we assume the existence of a solution to equation (E) in the class

$$u \in W^{1,1}_{loc}(\mathbb{R}_+, V) \cap W^{2,1}_{loc}(\mathbb{R}_+, H). \tag{4}$$

We define the energy function $\mathcal{E}$ along each trajectory by

$$\mathcal{E}(t) = \frac{1}{2}\left|\frac{du}{dt}(t)\right|^2 + \frac{1}{2}\,a(u(t), u(t)) + F(u(t)).$$

The major contribution of this paper is to provide a result of (weak) convergence in V for the trajectories of (E): if the quantity $\gamma(t)$ behaves as k/t^α, for some $\alpha \in]0,1[$, $k > 0$ and t large enough, there exists an equilibrium $u_\infty \in S$ such that $u(t) \rightharpoonup u_\infty$ weakly in V as $t \to +\infty$. The exact statement is in fact slightly more general, see Theorem 3.3. The main ingredients of the proof are the Opial lemma along with accurate estimates of the energy decay, *cf.* Proposition 3.2. Strong convergence in V holds true under compactness or symmetry conditions. The technique of the proof is new and is also applicable to the differential equation (3).

The second contribution of the paper is to give sharp estimates for the speed of convergence of the energy $\mathcal{E}(t)$ as $t \to +\infty$. In the linear case ($f = 0$) and under some ellipticity-like condition, we obtain the following estimate

$$\mathcal{E}(t) \sim K\, e^{-\int_0^t \gamma(s)ds} \quad \text{as} \quad t \to +\infty, \qquad \text{for some } K > 0. \tag{5}$$

Notice that this estimate fails to be true if the trajectory is contained in Ker$A$, see Theorem 2.1 for a precise statement. In the nonlinear case, the same kind of estimate is obtained at a slightly lower degree of precision[4], *cf.* Theorem 3.4.

Outline of the paper. Section 2 is concerned with the linear hyperbolic equation (E_0) obtained by taking $f = 0$ in (E). We analyze the behavior of the trajectories by studying respectively their components with respect to the spaces Ker$A$ and $(\mathrm{Ker}A)^\perp$. A sharp estimate of the energy decay is given under some ellipticity-like condition. In section 3, we deal with the general equation (E) by assuming that the nonlinearity f is monotone. It is shown in paragraph 3.1 that the energy $\mathcal{E}(t)$ vanishes as $t \to +\infty$, which allows to prove (weak) convergence of the trajectories in the case of a unique minimum. The general problem of convergence for a continuum of minima is treated in paragraph 3.2, which is the core of the paper. Additional results of strong convergence in V are given under some compactness or symmetry assumptions. Finally, the abstract results are applied to particular semilinear evolution problems in section 4.

2 Linear hyperbolic equation

Let $a : V \times V \to \mathbb{R}$ be a continuous bilinear form satisfying (h_1)-(h_2) and let $A : V \to V'$ be the associate operator. Given a map $\gamma \in W^{1,1}_{loc}(\mathbb{R}_+, \mathbb{R}_+)$, we consider the following linear hyperbolic equation

$$(E_0) \qquad \frac{d^2u}{dt^2}(t) + \gamma(t)\frac{du}{dt}(t) + Au(t) = 0, \quad t \geq 0.$$

We assume the existence of a solution to equation (E_0) in the class (4). We define the energy function $\mathcal{E}$ along each trajectory by

[4] In this case, a factor $\frac{2}{3}$ has to be introduced in the exponent of formula (5).

$$\mathcal{E}(t) = \frac{1}{2}\left|\frac{du}{dt}(t)\right|^2 + \frac{1}{2}a(u(t), u(t)).$$

We have $\mathcal{E} \in W^{1,1}_{loc}(\mathbb{R}_+)$ and

$$\begin{aligned}\dot{\mathcal{E}}(t) &= \left(\frac{d^2u}{dt^2}(t), \frac{du}{dt}(t)\right) + \left\langle Au(t), \frac{du}{dt}(t)\right\rangle_{V',V} \\ &= -\gamma(t)\left|\frac{du}{dt}(t)\right|^2 \leq 0 \quad \text{a.e. on } \mathbb{R}_+,\end{aligned}$$

hence the function $\mathcal{E}$ is a Lyapunov function for the system (E_0). The purpose of this section is to establish results of convergence for the trajectory u, along with estimates of the energy decay. For every $t \geq 0$, we set $\widehat{u}(t) = Pu(t)$, where P denotes the orthogonal projection onto the subspace[5] Ker$A$ in the sense of H. Since $\widehat{u}(t) \in$ Ker$A$ for every $t \geq 0$, we have

$$\forall t \geq 0, \quad \frac{d^2\widehat{u}}{dt^2}(t) + \gamma(t)\frac{d\widehat{u}}{dt}(t) = 0.$$

By integrating this equality twice, we find

$$\begin{aligned}\forall t \geq 0, \quad \widehat{u}(t) &= \widehat{u}(0) + \left(\int_0^t e^{-\int_0^s \gamma(\tau)d\tau} ds\right)\frac{d\widehat{u}}{dt}(0) \qquad (6)\\ &= Pu_0 + \left(\int_0^t e^{-\int_0^s \gamma(\tau)d\tau} ds\right)Pv_0.\end{aligned}$$

If $Pv_0 \neq 0$, the above equality shows that the asymptotic behavior of the component $\widehat{u}$ is strongly related with the convergence of the integral $\int_0^{+\infty} e^{-\int_0^s \gamma(\tau)d\tau} ds$. The next proposition summarizes the different possible cases.

Proposition 2.1 *Let us set* $\overline{\omega} = \int_0^{+\infty} e^{-\int_0^s \gamma(\tau)d\tau} ds \in \mathbb{R}_+ \cup \{+\infty\}$.

- *If* $v_0 \in (KerA)^\perp$, *then* $\widehat{u}(t) = Pu_0$ *for every* $t \geq 0$.
- *If* $v_0 \notin (KerA)^\perp$, *then the solution* $\widehat{u}$ *converges if and only if* $\overline{\omega} < +\infty$. *More precisely, we have* $\lim_{t\to+\infty} |\widehat{u}(t)| = +\infty$ *if* $\overline{\omega} = +\infty$ *while* $\lim_{t\to+\infty} \widehat{u}(t) = P(u_0+\overline{\omega}v_0)$ *if* $\overline{\omega} < +\infty$.

Our purpose is now to evaluate the energy decay along each trajectory $u(.)$. We start with a preliminary result corresponding to the case Ker$A = \{0\}$.

Lemma 2.1 *Assume that the bilinear form* $a(.,.)$ *satisfies* (h_1)-(h_2) *and that*

$$\exists \eta > 0, \forall u \in V, \quad a(u,u) \geq \eta |u|^2. \qquad (7)$$

Let $\gamma \in W^{1,1}_{loc}(\mathbb{R}_+, \mathbb{R}_+)$ *be a function such that* $\lim_{t\to+\infty}\gamma(t) = 0$ *and* $\dot{\gamma} \in L^1(0,+\infty)$. *Let* u *be a solution in the class (4) to equation* (E_0). *Then, either the solution* u *is stationary, or there exists* $K > 0$ *such that*

$$\mathcal{E}(t) \sim K e^{-\int_0^t \gamma(s)ds} \quad as \quad t \to +\infty.$$

[5] By using assumptions (h_1)-(h_2), it is easy to check that Ker$A$ is closed in H. See also Remark 3.2.

Proof. The main idea of the proof consists in using the function $\mathcal{F}$ defined by[6]

$$\begin{aligned}\mathcal{F}(t) &= \frac{1}{2}\left|\frac{du}{dt}(t)\right|^2 + \frac{1}{2}a(u(t),u(t)) + \frac{\gamma(t)}{2}\left(\frac{du}{dt}(t),u(t)\right)\\ &= \mathcal{E}(t) + \frac{\gamma(t)}{2}\left(\frac{du}{dt}(t),u(t)\right).\end{aligned}$$

We have $\mathcal{F} \in W^{1,1}_{loc}(\mathbb{R}_+)$ and by differentiating the function $\mathcal{F}$, we find for almost every $t \geq 0$

$$\begin{aligned}\dot{\mathcal{F}}(t) &= \dot{\mathcal{E}}(t) + \frac{\dot\gamma(t)}{2}\left(\frac{du}{dt}(t),u(t)\right) + \frac{\gamma(t)}{2}\left(\frac{d^2u}{dt^2}(t),u(t)\right) + \frac{\gamma(t)}{2}\left|\frac{du}{dt}(t)\right|^2\\ &= -\frac{\gamma(t)}{2}\left|\frac{du}{dt}(t)\right|^2 - \frac{\gamma(t)}{2}a(u(t),u(t)) + \left(\frac{\dot\gamma(t)}{2} - \frac{\gamma(t)^2}{2}\right)\left(\frac{du}{dt}(t),u(t)\right).\end{aligned}$$

Therefore we have

$$\dot{\mathcal{F}}(t) + \gamma(t)\mathcal{F}(t) = \frac{\dot\gamma(t)}{2}\left(\frac{du}{dt}(t),u(t)\right) \qquad \text{a.e. on } \mathbb{R}_+. \tag{8}$$

Since $\left|\left(\frac{du}{dt}(t),u(t)\right)\right| \leq \frac{1}{2}\left|\frac{du}{dt}(t)\right|^2 + \frac{1}{2}|u(t)|^2$ and $a(u(t),u(t)) \geq \eta\,|u(t)|^2$ by assumption (7), we have

$$\left|\left(\frac{du}{dt}(t),u(t)\right)\right| \leq C\,\mathcal{E}(t), \qquad \text{for some } C > 0. \tag{9}$$

Recalling that $\lim_{t\to+\infty}\gamma(t) = 0$, the expression of $\mathcal{F}$ shows that

$$\mathcal{F}(t) \sim \mathcal{E}(t) \qquad \text{as } t \to +\infty. \tag{10}$$

We deduce from (8), (9) and (10) the existence of $D > 0$ and $t_0 \geq 0$ such that

$$\left|\dot{\mathcal{F}}(t) + \gamma(t)\mathcal{F}(t)\right| \leq D\,|\dot\gamma(t)|\mathcal{F}(t) \qquad \text{a.e. on } [t_0,+\infty[.$$

Let us multiply each member of this inequality by $e^{\int_0^t \gamma(s)ds}$ and set $\mathcal{G}(t) = e^{\int_0^t \gamma(s)ds}\mathcal{F}(t)$. We obtain

$$|\dot{\mathcal{G}}(t)| \leq D\,|\dot\gamma(t)|\mathcal{G}(t) \qquad \text{a.e. on } [t_0,+\infty[. \tag{11}$$

Observe that if $\mathcal{G}(t_1) = 0$ for some $t_1 \geq t_0$, then we have $\mathcal{F}(t_1) = 0$ and $\mathcal{E}(t_1) = 0$. Since the map $\mathcal{E}$ is nonincreasing, we conclude that $\mathcal{E}(t) = 0$ for every $t \geq t_1$, *i.e.* the solution u is stationary. Now assume that $\mathcal{G}(t) > 0$ for every $t \geq t_0$ and divide

[6] The use of such an auxiliary function is classical, see for example [13, Lemma 3.2.6] in the case of an autonomous damping.

each member of equality (11) by $\mathcal{G}(t)$. Since $\dot\gamma \in L^1(0,+\infty)$ by assumption, we deduce that

$$\left|\frac{d}{dt}\ln\mathcal{G}\right|(t) = \frac{|\dot{\mathcal{G}}(t)|}{\mathcal{G}(t)} \quad \in L^1(0,+\infty).$$

It ensues that $\lim_{t\to+\infty} \ln\mathcal{G}(t)$ exists in $\mathbb{R}$. We deduce that $\lim_{t\to+\infty} e^{\int_0^t \gamma(s)ds}\mathcal{F}(t) = K > 0$. The conclusion immediately follows from estimate (10).

Remark 2.1 *A result similar to Lemma 2.1 can be obtained by eliminating the first order term in* (E_0) *via the change of variable* $v(t) = e^{\frac{1}{2}\int_0^t \gamma(s)ds}u(t)$. *The details are left to the reader.*

Remark 2.2 (Case γ constant) *Assuming that* $\gamma(t) \equiv \gamma > 0$ *and that* $a(u,u) \geq \eta\,|u|^2$ *for every* $u \in V$, *the estimate* $\mathcal{E}(t) = O\left(e^{-\gamma t}\right)$ *remains true as* $t \to +\infty$ *if* $\gamma < 2\,\eta^{1/2}$, *see [13, Lemma 3.2.6]. However, it fails to be valid if* $\gamma \geq 2\,\eta^{1/2}$, *see [13, Proposition 3.2.5].*

We now assume the following ellipticity-like condition

$$\forall u \in V, \quad a(u,u) \geq \eta\,|u - Pu|^2, \quad \text{for some } \eta > 0. \tag{12}$$

Remark 2.3 *Under* (h_2), *this condition is equivalent to the following one*[7]

$$\forall u \in V, \quad a(u,u) \geq \eta'\,\|u - Pu\|^2, \quad \textit{for some } \eta' > 0. \tag{13}$$

Indeed, assume that condition (12) is satisfied. Recalling that $Pu \in KerA$, *we deduce from* (h_2) *that*

$$\forall u \in V, \quad a(u,u) + \lambda\,|u - Pu|^2 \geq \mu\,\|u - Pu\|^2.$$

It ensues that $\left(1 + \frac{\lambda}{\eta}\right) a(u,u) \geq \mu\,\|u - Pu\|^2$ *for every* $u \in V$ *and finally (13) is fulfilled with* $\eta' = \frac{\eta\,\mu}{\eta+\lambda}$.

Remark 2.4 *Suppose that the injection* $V \hookrightarrow H$ *is compact and that* (h_1)-(h_2) *hold true. The eigenvalues of* A *then define a nondecreasing sequence of nonnegative scalars tending to* $+\infty$ *and there exists an orthonormal basis of* H *consisting of the corresponding eigenvectors, see for example [17, 20]. If* η *denotes the smallest eigenvalue of* A *greater than* 0, *it is clear that* $a(u,u) \geq \eta|u|^2$ *for every* $u \in (KerA)^{\perp} \cap V$ *and therefore condition (12) holds true.*

The next result allows to estimate the energy decay under condition (12).

---

[7] Condition (13) is used in [21, Section 4], where estimates of the energy decay are provided in the case of an autonomous damping.

Theorem 2.1 *Assume that the bilinear form* $a(.,.)$ *satisfies conditions* (h_1)-(h_2) *and (12). Let* $\gamma \in W^{1,1}_{loc}(\mathbb{R}_+, \mathbb{R}_+)$ *be a function such that* $\lim_{t\to+\infty} \gamma(t) = 0$ *and* $\dot\gamma \in L^1(0, +\infty)$*. Let* u *be a solution in the class (4) to equation* (E_0)*. Then, either the trajectory is contained in KerA, or there exists* $K > 0$ *such that*

$$\mathcal{E}(t) \sim K\, e^{-\int_0^t \gamma(s)ds} \quad as \quad t \to +\infty. \tag{14}$$

Proof. For every $t \geq 0$, we set $\widehat{u}(t) = Pu(t)$ and $\widetilde{u}(t) = u(t) - Pu(t)$. Since $\widehat{u}(t) \in \text{Ker}A$, $\frac{d\widehat{u}}{dt}(t) \in \text{Ker}A$ and $\frac{d\widetilde{u}}{dt}(t) \in (\text{Ker}A)^\perp$, we have for every $t \geq 0$

$$\begin{aligned}\mathcal{E}(t) &= \frac{1}{2}\left|\frac{d\widehat{u}}{dt}(t) + \frac{d\widetilde{u}}{dt}(t)\right|^2 + \frac{1}{2}a(\widehat{u}(t) + \widetilde{u}(t), \widehat{u}(t) + \widetilde{u}(t)) \\ &= \frac{1}{2}\left|\frac{d\widehat{u}}{dt}(t)\right|^2 + \frac{1}{2}\left|\frac{d\widetilde{u}}{dt}(t)\right|^2 + \frac{1}{2}a(\widetilde{u}(t), \widetilde{u}(t)).\end{aligned} \tag{15}$$

From equality (6), we deduce that for every $t \geq 0$

$$\left|\frac{d\widehat{u}}{dt}(t)\right|^2 = e^{-2\int_0^t \gamma(s)ds} \left|\frac{d\widehat{u}}{dt}(0)\right|^2. \tag{16}$$

Let us now set $V_1 = (\text{Ker}A)^\perp \cap V$, $a_1 = a_{|V_1\times V_1}$ and $A_1 = A_{|V_1}$. It is clear that $\widetilde{u}$ is a solution of

$$\frac{d^2\widetilde{u}}{dt^2}(t) + \gamma(t)\frac{d\widetilde{u}}{dt}(t) + A_1\widetilde{u}(t) = 0.$$

On the other hand, condition (12) implies that $a_1(u, u) \geq \eta\, |u|^2$ for every $u \in V_1$. By applying Lemma 2.1 to the solution $\widetilde{u}$, we obtain that either the map $\widetilde{u}$ is stationary or there exists $K_1 > 0$ such that

$$\frac{1}{2}\left|\frac{d\widetilde{u}}{dt}(t)\right|^2 + \frac{1}{2}a(\widetilde{u}(t), \widetilde{u}(t)) \sim K_1\, e^{-\int_0^t \gamma(s)ds} \quad \text{as} \quad t \to +\infty. \tag{17}$$

We now combine equalities (15), (16) with estimate (17). If $\int_0^{+\infty} \gamma(s)ds = +\infty$, we immediately obtain (14) with $K = K_1$. If $\int_0^{+\infty} \gamma(s)ds < +\infty$, then

$$\lim_{t\to+\infty} \mathcal{E}(t) = \frac{1}{2}\, e^{-2\int_0^{+\infty} \gamma(s)ds} \left|\frac{d\widehat{u}}{dt}(0)\right|^2 + K_1\, e^{-\int_0^{+\infty} \gamma(s)ds},$$

hence (14) is satisfied with $K = \frac{1}{2}\, e^{-\int_0^{+\infty} \gamma(s)ds} \left|\frac{d\widehat{u}}{dt}(0)\right|^2 + K_1$.

Remark 2.5 *If the trajectory* $u(.)$ *is contained in KerA, estimate (14) is no more valid. In this case, we infer from equality (16) that* $\mathcal{E}(t) = \frac{1}{2}\, e^{-2\int_0^t \gamma(s)ds} \left|\frac{d\widehat{u}}{dt}(0)\right|^2$ *for every* $t \geq 0$*.*

Corollary 2.1 *Under the hypotheses of Theorem 2.1, assume moreover that* $\gamma \notin L^1(0, +\infty)$*. Then we have* $\lim_{t\to+\infty} \mathcal{E}(t) = 0$*. If* $KerA = \{0\}$*, then* $u(t) \to 0$ *strongly in* V *as* $t \to +\infty$*.*

Proof. The first assertion is an immediate consequence of estimate (14), while the second one follows from

$$\forall t \geq 0, \qquad \mathcal{E}(t) \geq \frac{1}{2}\, a(u(t), u(t)) \geq \frac{\eta'}{2}\, \|u(t)\|^2,$$

see inequality (13).

When Ker$A \neq \{0\}$, convergence of the trajectories is obtained under the following stronger assumption

$$\int_0^{+\infty} e^{-\frac{1}{2}\int_0^s \gamma(\tau)d\tau} ds < +\infty. \tag{18}$$

Corollary 2.2 *Under the hypotheses of Theorem 2.1, assume moreover that condition (18) is satisfied. Then, there exists $u_\infty \in KerA$ such that $u(t) \to u_\infty$ strongly in V as $t \to +\infty$.*

Proof. First assume that the trajectory is contained in Ker$A$. Observing that $\overline{\omega} = \int_0^{+\infty} e^{-\int_0^s \gamma(\tau)d\tau} ds < +\infty$, we deduce from Proposition 2.1 that $u(t)$ converges strongly in H as $t \to +\infty$. If the trajectory is not contained in Ker$A$, we derive from estimate (14) that

$$\left|\frac{du}{dt}(t)\right| = O\left(e^{-\frac{1}{2}\int_0^t \gamma(s)ds}\right) \quad \text{as} \quad t \to +\infty,$$

hence $\frac{du}{dt} \in L^1(\mathbb{R}_+, H)$ in view of condition (18). The trajectory u has a finite length, hence strongly converges in H toward some $u_\infty \in$ Ker$A$. Using now the semi-coercivity condition (h_2), we have

$$\begin{aligned}\mu\|u(t) - u_\infty\|^2 &\leq \lambda|u(t) - u_\infty|^2 + a(u(t) - u_\infty, u(t) - u_\infty)\\ &= \lambda|u(t) - u_\infty|^2 + a(u(t), u(t)).\end{aligned}$$

Since $\lim_{t\to+\infty} |u(t) - u_\infty| = 0$ and $\lim_{t\to+\infty} a(u(t), u(t)) = 0$ in view of Corollary 2.1, we conclude that $\lim_{t\to+\infty} \|u(t) - u_\infty\| = 0$.

Example 2.1 Suppose that there exist α, $k > 0$ such that $\gamma(t) = \frac{k}{t^\alpha}$ for t large enough. If the bilinear form $a(.,.)$ satisfies conditions (h_1)-(h_2) and (12), we deduce from Theorem 2.1 and Corollary 2.2 that

- if $\alpha > 1$, then $\lim\limits_{t\to+\infty} \mathcal{E}(t) > 0$;
- if $\alpha = 1$, then $\mathcal{E}(t) \sim \frac{K}{t^k}$ as $t \to +\infty$ and the trajectory $u(.)$ strongly converges in V as soon as $k > 2$;
- if $\alpha \in (0, 1)$, then $\mathcal{E}(t) \sim K e^{-\frac{k}{1-\alpha}t^{1-\alpha}}$ as $t \to +\infty$ and the trajectory $u(.)$ strongly converges in V for every $k > 0$.

Other results of convergence will be provided in the more general framework of semilinear equations.

3 Monotone conservative nonlinearity

The assumptions concerning the spaces V, H, the linear operator $A : V \to V'$ and the map $\gamma : \mathbb{R}_+ \to \mathbb{R}_+$ are the same as in section 2. We consider the following semilinear hyperbolic equation

$$(E) \qquad \frac{d^2u}{dt^2}(t) + \gamma(t)\frac{du}{dt}(t) + Au(t) + f(u(t)) = 0, \quad t \geq 0.$$

We suppose that the nonlinearity $f : V \to H$ is conservative, *i.e.*

$$(k_1) \qquad \exists F \in \mathcal{C}^1(V, \mathbb{R}) \text{ such that } \forall u, v \in V, \quad \langle F'(u), v\rangle_{V',V} = (f(u), v).$$

Moreover, we assume that the map f is monotone

$$(k_2) \qquad \forall u, v \in V, \quad (f(u) - f(v), u - v) \geq 0,$$

which is equivalent to the convexity of the potential F. Defining $\Phi : V \to \mathbb{R}$ by

$$\Phi(v) = \tfrac{1}{2}a(v, v) + F(v),$$

we obtain a function of class $\mathcal{C}^1$ whose first derivative is given by $\langle \Phi'(u), v\rangle_{V',V} = a(u, v) + (f(u), v)$, or equivalently $\Phi'(u) = Au + f(u)$. Moreover, Φ is convex, which amounts to

$$\forall u, v \in V, \ a(u, v - u) + (f(u), v - u) \leq \Phi(v) - \Phi(u). \tag{19}$$

Consequently, minimum and stationary points of Φ coincide, *i.e.*

$$\text{Argmin}\Phi = \{v \in V \mid Av + f(v) = 0\}, \tag{20}$$

where $\text{Argmin}\Phi = \{v \in V \mid \Phi(v) = \inf \Phi\}$. We suppose that

$$(k_3) \qquad S = \text{Argmin}\Phi \neq \emptyset.$$

It is clear in view of equation (E) that nothing is changed if some constant is added to the potential Φ. Without loss of generality, we will systematically assume that $\inf \Phi = 0$.

Remark 3.1 *Assume that a is coercive, i.e. (h_2) holds with $\lambda = 0$. Then the map $u \mapsto a(u, u)$ is strongly convex and since the function F is convex, the map Φ is also strongly convex. This implies immediately that the set ArgminΦ is a singleton, hence the non-vacuity condition (k_3) holds true. Now assume that (h_2) holds with $\lambda > 0$. To overcome the lack of coercivity, suppose that there exist $\varepsilon > 0$ and $C \geq 0$ such that $F(u) \geq \varepsilon\,|u|^2 - C$ for every $u \in V$. Without loss of generality, we can assume that $\varepsilon \leq \frac{\lambda}{2}$. For every $u \in V$, we have*

$$\begin{aligned}\Phi(u) = \frac{1}{2}a(u,u) + F(u) &\geq \frac{\varepsilon}{\lambda}a(u,u) + F(u)\\ &\geq \frac{\varepsilon\,\mu}{\lambda}\,\|u\|^2 - \varepsilon\,|u|^2 + \varepsilon\,|u|^2 - C\\ &= \frac{\varepsilon\,\mu}{\lambda}\,\|u\|^2 - C,\end{aligned}$$

which shows that $\lim_{\|u\|\to+\infty}\Phi(u) = +\infty$. *Since the function* Φ *is convex and continuous, this classically implies condition* (k_3).

It is immediate to check that the set S is convex, closed in V and that $S \subset D(A)$.

Remark 3.2 *Under assumption* (h_2), *let us show that* S *is closed in* H. *Let* (u_n) *be a sequence in* S *such that* $\lim_{n\to+\infty} u_n = \overline{u}$ *strongly in* H, *for some* $\overline{u} \in H$. *Since the function* F *is convex, there exist* $b, c \in \mathbb{R}$ *such that, for all* $u \in V$, $F(u) \geq -b|u| - c$. *Therefore we have for all* $u \in V$,

$$\frac{1}{2}a(u,u) \leq \Phi(u) + b|u| + c. \tag{21}$$

Recalling that $\Phi(u_n) = 0$ *for every* $n \in \mathbb{N}$, *we deduce that* $\frac{1}{2}a(u_n,u_n) \leq b|u_n| + c$, *hence the sequence* $(a(u_n,u_n))$ *is bounded. From hypothesis* (h_2), *we infer that the sequence* (u_n) *is bounded in* V. *It ensues that there exist* $\widehat{u} \in V$ *and a subsequence* (u_{n_k}) *such that* $\lim_{k\to+\infty} u_{n_k} = \widehat{u}$ *weakly in* V. *We immediately have* $\widehat{u} = \overline{u}$ *and the weak lower semicontinuity of* Φ *implies that* $\Phi(\overline{u}) \leq \liminf_{k\to+\infty}\Phi(u_{n_k}) = 0$, *hence* $\overline{u} \in S$.

Remark 3.3 (Case $f(0) = 0$) If $f(0) = 0$ then we have

$$S = \mathrm{Ker}A \cap \{v \in V \mid f(v) = 0\} \neq \emptyset.$$

Indeed, if $w \in S$ then in particular $(Aw,w) + (f(w),w) = 0$, and by monotonicity of f we have $(f(w) - f(0), w) \geq 0$, hence $(Aw,w) = (f(w),w) = 0$ and therefore $Aw = 0$.

In the sequel, we assume the existence of a solution to equation (E) in the class (4). We define the energy function $\mathcal{E}$ along each trajectory by

$$\mathcal{E}(t) = \frac{1}{2}\left|\frac{du}{dt}(t)\right|^2 + \Phi(u(t)).$$

We have $\mathcal{E} \in W^{1,1}_{loc}(\mathbb{R}_+)$ and

$$\begin{aligned}\dot{\mathcal{E}}(t) &= \left(\frac{d^2u}{dt^2}(t), \frac{du}{dt}(t)\right) + \left\langle Au(t) + f(u(t)), \frac{du}{dt}(t)\right\rangle_{V',V}\\ &= -\gamma(t)\left|\frac{du}{dt}(t)\right|^2 \leq 0 \qquad \text{a.e. on } \mathbb{R}_+,\end{aligned}$$

hence the function $\mathcal{E}$ is a Lyapunov function for the equation (E). We deduce that for every $t \geq 0$

$$\frac{1}{2}\left|\frac{du}{dt}(t)\right|^2 \leq \mathcal{E}(t) \leq \mathcal{E}(0) \quad \text{and} \quad \Phi(u(t)) \leq \mathcal{E}(t) \leq \mathcal{E}(0). \tag{22}$$

In particular, we have $\frac{du}{dt} \in L^\infty(\mathbb{R}_+, H)$. In the sequel, we will consider solutions which are bounded in H, *i.e.* satisfying $u \in L^\infty(\mathbb{R}_+, H)$.

Remark 3.4 *Under assumption (h_2), it is easy to see that $u \in L^\infty(\mathbb{R}_+, H)$ implies $u \in L^\infty(\mathbb{R}_+, V)$. Indeed, let us assume that $\{u(t); t \geq 0\}$ is bounded in H. From inequality (21), we have $\frac{1}{2}a(u(t), u(t)) \leq \Phi(u(t)) + b|u(t)| + c$ for all $t \in \mathbb{R}_+$. Recalling that $\Phi(u(t)) \leq \mathcal{E}(0)$ in view of (22), we infer that $\{a(u(t), u(t)); t \geq 0\}$ is bounded. From hypothesis (h_2), we conclude that $\{u(t); t \geq 0\}$ is bounded in V.*

3.1 Summability of the energy. Case of a unique equilibrium

We now prove that the map $\gamma\,\mathcal{E}$ is summable over $\mathbb{R}_+$ and that $\lim_{t\to+\infty} \mathcal{E}(t) = 0$.

Proposition 3.1 *Assume that the bilinear form $a(.,.)$ and the function f satisfy respectively hypotheses (h_1)-(h_2) and (k_1)-(k_3). Let $\gamma \in W^{1,1}_{loc}(\mathbb{R}_+, \mathbb{R}_+)$ be a map such that $\dot\gamma \in L^1(0, +\infty)$. Let u be a solution in the class (4) to equation (E) and assume that $u \in L^\infty(\mathbb{R}_+, H)$. Then*

(i) $\int_0^{+\infty} \gamma(t)\,\mathcal{E}(t)\,dt < +\infty$.
(ii)If moreover $\gamma \notin L^1(0, +\infty)$, then $\lim_{t\to+\infty} \mathcal{E}(t) = 0$, hence

$$\lim_{t\to+\infty}\left|\frac{du}{dt}(t)\right| = 0 \quad \textit{and} \quad \lim_{t\to+\infty}\Phi(u(t)) = 0. \tag{23}$$

Proof. (i) The proof follows the same arguments as those of [8, Prop. 3.1]. Let us take $v \in S$ and define the function $p : \mathbb{R}_+ \to \mathbb{R}_+$ by $p(t) = \frac{1}{2}\,|u(t) - v|^2$. By differentiating, we find for every $t \geq 0$

$$\dot p(t) = \left(\frac{du}{dt}(t), u(t) - v\right).$$

Since $\frac{du}{dt} \in W^{1,1}_{loc}(\mathbb{R}_+, H)$ by assumption, it is immediate to check that $\dot p \in W^{1,1}_{loc}(\mathbb{R}_+)$. Hence the map $\dot p$ is differentiable almost everywhere on $\mathbb{R}_+$ and we have

$$\ddot p(t) = \left(\frac{d^2u}{dt^2}(t), u(t) - v\right) + \left|\frac{du}{dt}(t)\right|^2 \quad \text{a.e. on } \mathbb{R}_+.$$

By combining the expressions of $\dot p$, $\ddot p$ and by using the convexity of the function Φ, we obtain

$$\ddot{p}(t)+\gamma(t)\dot{p}(t) = a(u(t), v-u(t)) + (f(u(t)), v-u(t)) + \left|\frac{du}{dt}(t)\right|^2$$
$$\leq -\Phi(u(t)) + \left|\frac{du}{dt}(t)\right|^2 \qquad \text{a.e. on } \mathbb{R}_+. \tag{24}$$

It follows that

$$\ddot{p}(t)+\gamma(t)\dot{p}(t)+\mathcal{E}(t) \leq \frac{3}{2}\left|\frac{du}{dt}(t)\right|^2 \qquad \text{a.e. on } \mathbb{R}_+. \tag{25}$$

Let us multiply this inequality by $\gamma(t)$ and integrate on $[0,t]$. By using the fact that $\dot{\mathcal{E}}(t) = -\gamma(t)\left|\frac{du}{dt}(t)\right|^2$ almost everywhere on $\mathbb{R}_+$, we derive that

$$\int_0^t \gamma(s)\,\mathcal{E}(s)\,ds \leq \frac{3}{2}\left(\mathcal{E}(0)-\mathcal{E}(t)\right) - \int_0^t \gamma(s)\ddot{p}(s)ds - \int_0^t \gamma(s)^2\dot{p}(s)ds. \tag{26}$$

For the last two integrals, let us use a technique of integration by parts.

$$-\int_0^t \gamma(s)\ddot{p}(s)ds = \gamma(0)\dot{p}(0) - \gamma(t)\dot{p}(t) + \int_0^t \dot{\gamma}(s)\dot{p}(s)ds. \tag{27}$$

Recall that the map u is bounded in H by assumption. On the other hand, the map $\frac{du}{dt}$ is bounded in H, see (22). Hence we infer the existence of $M>0$ such that $p(t)\leq M$ and $|\dot{p}(t)|\leq M$ for every $t\geq 0$. Therefore

$$-\int_0^t \gamma(s)\ddot{p}(s)ds \leq M\gamma(0) + M\gamma(t) + M\int_0^t |\dot{\gamma}(s)|\,ds.$$

Since $\dot{\gamma}\in L^1(0,+\infty)$ by assumption, the right-hand side is majorized by some $M'\geq 0$. On the other hand, we have

$$-\int_0^t \gamma(s)^2\dot{p}(s)\,ds = \gamma(0)^2\,p(0) - \gamma(t)^2\,p(t) + 2\int_0^t \gamma(s)\dot{\gamma}(s)\,p(s)\,ds \tag{28}$$
$$\leq M\,\gamma(0)^2 + 2\,M\int_0^t \gamma(s)|\dot{\gamma}(s)|\,ds.$$

Using again the assumption $\dot{\gamma}\in L^1(0,+\infty)$, we obtain that the right-hand side is majorized by some $M''\geq 0$. Coming back to inequality (26), we conclude that $\int_0^t \gamma(s)\,\mathcal{E}(s)\,ds \leq \frac{3}{2}\mathcal{E}(0) + M' + M''$ for every $t\geq 0$ and the expected estimate follows.

(ii) Let us argue by contradiction and assume that $\lim_{t\to+\infty}\mathcal{E}(t) = l > 0$. The map $\mathcal{E}$ is nonincreasing, hence $\mathcal{E}(t)\geq l$ for every $t\geq 0$. Since $\gamma\notin L^1(0,+\infty)$, we deduce that

$$\int_0^{+\infty} \gamma(t)\,\mathcal{E}(t)\,dt \geq l\int_0^{+\infty}\gamma(t)\,dt = +\infty,$$

a contradiction with the result of (i). The last assertion is immediate.

In view of the previous result, we can prove weak convergence of the trajectories in the case of a unique equilibrium. The general case of multiple equilibria is more delicate and will be discussed in section 3.2.

Corollary 3.1 (Case of a unique equilibrium) *Under the hypotheses of Proposition 3.1, assume moreover that* $\mathrm{Argmin}\Phi = \{\overline{u}\}$ *for some* $\overline{u} \in V$*. Then the solution* $u(t)$ *weakly converges in* V *toward* $\overline{u}$ *as* $t \to +\infty$*. Furthermore, if* $u(t)$ *strongly converges*[8] *in* H *then it strongly converges in* V*.*

Proof. By assumption, the solution u is bounded in H. In view of hypothesis (h_2) and Remark 3.4, it is also bounded in V. Hence there exist $u_\infty \in V$ and a subsequence (t_n) tending to $+\infty$ such that $\lim_{n\to+\infty} u(t_n) = u_\infty$ weakly in V. Since Φ is convex and continuous for the strong topology of V, it is lower semicontinuous for the weak topology of V. Hence, we have $\Phi(u_\infty) \leq \liminf_{n\to+\infty} \Phi(u(t_n))$. From the second part of (23) we deduce that $\Phi(u_\infty) \leq 0$, *i.e.* $u_\infty \in \mathrm{Argmin}\Phi = \{\overline{u}\}$. Hence $\overline{u}$ is the unique limit point of the map $t \mapsto u(t)$ as $t \to +\infty$ for the weak topology of V. It ensues that $\lim_{t\to+\infty} u(t) = \overline{u}$ weakly in V. Let us now prove the second point. The argument is given in [3, p. 548-549] but we recall it for the sake of completeness. From (h_2), we have

$$\begin{aligned}\mu\,\|u(t)-\overline{u}\|^2 &\leq \lambda\,|u(t)-\overline{u}|^2 + a(u(t)-\overline{u}, u(t)-\overline{u}) \\ &= \lambda\,|u(t)-\overline{u}|^2 + 2\,\Phi(u(t)) - 2\,F(u(t)) - 2\,a(u(t),\overline{u}) + a(\overline{u},\overline{u}).\end{aligned} \tag{29}$$

Since $u(t) \to \overline{u}$ strongly in H and weakly in V, we have $\lim_{t\to+\infty} |u(t)-\overline{u}|^2 = 0$ and $\lim_{t\to+\infty} a(u(t),\overline{u}) = a(\overline{u},\overline{u})$. On the other hand, by weak lower semicontinuity of the continuous convex function $F : V \to \mathbb{R}$, we infer that $\liminf_{t\to+\infty} F(u(t)) \geq F(\overline{u})$. Recalling finally property (23), we deduce from inequality (29) that

$$\mu \limsup_{t\to+\infty} \|u(t)-\overline{u}\|^2 \leq -2\,F(\overline{u}) - a(\overline{u},\overline{u}) = 0.$$

We conclude that $u(t) \to \overline{u}$ strongly in V.

3.2 Convergence of the trajectories

Case of a non vanishing damping

When the damping coefficient $\gamma(t)$ is constant, *i.e.* $\gamma(t) \equiv \gamma > 0$, the solutions of (E) weakly converge in V toward an equilibrium point, see [3]. We are going to show that this property still holds true if

$$(l_1) \qquad \begin{cases} \lim_{t\to+\infty} \gamma(t) = \gamma_\infty > 0 \\ \dot{\gamma} \in L^1(0,+\infty). \end{cases}$$

[8] This assumption is satisfied if the injection $V \hookrightarrow H$ is compact.

Theorem 3.1 *Assume that the bilinear form $a(.,.)$ and the function f satisfy respectively (h_1)-(h_2) and (k_1)-(k_3). Let $\gamma \in W^{1,1}_{loc}(\mathbb{R}_+, \mathbb{R}_+)$ be a map satisfying (l_1). Let u be a solution in the class (4) to equation (E). Then, there exists $u_\infty \in S$ such that $u(t) \rightharpoonup u_\infty$ weakly in V as $t \to +\infty$. Furthermore, if $u(t)$ strongly converges in H then it strongly converges in V.*

Proof. Let $v \in S$ and define the map $p : \mathbb{R}_+ \to \mathbb{R}_+$ by $p(t) = \frac{1}{2}|u(t) - v|^2$ as in the proof of Proposition 3.1. Inequality (24) implies that

$$\ddot{p}(t) + \gamma(t)\dot{p}(t) \leq \left|\frac{du}{dt}(t)\right|^2 \qquad \text{a.e. on } \mathbb{R}_+.$$

Let us multiply each member of this inequality by $e^{\int_0^t \gamma(\tau)\,d\tau}$ and integrate on $[0,t]$. Recalling that $\dot{p} \in W^{1,1}_{loc}(\mathbb{R}_+)$, we obtain

$$\dot{p}(t) \leq e^{-\int_0^t \gamma(\tau)\,d\tau}\dot{p}(0) + e^{-\int_0^t \gamma(\tau)\,d\tau} \int_0^t e^{\int_0^s \gamma(\tau)\,d\tau} \left|\frac{du}{ds}(s)\right|^2 ds. \tag{30}$$

We now show that the right member of the above inequality is a summable function. Since $\lim_{t\to+\infty} \gamma(t) = \gamma_\infty > 0$, there exists $t_0 > 0$ such that $\gamma(t) \geq \gamma_\infty/2$ for every $t \geq t_0$. From Lemma 3.1 (i) below, we have

$$\int_0^{+\infty} e^{-\int_0^t \gamma(\tau)\,d\tau}\, dt < +\infty. \tag{31}$$

Lemma 3.1 *Let us assume that there exist $k > 0$ and $t_0 > 0$ such that $\gamma(t) \geq k$ for every $t \geq t_0$. Then we have*

(i) $\displaystyle\int_0^{+\infty} e^{-\int_0^t \gamma(\tau)\,d\tau}\, dt < +\infty;$

(ii) $\displaystyle\int_s^{+\infty} e^{-\int_0^t \gamma(\tau)\,d\tau}\, dt \leq \frac{1}{k}\, e^{-\int_0^s \gamma(\tau)\,d\tau}$ *for s large enough.*

Lemma 3.1 is a particular case of a more general result that will be proved next, see Lemma 3.3. Coming back to inequality (30), we find by applying Fubini theorem

$$\int_0^{+\infty} e^{-\int_0^t \gamma(\tau)\,d\tau} \int_0^t e^{\int_0^s \gamma(\tau)\,d\tau} \left|\frac{du}{ds}(s)\right|^2 ds\,dt = \int_0^{+\infty} \left|\frac{du}{ds}(s)\right|^2 e^{\int_0^s \gamma(\tau)\,d\tau} \int_s^{+\infty} e^{-\int_0^t \gamma(\tau)\,d\tau}\, dt\,ds. \tag{32}$$

From Lemma 3.1 (ii), we obtain

$$e^{\int_0^s \gamma(\tau)\,d\tau} \int_s^{+\infty} e^{-\int_0^t \gamma(\tau)\,d\tau}\, dt \leq \frac{2}{\gamma_\infty} \leq \frac{4}{\gamma_\infty^2}\,\gamma(s).$$

Recalling that $\dot{\mathcal{E}}(t) = -\gamma(t)\,\left|\frac{du}{dt}(t)\right|^2$, we have the estimate $\int_0^{+\infty} \gamma(s)\,\left|\frac{du}{ds}(s)\right|^2 ds < +\infty$. Hence we deduce from equality (32) that

$$\int_0^{+\infty} e^{-\int_0^t \gamma(\tau)\,d\tau} \int_0^t e^{\int_0^s \gamma(\tau)\,d\tau} \left|\frac{du}{ds}(s)\right|^2 ds\,dt < +\infty. \tag{33}$$

By combining inequality (30) with estimates (31) and (33), we infer that $[\dot{p}]_+ \in L^1(0,+\infty)$ and hence $\lim_{t\to+\infty} p(t)$ exists. In particular the map u is bounded in H. The end of the proof is the same as in [3, Theorem 3.1] but the arguments are given for the sake of completeness. Since $u \in L^\infty(\mathbb{R}_+, H)$, we deduce from hypothesis (h_2) and Remark 3.4 that $u \in L^\infty(\mathbb{R}_+, V)$. Let $\overline{u} \in V$ be a weak cluster point of $\{u(t); t \to +\infty\}$ for the weak topology of V. There exists a sequence $t_n \to +\infty$ such that $u(t_n) \rightharpoonup \overline{u}$ weakly in V as $n \to +\infty$. Since the function Φ is lower semicontinuous for the weak topology of V, we have[9] in view of Proposition 3.1

$$\Phi(\overline{u}) \le \liminf_{n\to+\infty} \Phi(u(t_n)) = \lim_{t\to+\infty} \Phi(u(t)) = 0,$$

which implies that $\overline{u} \in S$. Let us prove that $\{u(t); t \to +\infty\}$ has a unique cluster point for the weak topology in V. We apply the following argument due to Opial [18]. Let $\overline{u}_1, \overline{u}_2 \in S$ be two cluster points of $\{u(t); t \to +\infty\}$ for the weak topology of V. According to the first part of the proof, we can assert that $\lim_{t\to+\infty} |u(t) - \overline{u}_i|^2 = l_i$ exists for each $i = 1, 2$. Moreover there exists a sequence $t_n \to +\infty$ such that $u(t_n) \rightharpoonup \overline{u}_1$ weakly in V as $n \to +\infty$. Since the injection $V \hookrightarrow H$ is continuous, $u(t_n) \rightharpoonup \overline{u}_1$ weakly in H as $n \to +\infty$. From the equality

$$|u(t) - \overline{u}_1|^2 - |u(t) - \overline{u}_2|^2 = |\overline{u}_1 - \overline{u}_2|^2 + 2(\overline{u}_1 - \overline{u}_2, \overline{u}_2 - u(t)),$$

we infer that $l_1 - l_2 = -|\overline{u}_1 - \overline{u}_2|^2$. On the other hand, if we take $t_m \to +\infty$ such that $u(t_m) \rightharpoonup \overline{u}_2$ weakly in V as $m \to +\infty$, we find $l_1 - l_2 = |\overline{u}_1 - \overline{u}_2|^2$. As a consequence, $|\overline{u}_1 - \overline{u}_2|^2 = 0$. This establishes the uniqueness of the cluster points of $\{u(t); t \to +\infty\}$ for the weak topology of V. Hence $u(t) \rightharpoonup u_\infty$ weakly in V as $t \to +\infty$ for some $u_\infty \in V$.

For the second point, the reader is referred to the corresponding argument in the proof of Corollary 3.1.

An interesting situation ensuring strong convergence in V is the case where the non-linearity satisfies the symmetry property $F(-u) = F(u)$ for all $u \in V$.

Theorem 3.2 *Under the hypotheses of Theorem 3.1, assume moreover that the function F is even, i.e. $F(-u) = F(u)$ for all $u \in V$. Then there exists $u_\infty \in S$ such that $u(t) \to u_\infty$ strongly in V.*

Proof. The argument was originated by Bruck, see [6, Theorem 5]. It has been adapted to the framework of second-order in time equations, see for example [2, Theorem 2.4 (i)] or [3, Remark 3.2] in the case of a constant damping parameter γ. Let us fix $t_0 > 0$ and define the map $q : [0, t_0] \to \mathbb{R}$ by

[9] Observe that Proposition 3.1 applies rightfully since we have proved that $u \in L^\infty(\mathbb{R}_+, H)$.

$$q(t) = |u(t)|^2 - |u(t_0)|^2 - \frac{1}{2}|u(t) - u(t_0)|^2.$$

A first differentiation gives for all $t \in [0, t_0]$

$$\dot{q}(t) = \left(\frac{du}{dt}(t), u(t) + u(t_0)\right).$$

Since $\frac{du}{dt} \in W^{1,1}_{loc}(\mathbb{R}_+, H)$ by assumption, it is immediate to check that the map $\dot{q}$ is absolutely continuous, hence differentiable almost everywhere on $[0, t_0]$ and we have

$$\ddot{q}(t) = \left(\frac{d^2u}{dt^2}(t), u(t) + u(t_0)\right) + \left|\frac{du}{dt}(t)\right|^2 \qquad \text{a.e. on } [0, t_0].$$

By combining the expressions of $\dot{q}$, $\ddot{q}$, we obtain for almost every $t \in [0, t_0]$

$$\begin{aligned}\ddot{q}(t) + \gamma(t)\dot{q}(t) &= -a(u(t), u(t) + u(t_0)) - (f(u(t)), u(t) + u(t_0)) + \left|\frac{du}{dt}(t)\right|^2 \\ &= -\langle \Phi'(u(t)), u(t) + u(t_0)\rangle_{V',V} + \left|\frac{du}{dt}(t)\right|^2. \qquad (34)\end{aligned}$$

Since the function Φ is convex and even, we have for all $u, v \in V$

$$\Phi(v) - \Phi(u) = \Phi(-v) - \Phi(u) \geq -\langle \Phi'(u), v + u\rangle_{V',V}.$$

Hence inequality (34) gives

$$\ddot{q}(t) + \gamma(t)\dot{q}(t) \leq \Phi(u(t_0)) - \Phi(u(t)) + \left|\frac{du}{dt}(t)\right|^2 \qquad \text{a.e. on } [0, t_0]. \qquad (35)$$

Recalling that the energy function $\mathcal{E}(t)$ is nonincreasing, we have $\frac{1}{2}\left|\frac{du}{dt}(t)\right|^2 + \Phi(u(t)) \geq \left|\frac{du}{dt}(t_0)\right|^2 + \Phi(u(t_0))$ for every $t \in [0, t_0]$. Therefore

$$\forall t \in [0, t_0], \quad \Phi(u(t_0)) - \Phi(u(t)) \leq \frac{1}{2}\left|\frac{du}{dt}(t)\right|^2.$$

Using inequality (35), we deduce that

$$\ddot{q}(t) + \gamma(t)\dot{q}(t) \leq \frac{3}{2}\left|\frac{du}{dt}(t)\right|^2 \qquad \text{a.e. on } [0, t_0].$$

Let us multiply each member of this inequality by $e^{\int_0^t \gamma(\tau)\, d\tau}$ and integrate on $[0, t]$. Since the map $\dot{q}$ is absolutely continuous, we find

$$\dot{q}(t) \leq e^{-\int_0^t \gamma(\tau)\, d\tau}\dot{q}(0) + \frac{3}{2}e^{-\int_0^t \gamma(\tau)\, d\tau}\int_0^t e^{\int_0^s \gamma(\tau)\, d\tau}\left|\frac{du}{ds}(s)\right|^2 ds.$$

Let us integrate this inequality on $[t, t_0]$, we obtain

$$-q(t) \leq \dot{q}(0) \int_t^{t_0} e^{-\int_0^s \gamma(\tau)\, d\tau}\, ds + \frac{3}{2}(h(t_0) - h(t)),$$

where we have set

$$h(t) = \int_0^t e^{-\int_0^s \gamma(\tau)\, d\tau} \int_0^s e^{\int_0^\sigma \gamma(\tau)\, d\tau} \left|\frac{du}{dt}(\sigma)\right|^2 d\sigma\, ds.$$

We deduce from the previous inequality that

$$\frac{1}{2}|u(t)-u(t_0)|^2 \leq |u(t)|^2 - |u(t_0)|^2 + \dot{q}(0) \int_t^{t_0} e^{-\int_0^s \gamma(\tau)\, d\tau}\, ds + \frac{3}{2}(h(t_0)-h(t)). \quad (36)$$

In the proof of Theorem 3.1, we showed that $\lim_{t\to+\infty} |u(t) - v|^2$ exists for all $v \in \operatorname{Argmin}\Phi$. Since Φ is convex and even, we have $0 \in \operatorname{Argmin}\Phi$, hence $\lim_{t\to+\infty} |u(t)|^2$ exists. On the other hand, the integral $\int_0^{+\infty} e^{-\int_0^s \gamma(\tau)\, d\tau}\, ds$ is finite from (31), while $\lim_{t\to+\infty} h(t)$ exists in view of estimate (33). We then deduce from inequality (36) that $\{u(t); t \to +\infty\}$ is a Cauchy net in H hence strongly converges in H. It suffices to use the second part of Theorem 3.1 to obtain the strong convergence in V.

Case of a vanishing damping

It is assumed in this paragraph that the damping parameter $\gamma(t)$ vanishes as $t \to +\infty$. The trajectories of (E) are clearly more volatile in this framework. Our purpose is to obtain results of convergence for the trajectories, assuming that $\gamma(t)$ tends slowly enough toward 0. We are going to show that the convergence properties stated in the previous paragraph still hold true if the quantity $\gamma(t)$ behaves as k/t^α, for some $\alpha \in]0, 1[$, $k > 0$ and t large enough. The main step consists in establishing a refinement of Proposition 3.1 via sharp estimates for the energy decay. Let us start with a technical lemma that will be crucial in the sequel.

Lemma 3.2 *Assume that the bilinear form $a(.,.)$ and the function f satisfy respectively hypotheses (h_1)-(h_2) and (k_1)-(k_3). Let $\gamma \in W^{1,1}_{loc}(\mathbb{R}_+, \mathbb{R}_+)$ be a function such that $\lim_{t\to+\infty} \gamma(t) = 0$. Let u be a solution in the class (4) to equation (E) and assume that $u \in L^\infty(\mathbb{R}_+, H)$. We are given some $t_0 \geq 0$ along with a non constant map $\lambda \in \mathcal{C}^3([t_0, +\infty[, \mathbb{R})$ such that $\lambda(t) \geq 0$, $\dot{\lambda}(t) \geq 0$, $\ddot{\lambda}(t) \geq 0$ and $\dddot{\lambda}(t) \leq 0$ for every $t \geq t_0$. Assume that the map $t \mapsto \dot{\lambda}(t) \left|\frac{du}{dt}(t)\right|$ is bounded, that $\int_{t_0}^{+\infty} \dot{\lambda}(t)\, |\dot{\gamma}(t)|\, dt < +\infty$ and that $\lambda(t)\, \gamma(t) \geq 2\, \dot{\lambda}(t)$ for every $t \geq t_0$. Then the following estimates hold true*

(i) $\int_{t_0}^{+\infty} \dot{\lambda}(t)\, \mathcal{E}(t)\, dt < +\infty$.

(ii) $\lim_{t\to+\infty} \lambda(t)\,\mathcal{E}(t) = 0$.
(iii) $\int_{t_0}^{+\infty} \lambda(t)\,\gamma(t)\,\left|\frac{du}{dt}(t)\right|^2 dt < +\infty$.

Proof. Let us consider the map p defined by $p(t) = \frac{1}{2}|u(t) - v|^2$ for some $v \in S$, see the proof of Proposition 3.1. Recall that we have from inequality (25)

$$\mathcal{E}(t) \le \frac{3}{2}\left|\frac{du}{dt}(t)\right|^2 - \ddot{p}(t) - \gamma(t)\,\dot{p}(t) \qquad \text{a.e. on } \mathbb{R}_+. \tag{37}$$

Now define the map $\mathcal{E}_\lambda : [t_0, +\infty[\to \mathbb{R}_+$ by $\mathcal{E}_\lambda(t) = \lambda(t)\,\mathcal{E}(t)$. It is clear that $\mathcal{E}_\lambda \in W^{1,1}_{loc}([t_0, +\infty[)$. Since $\dot{\mathcal{E}}(t) = -\gamma(t)\left|\frac{du}{dt}(t)\right|^2$ for almost every $t \ge 0$, we have

$$\dot{\mathcal{E}}_\lambda(t) = \dot{\lambda}(t)\,\mathcal{E}(t) - \lambda(t)\,\gamma(t)\left|\frac{du}{dt}(t)\right|^2 \qquad \text{a.e. on } [t_0, +\infty[. \tag{38}$$

From the assumption $\lambda(t)\,\gamma(t) \ge 2\,\dot{\lambda}(t)$ for every $t \ge t_0$, we deduce that

$$\dot{\lambda}(t)\left|\frac{du}{dt}(t)\right|^2 \le \frac{1}{2}\dot{\lambda}(t)\,\mathcal{E}(t) - \frac{1}{2}\dot{\mathcal{E}}_\lambda(t) \qquad \text{a.e. on } [t_0, +\infty[. \tag{39}$$

By combining inequalities (37) and (39), we infer that

$$\dot{\lambda}(t)\,\mathcal{E}(t) \le -3\,\dot{\mathcal{E}}_\lambda(t) - 4\,\dot{\lambda}(t)\,[\ddot{p}(t) + \gamma(t)\,\dot{p}(t)] \qquad \text{a.e. on } [t_0, +\infty[.$$

Let us integrate this inequality on $[t_0, t]$; we find

$$\int_{t_0}^{t} \dot{\lambda}(s)\,\mathcal{E}(s)ds \le 3\,\mathcal{E}_\lambda(t_0) - 4\int_{t_0}^{t} \dot{\lambda}(s)\,\ddot{p}(s)\,ds - 4\int_{t_0}^{t} \dot{\lambda}(s)\,\gamma(s)\,\dot{p}(s)\,ds. \tag{40}$$

For the last two integrals, let us use a technique of integration by parts.

$$\begin{aligned}
-\int_{t_0}^{t} \dot{\lambda}(s)\,\ddot{p}(s)\,ds &= -\dot{\lambda}(t)\,\dot{p}(t) + \dot{\lambda}(t_0)\,\dot{p}(t_0) + \int_{t_0}^{t} \ddot{\lambda}(s)\,\dot{p}(s)\,ds \\
&= -\dot{\lambda}(t)\,\dot{p}(t) + \dot{\lambda}(t_0)\,\dot{p}(t_0) + \ddot{\lambda}(t)\,p(t) - \ddot{\lambda}(t_0)\,p(t_0) \\
&\quad - \int_{t_0}^{t} \dddot{\lambda}(s)\,p(s)\,ds.
\end{aligned}$$

The map u is bounded in H by assumption, hence there exist M, $M' > 0$ such that $p(t) \le M$ and $|\dot{p}(t)| \le M'\left|\frac{du}{dt}(t)\right|$ for every $t \ge 0$. Therefore we deduce from the above equality that

$$-\int_{t_0}^{t} \dot{\lambda}(s)\,\ddot{p}(s)\,ds \le M'\,\dot{\lambda}(t)\left|\frac{du}{dt}(t)\right| + M'\,\dot{\lambda}(t_0)\left|\frac{du}{dt}(t_0)\right| + M\,\ddot{\lambda}(t) + M\int_{t_0}^{t} |\dddot{\lambda}(s)|\,ds.$$

Recalling that $\dddot{\lambda}(t) \le 0$ and that the map $t \mapsto \dot{\lambda}(t)\left|\frac{du}{dt}(t)\right|$ is bounded by some $M'' > 0$, we obtain

$$-\int_{t_0}^{t} \dot\lambda(s)\,\ddot p(s)\,ds \le 2\,M'\,M'' + M\,\ddot\lambda(t) + M\,(\ddot\lambda(t_0) - \ddot\lambda(t)) = 2\,M'\,M'' + M\,\ddot\lambda(t_0). \tag{41}$$

On the other hand, we have

$$\begin{aligned}
-\int_{t_0}^{t} \dot\lambda(s)\,\gamma(s)\,\dot p(s)\,ds &= -\dot\lambda(t)\,\gamma(t)\,p(t) + \dot\lambda(t_0)\,\gamma(t_0)\,p(t_0) + \int_{t_0}^{t} \ddot\lambda(s)\,\gamma(s)\,p(s)\,ds \\
&\quad + \int_{t_0}^{t} \dot\lambda(s)\,\dot\gamma(s)\,p(s)\,ds \\
&\le M\,\dot\lambda(t_0)\,\gamma(t_0) + M\int_{t_0}^{t} \ddot\lambda(s)\,\gamma(s)\,ds \qquad (42)\\
&\quad + M\int_{t_0}^{t} \dot\lambda(s)\,|\dot\gamma(s)|\,ds.
\end{aligned}$$

Observe that

$$\begin{aligned}
\int_{t_0}^{t} \ddot\lambda(s)\,\gamma(s)\,ds &= \dot\lambda(t)\gamma(t) - \dot\lambda(t_0)\gamma(t_0) - \int_{t_0}^{t} \dot\lambda(s)\,\dot\gamma(s)\,ds \\
&\le \dot\lambda(t)\gamma(t) + \int_{t_0}^{t} \dot\lambda(s)\,|\dot\gamma(s)|\,ds. \qquad (43)
\end{aligned}$$

Since $\lim_{t\to+\infty} \gamma(t) = 0$, we have for every $t \ge t_0$

$$\dot\lambda(t)\gamma(t) = \dot\lambda(t)\int_{t}^{+\infty} -\dot\gamma(s)\,ds \le \dot\lambda(t)\int_{t}^{+\infty} |\dot\gamma(s)|\,ds \le \int_{t}^{+\infty} \dot\lambda(s)\,|\dot\gamma(s)|\,ds,$$

the last equality being a consequence of the fact that the map $\dot\lambda$ is non decreasing. The finiteness of the integral $\int_t^{+\infty} \dot\lambda(s)\,|\dot\gamma(s)|\,ds$ is ensured by assumption. In view of (43), we deduce that

$$\int_{t_0}^{t} \ddot\lambda(s)\,\gamma(s)\,ds \le \int_{t_0}^{+\infty} \dot\lambda(s)\,|\dot\gamma(s)|\,ds < +\infty.$$

Coming back to (42), we infer that

$$-\int_{t_0}^{t} \dot\lambda(s)\,\gamma(s)\,\dot p(s)\,ds \le M\,\dot\lambda(t_0)\,\gamma(t_0) + 2\,M\int_{t_0}^{+\infty} \dot\lambda(s)\,|\dot\gamma(s)|\,ds < +\infty. \tag{44}$$

By combining inequalities (40), (41) and (44), we conclude that the quantity $\int_{t_0}^{t} \dot\lambda(s)\,\mathcal{E}(s)ds$ is uniformly majorized with respect to t, whence (i).

Let us now come back to equation (38). By taking the positive part of each member, we find $(\dot{\mathcal{E}}_\lambda)_+(t) \le \dot\lambda(t)\mathcal{E}(t)$. This implies that $(\dot{\mathcal{E}}_\lambda)_+ \in L^1(0,+\infty)$ and therefore $l = \lim_{t\to+\infty} \lambda(t)\,\mathcal{E}(t)$ exists in $\mathbb{R}_+$. We have to prove that $l = 0$. Let us argue by contradiction and assume that $l > 0$. Then $\mathcal{E}(t) \sim$

$l/\lambda(t)$ for t large enough. From (i), we deduce that $\int_{t_0}^{+\infty} \dot{\lambda}(s)/\lambda(s)\, ds < +\infty$, *i.e.* $\lim_{t\to+\infty} \ln \lambda(t) < +\infty$. Hence the nondecreasing convex map λ has a finite limit as $t \to +\infty$, which implies that it is constant. But it contradicts the assumption and we conclude that $l = 0$, which shows (ii).

By integrating equality (38) on $[t_0, t]$, we obtain

$$\begin{aligned}\int_{t_0}^{t} \lambda(s)\, \gamma(s) \left|\frac{du}{ds}(s)\right|^2 ds &= \int_{t_0}^{t} \dot{\lambda}(s)\, \mathcal{E}(s)\, ds + \mathcal{E}_\lambda(t_0) - \mathcal{E}_\lambda(t)\\ &\le \int_{t_0}^{+\infty} \dot{\lambda}(s)\, \mathcal{E}(s)\, ds + \mathcal{E}_\lambda(t_0) < +\infty.\end{aligned}$$

Letting $t \to +\infty$, we immediately obtain (iii).

A repeated application of Lemma 3.2 allows to derive sharp estimates for the energy decay under some suitable conditions. These estimates will be the keystone for proving convergence of the trajectories.

Proposition 3.2 *Assume that the bilinear form $a(.,.)$ and the function f satisfy respectively hypotheses (h_1)-(h_2) and (k_1)-(k_3). Let $\gamma \in W^{1,1}_{loc}(\mathbb{R}_+, \mathbb{R}_+)$ be a function such that $\lim_{t\to+\infty} \gamma(t) = 0$. Assume that $\displaystyle\int_0^{+\infty} t^{1-\left(\frac{1}{2}\right)^n} |\dot{\gamma}(t)|\, dt < +\infty$ for some $n \in \mathbb{N}$ and that there exists $t_0 > 0$ such that $\gamma(t) \ge \frac{4}{t}$ for every $t \ge t_0$. Let u be a solution in the class (4) to equation (E) and assume that $u \in L^\infty(\mathbb{R}_+, H)$. Then we have*

(i) $\displaystyle\int_0^{+\infty} t^{1-\left(\frac{1}{2}\right)^n}\, \mathcal{E}(t)\, dt < +\infty.$

(ii) $\displaystyle\lim_{t\to+\infty} t^{2-\left(\frac{1}{2}\right)^n}\, \mathcal{E}(t) = 0.$

(iii) $\displaystyle\int_0^{+\infty} t^{2-\left(\frac{1}{2}\right)^n}\, \gamma(t) \left|\frac{du}{dt}(t)\right|^2 dt < +\infty.$

Proof. First we use Lemma 3.2 with the map λ_0 defined by $\lambda_0(t) = t$ for every $t \ge 0$. Let us verify that the assumptions of Lemma 3.2 are satisfied. Recall that the map $t \mapsto \left|\frac{du}{dt}(t)\right|$ is bounded, see (22). On the other hand, the finiteness of the integral $\int_0^{+\infty} |\dot{\gamma}(t)|\, dt$ is a consequence of the assumption $\int_0^{+\infty} t^{1-\left(\frac{1}{2}\right)^n} |\dot{\gamma}(t)|\, dt < +\infty$. Finally, the assumption $\lambda_0(t)\, \gamma(t) \ge 2\, \dot{\lambda}_0(t)$ is trivially verified since $\gamma(t) \ge \frac{4}{t}$ for every $t \ge t_0$. Lemma 3.2 (ii) then shows that $\lim_{t\to+\infty} t\, \mathcal{E}(t) = 0$. Since $\mathcal{E}(t) \ge \frac{1}{2} \left|\frac{du}{dt}(t)\right|^2$, we deduce that $\lim_{t\to+\infty} t^{1/2} \left|\frac{du}{dt}(t)\right| = 0$. This suggests to apply Lemma 3.2 with the map λ_1 defined by $\lambda_1(t) = t^{3/2}$. The boundedness of the map $\dot{\lambda}_1 \left|\frac{du}{dt}\right|$ is guaranteed by the previous step. The other assumptions of Lemma 3.2 are trivially satisfied. Lemma 3.2 (ii) then shows that $\lim_{t\to+\infty} t^{3/2}\, \mathcal{E}(t) = 0$, thus implying that $\lim_{t\to+\infty} t^{3/4} \left|\frac{du}{dt}(t)\right| = 0$. By using recursively Lemma 3.2, we let the reader check that $\lim_{t\to+\infty} t^{1-\left(\frac{1}{2}\right)^n} \left|\frac{du}{dt}(t)\right| = 0$.

Define the map λ_n by $\lambda_n(t) = t^{2-\left(\frac{1}{2}\right)^n}$. The boundedness of the map $\dot{\lambda}_n \left|\frac{du}{dt}\right|$ is implied by the previous step, while the integral $\int_0^{+\infty} \dot{\lambda}_n(t)\, |\dot{\gamma}(t)|\, dt$ is finite by assumption. Lemma 3.2 applied with the map λ_n yields conclusions (i), (ii) and (iii) of Proposition 3.2.

Given $n \in \mathbb{N}$, $k > 0$ and $t_0 > 0$, the following condition plays a central role in the sequel

$$(l_2) \qquad \begin{cases} \lim_{t\to+\infty} \gamma(t) = 0 \\ \displaystyle\int_0^{+\infty} t^{1-\left(\frac{1}{2}\right)^n} |\dot{\gamma}(t)|\, dt < +\infty \\ \forall t \geq t_0, \quad \gamma(t) \geq \frac{k}{t^{1-\left(\frac{1}{2}\right)^{n+1}}}. \end{cases}$$

Hypothesis (l_2) automatically implies $\dot{\gamma} \in L^1(0, +\infty)$ together with $\gamma \notin L^1(0, +\infty)$.

Remark 3.5 *Assume that the map $\gamma : \mathbb{R}_+ \to \mathbb{R}_+$ is nonincreasing and that there exist $\alpha \in]0, 1[$, k, $k' > 0$ and $t_0 > 0$ such that*

$$\forall t \geq t_0, \quad \frac{k}{t^\alpha} \leq \gamma(t) \leq \frac{k'}{t^\alpha}. \tag{45}$$

Let us show that condition (l_2) is satisfied if the integer $n \in \mathbb{N}$ is chosen such that[10] $\alpha \in \left]1 - \left(\frac{1}{2}\right)^n, 1 - \left(\frac{1}{2}\right)^{n+1}\right]$. *Since $\alpha \leq 1 - \left(\frac{1}{2}\right)^{n+1}$, we have*

$$\frac{k}{t^{1-\left(\frac{1}{2}\right)^{n+1}}} \leq \frac{k}{t^\alpha} \leq \gamma(t),$$

and the third condition of (l_2) is proved. Recalling that $\dot{\gamma}(t) \leq 0$, an immediate integration by parts gives

$$\begin{aligned} \int_{t_0}^t s^{1-\left(\frac{1}{2}\right)^n} |\dot{\gamma}(s)|\, ds &= -\int_{t_0}^t s^{1-\left(\frac{1}{2}\right)^n} \dot{\gamma}(s)\, ds \\ &= t_0^{1-\left(\frac{1}{2}\right)^n} \gamma(t_0) - t^{1-\left(\frac{1}{2}\right)^n} \gamma(t) + \left(1 - \left(\frac{1}{2}\right)^n\right) \int_{t_0}^t \frac{\gamma(s)}{s^{\left(\frac{1}{2}\right)^n}}\, ds. \end{aligned}$$

Since $0 \leq \gamma(t) \leq \frac{k'}{t^\alpha}$ for every $t \geq t_0$, we infer that

$$\int_{t_0}^t s^{1-\left(\frac{1}{2}\right)^n} |\dot{\gamma}(s)|\, ds \leq t_0^{1-\left(\frac{1}{2}\right)^n} \gamma(t_0) + k' \left(1 - \left(\frac{1}{2}\right)^n\right) \int_{t_0}^t \frac{ds}{s^{\alpha+\left(\frac{1}{2}\right)^n}}.$$

[10] Its explicit expression is given by $n = -\left[\frac{\ln(1-\alpha)}{\ln 2}\right] - 1$, where $[x]$ denotes the integer part of $x \in \mathbb{R}$.

From the choice of n, we have $\alpha + \left(\frac{1}{2}\right)^n > 1$*, hence the integral* $\int_{t_0}^{+\infty} \frac{ds}{s^{\alpha+\left(\frac{1}{2}\right)^n}}$ *is convergent. In view of the above inequality, we conclude that* $\int_{t_0}^{+\infty} s^{1-\left(\frac{1}{2}\right)^n} |\dot{\gamma}(s)|\, ds < +\infty$*. Notice that if* $n = 0$*, this condition reduces to* $\int_0^{+\infty} |\dot{\gamma}(t)|\, dt < +\infty$*, which is automatically satisfied since* $\dot{\gamma} \leq 0$*. It follows that if* $\alpha \in \left]0, \frac{1}{2}\right]$*, one may take* $k' = +\infty$ *in condition (45) (no required upper bound).*

Let us now state the main result of this section.

Theorem 3.3 *Assume that the bilinear form* $a(.,.)$ *and the function* f *satisfy respectively* (h_1)*-*(h_2) *and* (k_1)*-*(k_3)*. Let* $\gamma \in W^{1,1}_{loc}(\mathbb{R}_+, \mathbb{R}_+)$ *be a map satisfying* (l_2)*. Let* u *be a solution in the class (4) to equation* (E) *and assume that* $u \in L^\infty(\mathbb{R}_+, H)$*. Then, there exists* $u_\infty \in S$ *such that* $u(t) \rightharpoonup u_\infty$ *weakly in* V *as* $t \to +\infty$*. Furthermore, if* $u(t)$ *strongly converges*[11] *in* H *then it strongly converges in* V*. Finally, if the potential function* F *is even, the convergence is strong in* V*.*

Proof. The proof follows the same lines as the ones of Theorem 3.1. Given $v \in S$, we define the map $p : \mathbb{R}_+ \to \mathbb{R}_+$ by $p(t) = \frac{1}{2}|u(t) - v|^2$. Recall that

$$\dot{p}(t) \leq e^{-\int_0^t \gamma(\tau)\, d\tau}\dot{p}(0) + e^{-\int_0^t \gamma(\tau)\, d\tau} \int_0^t e^{\int_0^s \gamma(\tau)\, d\tau} \left|\frac{du}{ds}(s)\right|^2 ds \tag{46}$$

(see formula 30). We have to show that the right member of the above inequality is a summable function. From Lemma 3.3 (i) below applied with $\theta = 1 - \left(\frac{1}{2}\right)^{n+1}$, we have

$$\int_0^{+\infty} e^{-\int_0^t \gamma(\tau)\, d\tau}\, dt < +\infty. \tag{47}$$

Lemma 3.3 *Let us assume that there exist* $\theta \in [0, 1[$*,* $k > 0$ *and* $t_0 > 0$ *such that* $\gamma(t) \geq \frac{k}{t^\theta}$ *for every* $t \geq t_0$*. Then*

(i) $\int_0^{+\infty} e^{-\int_0^t \gamma(\tau)\, d\tau}\, dt < +\infty$*;*

(ii)For every $c > 1$*, we have for* s *large enough*

$$\int_s^{+\infty} e^{-\int_0^t \gamma(\tau)\, d\tau}\, dt \leq \frac{c}{k}\, s^\theta\, e^{-\int_0^s \gamma(\tau)\, d\tau}. \tag{48}$$

If $\theta = 0$*, one can take* $c = 1$ *in the above inequality.*

The proof of Lemma 3.3 is postponed to the appendix. On the other hand, by applying Fubini theorem, we find

$$\int_0^{+\infty} e^{-\int_0^t \gamma(\tau)\, d\tau} \int_0^t e^{\int_0^s \gamma(\tau)\, d\tau} \left|\frac{du}{ds}(s)\right|^2 ds\, dt = \int_0^{+\infty} \left|\frac{du}{ds}(s)\right|^2 e^{\int_0^s \gamma(\tau)\, d\tau} \int_s^{+\infty} e^{-\int_0^t \gamma(\tau)\, d\tau}\, dt\, ds. \tag{49}$$

[11] This assumption is satisfied if the injection $V \hookrightarrow H$ is compact.

From Lemma 3.3 (ii) applied with $\theta = 1 - \left(\frac{1}{2}\right)^{n+1}$, we obtain

$$e^{\int_0^s \gamma(\tau)\, d\tau} \int_s^{+\infty} e^{-\int_0^t \gamma(\tau)\, d\tau}\, dt \leq \frac{2}{k}\, s^{1-\left(\frac{1}{2}\right)^{n+1}}.$$

Since $\gamma(s) \geq \frac{k}{s^{1-\left(\frac{1}{2}\right)^{n+1}}}$, we derive that

$$e^{\int_0^s \gamma(\tau)\, d\tau} \int_s^{+\infty} e^{-\int_0^t \gamma(\tau)\, d\tau}\, dt \leq \frac{2}{k^2}\, s^{2-\left(\frac{1}{2}\right)^{n}}\, \gamma(s).$$

From Proposition 3.2 (iii) we have $\int_0^{+\infty} s^{2-\left(\frac{1}{2}\right)^{n}}\, \gamma(s)\, \left|\frac{du}{ds}(s)\right|^2 ds < +\infty$, hence we deduce from equality (49) that

$$\int_0^{+\infty} e^{-\int_0^t \gamma(\tau)\, d\tau} \int_0^t e^{\int_0^s \gamma(\tau)\, d\tau} \left|\frac{du}{ds}(s)\right|^2 ds\, dt < +\infty. \tag{50}$$

By combining inequality (46) with estimates (47) and (50), we infer that $[\dot{p}]_+ \in L^1(0, +\infty)$ and hence $\lim_{t\to+\infty} p(t)$ exists. The end of the proof is the same as the one of Theorem 3.1. For the second point, the reader is referred to the corresponding argument in the proof of Corollary 3.1. Finally, if the potential function F is even, the arguments of the proof of Theorem 3.2 apply directly. Details are left to the reader.

Remark 3.6 *The assumption $u \in L^\infty(\mathbb{R}_+, H)$ arises in the statement of Theorem 3.3, while it is useless in the framework of Theorem 3.1. In the proof of this last one, the existence of $\lim_{t\to+\infty} |u(t) - v|^2$ relies on the general estimate $\gamma \left|\frac{du}{dt}\right|^2 \in L^1(0, +\infty)$, and gives the boundedness of u as a by-product. By contrast, in Theorem 3.3 the existence of $\lim_{t\to+\infty} |u(t) - v|^2$ needs a sharper estimate (see Proposition 3.2 (iii)), which uses some boundedness assumption for the map u. The question to know if the assumption $u \in L^\infty(\mathbb{R}_+, H)$ is really necessary in Theorem 3.3 remains open.*

In view of Remark 3.5, we obtain directly the following corollary of Theorem 3.3.

Corollary 3.2 *Assume that the bilinear form $a(.,.)$ and the function f satisfy the same hypotheses as in Theorem 3.3. Let $\gamma \in W^{1,1}_{loc}(\mathbb{R}_+, \mathbb{R}_+)$ be a nonincreasing map and suppose that there exist $\alpha \in]0, 1[$, k, $k' > 0$ and $t_0 > 0$ such that*[12]

$$\forall t \geq t_0, \quad \frac{k}{t^\alpha} \leq \gamma(t) \leq \frac{k'}{t^\alpha}.$$

Then we have the same conclusions as in Theorem 3.3.

[12] This condition is satisfied if there exists $k'' > 0$ such that $\gamma(t) \sim \frac{k''}{t^\alpha}$ as $t \to +\infty$. On the other hand, one can take $k' = +\infty$ if $\alpha \in \left]0, \frac{1}{2}\right]$, see Remark 3.5.

3.3 Decay estimates for a strong set of minima

Recall that the set $S = \text{Argmin}\Phi$ is convex and closed in H, see Remark 3.2. Let us denote by P_S the projection operator onto the set S in the sense of H. In this paragraph, we assume that the function $\Phi : V \to \mathbb{R}$ satisfies[13]

$$\exists \eta > 0 \quad \text{such that} \quad \forall u \in V, \quad \Phi(u) \geq \frac{\eta}{2}|u - P_S(u)|^2. \tag{51}$$

If $\gamma \notin L^1(0, +\infty)$, we know from Proposition 3.1 (ii) that $\lim_{t\to+\infty} \mathcal{E}(t) = 0$. Under assumption (51), we are able to evaluate the speed of convergence of $\mathcal{E}(t)$ as $t \to +\infty$.

Theorem 3.4 *Assume that the bilinear form $a(.,.)$ and the function f satisfy respectively (h_1)-(h_2) and (k_1)-(k_3). Let $\gamma \in W^{1,1}_{loc}(\mathbb{R}_+, \mathbb{R}_+)$ be a function satisfying $\lim_{t\to+\infty} \gamma(t) = 0$ and $\dot\gamma(t) = o(\gamma(t))$ as $t \to +\infty$. We suppose that the function $\Phi : V \to \mathbb{R}$ defined by $\Phi(u) = \frac{1}{2}a(u,u) + F(u)$ satisfies condition (51). Let u be a solution in the class (4) to equation (E). Then, for all $m \in]0, \frac{2}{3}[$, there exist $C > 0$ and $t_0 \geq 0$ such that:*

$$\forall t \geq t_0, \quad \mathcal{E}(t) \leq Ce^{-m\int_0^t \gamma(s)ds}.$$

Proof. Define the map $\varphi : \mathbb{R}_+ \to \mathbb{R}$ by $\varphi(t) = \frac{1}{2}d_H^2(u(t), S)$, where $d_H(., S)$ stands for the distance function from the set S in the sense of H. By differentiating, we find for every $t \geq 0$

$$\dot\varphi(t) = \left(\frac{du}{dt}(t), u(t) - P_S(u(t))\right). \tag{52}$$

Since $\frac{du}{dt} \in W^{1,1}_{loc}(\mathbb{R}_+, H)$ by assumption, it is immediate to check that $\dot\varphi \in W^{1,1}_{loc}(\mathbb{R}_+)$, hence the map $\dot\varphi$ is differentiable almost everywhere on $\mathbb{R}_+$. Consider now some $t > 0$ where the maps $\dot\varphi$ and $\frac{du}{dt}$ are both differentiable, and let us majorize the quantity $\ddot\varphi(t)$. For that purpose, we use a technique of differential quotient. For all $h \neq 0$, we have

$$\begin{aligned}\frac{1}{h}(\dot\varphi(t+h) - \dot\varphi(t)) = &\frac{1}{h}\left(\frac{du}{dt}(t), u(t+h) - P_S(u(t+h)) - u(t) + P_S(u(t))\right)\\ &+\frac{1}{h}\left(\frac{du}{dt}(t+h) - \frac{du}{dt}(t), u(t+h) - P_S(u(t+h))\right).\end{aligned}$$

The monotonicity of P_S implies that

$$-\frac{1}{h}\left(\frac{du}{dt}(t), P_S(u(t+h)) - P_S(u(t))\right)$$

[13] If $f = 0$, the set S coincides with $\text{Ker}A$ and we recover condition (12) of section 2.

$$\leq \frac{1}{h^2}\left(u(t+h)-u(t)-h\frac{du}{dt}(t), P_S(u(t+h))-P_S(u(t))\right).$$

Hence we obtain

$$\begin{aligned}\frac{1}{h}(\dot\varphi(t+h)-\dot\varphi(t)) \leq \frac{1}{h}&\left(\frac{du}{dt}(t), u(t+h)-u(t)\right)\\ &+\frac{1}{h^2}\left(u(t+h)-u(t)-h\frac{du}{dt}(t), P_S(u(t+h))-P_S(u(t))\right)\\ &+\frac{1}{h}\left(\frac{du}{dt}(t+h)-\frac{du}{dt}(t), u(t+h)-P_S(u(t+h))\right).\end{aligned}$$

Taking the limit as $h\to 0$, we derive that

$$\ddot\varphi(t) \leq \left|\frac{du}{dt}(t)\right|^2 + \left(\frac{d^2u}{dt^2}(t), u(t)-P_S(u(t))\right). \tag{53}$$

By combining formulae (52) and (53), and using the convexity of the function Φ, we deduce that for almost every $t\in\mathbb{R}_+$

$$\begin{aligned}\ddot\varphi(t)+\gamma(t)\dot\varphi(t) &\leq \left|\frac{du}{dt}(t)\right|^2 + \left(\frac{d^2u}{dt^2}(t)+\gamma(t)\,\frac{du}{dt}(t), u(t)-P_S(u(t))\right)\\ &= \left|\frac{du}{dt}(t)\right|^2 - a\left(u(t), u(t)-P_S(u(t))\right) - \left(f(u(t)), u(t)-P_S(u(t))\right)\\ &\leq \left|\frac{du}{dt}(t)\right|^2 - \Phi(u(t)) + \Phi(P_S(u(t))) = \left|\frac{du}{dt}(t)\right|^2 - \Phi(u(t)).\end{aligned}$$

It follows that

$$\ddot\varphi(t)+\gamma(t)\dot\varphi(t)+\mathcal{E}(t) \leq \frac{3}{2}\left|\frac{du}{dt}(t)\right|^2 \qquad \text{a.e. on } \mathbb{R}_+.$$

Multiplying this formula by $\frac{2}{3}\gamma(t)$ and recalling that $\dot{\mathcal{E}}(t) = -\gamma(t)\left|\frac{du}{dt}(t)\right|^2$ for almost every $t\in\mathbb{R}_+$, we obtain

$$\frac{2}{3}\,\gamma(t)\left(\ddot\varphi(t)+\gamma(t)\dot\varphi(t)\right)+\dot{\mathcal{E}}(t)+\frac{2}{3}\,\gamma(t)\,\mathcal{E}(t) \leq 0 \qquad \text{a.e. on } \mathbb{R}_+. \tag{54}$$

This suggests to define the function $\mathcal{F}:\mathbb{R}_+\to\mathbb{R}$ by

$$\begin{aligned}\mathcal{F}(t) &= \Phi(u(t)) + \frac{1}{2}\left|\frac{du}{dt}(t)\right|^2 + \frac{2}{3}\gamma(t)\left(\frac{du}{dt}(t), u(t)-P_S(u(t))\right) \qquad (55)\\ &= \mathcal{E}(t)+\frac{2}{3}\gamma(t)\,\dot\varphi(t).\end{aligned}$$

In view of inequality (54), we immediately find

$$\dot{\mathcal{F}}(t)+\frac{2}{3}\gamma(t)\mathcal{F}(t) \leq \frac{2}{3}\left(\dot{\gamma}(t)-\frac{1}{3}\gamma(t)^2\right)\left(\frac{du}{dt}(t), u(t)-P_S(u(t))\right) \qquad \text{a.e. on } \mathbb{R}_+. \tag{56}$$

Since $\left|\left(\frac{du}{dt}(t), u(t)-P_S(u(t))\right)\right| \leq \frac{1}{2}\left|\frac{du}{dt}(t)\right|^2+\frac{1}{2}\left|u(t)-P_S(u(t))\right|^2$ and $\Phi(u(t)) \geq \frac{\eta}{2}|u(t)-P_S(u(t))|^2$ by assumption, we have

$$\left|\left(\frac{du}{dt}(t), u(t)-P_S(u(t))\right)\right| \leq C\,\mathcal{E}(t), \qquad \text{for some } C>0. \tag{57}$$

Recalling that $\lim_{t\to+\infty}\gamma(t)=0$, the expression of $\mathcal{F}$ shows that

$$\mathcal{F}(t) \sim \mathcal{E}(t) \qquad \text{as } t\to+\infty. \tag{58}$$

Let us fix some $m \in]0,\frac{2}{3}[$. Using the fact that $\dot{\gamma}(t)=o(\gamma(t))$ and $\gamma(t)^2=o(\gamma(t))$ as $t\to+\infty$, we deduce from (56), (57) and (58) the existence of $t_0\geq 0$ such that,

$$\dot{\mathcal{F}}(t)+\frac{2}{3}\gamma(t)\mathcal{F}(t) \leq \left(\frac{2}{3}-m\right)\gamma(t)\mathcal{F}(t) \qquad \text{a.e. on } [t_0,+\infty[,$$

hence $\dot{\mathcal{F}}(t)+m\gamma(t)\mathcal{F}(t)\leq 0$ for almost every $t\geq t_0$. Let us multiply by $e^{m\int_0^t \gamma(s)ds}$ and integrate on $[t_0,t]$. Since the function $\mathcal{F}$ is absolutely continuous, we find $\mathcal{F}(t) \leq D\,e^{-m\int_0^t\gamma(s)ds}$, with $D=e^{m\int_0^{t_0}\gamma(s)ds}\,F(t_0)$. Conclusion follows from estimate (58).

Remark 3.7 *Under the hypotheses of Theorem 3.4, assume that there exists $k>3$ such that $\gamma(t)\geq\frac{k}{t}$ for t large enough. Fix $m\in\left]\frac{2}{k},\frac{2}{3}\right[$. From Theorem 3.4, there exist $C>0$ and $t_0\geq 0$ such that*

$$\forall t\geq t_0, \qquad \frac{1}{2}\left|\frac{du}{dt}(t)\right|^2 \leq \mathcal{E}(t) \leq \frac{C}{t^{mk}}.$$

Hence we have $\left|\frac{du}{dt}(t)\right| \leq \frac{(2C)^{1/2}}{t^{mk/2}}$ and since $mk>2$, we deduce that $\left|\frac{du}{dt}\right|\in L^1(0,+\infty)$. The trajectory u has a finite length, therefore it strongly converges in H toward some $u_\infty\in S$.

4 Application to particular semilinear evolution problems

We suppose that Ω is a bounded open subset of $\mathbb{R}^n$ with boundary $\partial\Omega$ sufficiently regular.

4.1 Hyperbolic problems of order two in space

Example 4.1 Given a map $\gamma:\mathbb{R}_+\to\mathbb{R}_+$ and a function $f\in\mathcal{C}^1(\mathbb{R})$, let us consider the following damped wave equation

$$\frac{d^2u}{dt^2} + \gamma(t)\,\frac{du}{dt} - \Delta u + f(u) = 0 \quad \text{on} \quad \Omega\times]0, +\infty[, \tag{59}$$

with Dirichlet boundary condition:

$$u = 0 \quad \text{on} \quad \partial\Omega\times]0, +\infty[. \tag{60}$$

The functional setting of the evolution problem (59)-(60) is given by

$$H = L^2(\Omega), \quad V = H_0^1(\Omega) \quad \text{and} \quad a(u, v) = \int_\Omega \nabla u(x)\nabla v(x)dx.$$

Hypothesis (h_1) is trivially verified while hypothesis (h_2) is satisfied with $\lambda = 0$, since the bilinear form a is coercive. On the other hand, we assume that the function f satisfies the following properties:

(i) There exist $C, \alpha \geq 0$ such that $(n-2)\alpha \leq 2$ and $|f'(r)| \leq C\ (1 + |r|^\alpha) \quad \forall r \in \mathbb{R}$.

(ii) f is nondecreasing.

Define the function $F \in \mathcal{C}^2(\mathbb{R})$ by $F(r) = \int_0^r f(s)\,ds$ for every $r \in \mathbb{R}$. For simplicity of notation, we write $F(u)$ for $\int_\Omega F(u(x))\,dx$. Hypothesis (k_1) is a consequence of assumption (i) above, see for example [10, pp. 73-75]. The monotonicity hypothesis (k_2) is ensured by point (ii). Finally the coercivity of the bilinear form a implies that the equilibrium set is a singleton $\{\overline{u}\}$, see Remark 3.1. In particular, the non-vacuity condition (k_3) is satisfied. If the map $\gamma \in W^{1,1}_{loc}(\mathbb{R}_+, \mathbb{R}_+)$ is such that $\dot\gamma \in L^1(0, +\infty)$ and $\gamma \notin L^1(0, +\infty)$, we derive from Corollary 3.1 that $u(t) \rightharpoonup \overline{u}$ weakly in $H_0^1(\Omega)$ as $t \to +\infty$. Since the injection $H_0^1(\Omega) \hookrightarrow L^2(\Omega)$ is compact, the second part of Corollary 3.1 shows that the convergence is strong in $H_0^1(\Omega)$. On the other hand, the coercivity of a implies that condition (51) is fulfilled. If the map γ satisfies $\lim_{t\to+\infty} \gamma(t) = 0$ and $\dot\gamma(t) = o(\gamma(t))$ as $t \to +\infty$, Theorem 3.4 then shows that for every $m \in \left]0, \frac{2}{3}\right[$,

$$\frac{1}{2}\int_\Omega \left\{\left|\frac{\partial u}{\partial t}(t, x)\right|^2 + |\nabla u(t, x)|^2\right\} dx + \int_\Omega F(u(t, x))\,dx = O\left(e^{-m\int_0^t \gamma(s)\,ds}\right) \qquad \text{as } t \to +\infty.$$

Example 4.2 Let us consider the damped wave equation (59) with Neumann boundary condition $\frac{\partial u}{\partial n} = 0$ on $\partial\Omega\times]0, +\infty[$. The functional setting of the evolution problem is given by:

$$H = L^2(\Omega), \quad V = H^1(\Omega) \quad \text{and} \quad a(u, v) = \int_\Omega \nabla u(x).\nabla v(x)dx.$$

The bilinear form a is semi-coercive, hypothesis (h_2) is satisfied with $\lambda = \mu = 1$. To overcome the lack of coercivity, assumptions (i)-(ii) above are supplemented with the following one

(iii) There exist $\varepsilon > 0$ and $D \geq 0$ such that $F(r) \geq \varepsilon\, r^2 - D$ for every $r \in \mathbb{R}$.

Assumption (iii) implies that condition (k_3) is verified, see Remark 3.1. Hypotheses (k_1)-(k_2) are fulfilled as in the previous example. If the map $\gamma \in W^{1,1}_{loc}(\mathbb{R}_+, \mathbb{R}_+)$ satisfies (l_1) (resp. (l_2)), we derive from Theorem 3.1 (resp. 3.3) that there exists a solution u_∞ of

$$\begin{cases} -\Delta u + f(u) = 0 \text{ in } \ \Omega \\ \qquad\qquad \frac{\partial u}{\partial n} = 0 \text{ on } \partial\Omega \end{cases}$$

such that $u(t) \rightharpoonup u_\infty$ weakly in $H^1(\Omega)$ as $t \to +\infty$. Since the injection $H^1(\Omega) \hookrightarrow L^2(\Omega)$ is compact, the second part of Theorem 3.1 (resp. 3.3) shows that the convergence is strong in $H^1(\Omega)$.

Example 4.3 Let us consider the following equation

$$\frac{d^2u}{dt^2} + \gamma(t)\,\frac{du}{dt} - \Delta u - \lambda_1 u + f(u) = 0 \quad \text{on} \quad \Omega\times]0, +\infty[, \tag{61}$$

with Dirichlet boundary condition. Here λ_1 stands for the smallest eigenvalue of the Laplacian-Dirichlet operator. The functional setting of the evolution problem is given by:

$$H = L^2(\Omega), \quad V = H^1_0(\Omega) \quad \text{and} \quad a(u, v) = \int_\Omega [\nabla u(x).\nabla v(x) - \lambda_1 u(x)v(x)]\, dx.$$

It is immediate to check that (h_1)-(h_2) are satisfied. Under the above assumptions (i), (ii) and (iii), we obtain as previously that conditions (k_1)-(k_3) hold true. If the map $\gamma \in W^{1,1}_{loc}(\mathbb{R}_+, \mathbb{R}_+)$ satisfies (l_1) (resp. (l_2)), we derive from Theorem 3.1 (resp. 3.3) that there exists a solution u_∞ of

$$\begin{cases} -\Delta u - \lambda_1\, u + f(u) = 0 \text{ in } \ \Omega \\ \qquad\qquad\qquad\quad u = 0 \text{ on } \partial\Omega \end{cases}$$

such that $u(t) \to u_\infty$ strongly in $H^1_0(\Omega)$ as $t \to +\infty$.

Example 4.4 The equation arising in the previous example can be generalized as follows

$$\frac{d^2u}{dt^2} + \gamma(t)\,\frac{du}{dt} - \Delta u - \sum_{i=1}^{+\infty} \eta_i P_i u + f(u) = 0 \quad \text{on} \quad \Omega\times]0, +\infty[,$$

see [21, Example 4.5]. We still assume Dirichlet boundary conditions. Let us explicit the notations: $(\lambda_i)_{i\geq 1}$ (respectively $(e_i)_{i\geq 1}$) is the sequence of eigenvalues (respectively eigenfunctions normalized in $L^2(\Omega)$) of $(-\Delta)$ in $H^1_0(\Omega)$. For each $i \geq 1$, P_i denotes the orthogonal projection on span$\{e_i\}$ in the sense of $L^2(\Omega)$. We assume that the nonnegative sequence $(\eta_i)_{i\geq 1}$ is bounded and that $\eta_i \leq \lambda_i$ for every $i \geq 1$. The functional setting of the evolution problem is given by

$$H = L^2(\Omega), \quad V = H_0^1(\Omega) \quad \text{and} \quad a(u,v) = \int_\Omega \nabla u(x).\nabla v(x)dx - \sum_{i=1}^{+\infty} \eta_i \int_\Omega P_i u(x).P_i v(x)dx.$$

It is easy to check that hypotheses (h_1)-(h_2) hold true. Under the additional assumptions (i), (ii) and (iii), we then obtain (k_1)-(k_3). If the map $\gamma \in W^{1,1}_{loc}(\mathbb{R}_+, \mathbb{R}_+)$ satisfies (l_1) or (l_2), we obtain as in the previous example the existence of an equilibrium u_∞ such that $u(t) \to u_\infty$ strongly in $H_0^1(\Omega)$ as $t \to +\infty$.

4.2 A higher-order example

Example 4.5 Let us consider the following equation

$$\frac{d^2u}{dt^2} + \gamma(t)\,\frac{du}{dt} + \Delta^2 u + f(u) = 0 \quad \text{on} \quad \Omega\times]0,+\infty[, \tag{62}$$

with the boundary condition:

$$u = \frac{\partial u}{\partial n} = 0 \quad \text{on} \quad \partial\Omega\times]0,+\infty[. \tag{63}$$

The functional setting of the evolution problem (62)-(63) is given by:

$$H = L^2(\Omega), \quad V = \left\{ u \in H^2(\Omega), \quad u = \frac{\partial u}{\partial n} = 0 \text{ on } \partial\Omega \right\} \quad \text{and} \quad a(u,v) = \int_\Omega \Delta u(x).\Delta v(x)dx.$$

Hypothesis (h_1) is trivially verified. Moreover, from the regularity results relative to the Laplacian-Dirichlet problem, there exists $\kappa > 0$ such that $\|u\|_{H^2(\Omega)} \leq \kappa\,|\Delta u|_{L^2(\Omega)}$. Hence condition (h_2) is satisfied with $\lambda = 0$, *i.e.* the bilinear form a is coercive. We assume that the function f satisfies assumption (ii) along with the following variant of (i)

(i') There exist $C, \alpha \geq 0$ such that $(n-4)\alpha \leq 4$ and $|f'(r)| \leq C\,(1 + |r|^\alpha) \quad \forall r \in \mathbb{R}$.

By using Sobolev's imbedding theorem, we let the reader check that hypothesis (k_1) is a consequence of assumption (i') above. The monotonicity hypothesis (k_2) is ensured by (ii). Finally in view of Remark 3.1, the coercivity of the bilinear form a implies that the equilibrium set is a singleton $\{\overline{u}\}$ and in particular (k_3) holds true. If the map $\gamma \in W^{1,1}_{loc}(\mathbb{R}_+, \mathbb{R}_+)$ is such that $\dot\gamma \in L^1(0,+\infty)$ and $\gamma \notin L^1(0,+\infty)$, we derive from Corollary 3.1 that $u(t) \to \overline{u}$ strongly in $H^2(\Omega)$ as $t \to +\infty$. On the other hand, the coercivity of a implies that condition (51) is fulfilled. If the map γ is such that $\lim_{t\to+\infty} \gamma(t) = 0$ and $\dot\gamma(t) = o(\gamma(t))$ as $t \to +\infty$, Theorem 3.4 then shows that for every $m \in \left]0, \frac{2}{3}\right[$,

$$\frac{1}{2}\int_\Omega \left\{ \left|\frac{\partial u}{\partial t}(t,x)\right|^2 + |\Delta u(t,x)|^2 \right\} dx + \int_\Omega F(u(t,x))\,dx = O\left(e^{-m\int_0^t \gamma(s)\,ds}\right) \quad \text{as } t \to +\infty.$$

Appendix

Proof of Lemma 3.3. (i) From the assumption $\gamma(t) \geq \frac{k}{t^\theta}$, we deduce the existence of $\alpha \in \mathbb{R}$ such that $\int_0^t \gamma(\tau)\, d\tau \geq \frac{k}{1-\theta} t^{1-\theta} + \alpha$ for every $t \geq t_0$. Therefore, we have

$$\int_0^{+\infty} e^{-\int_0^t \gamma(\tau)\, d\tau}\, dt \leq e^{-\alpha} \int_0^{+\infty} e^{-\frac{k}{1-\theta} t^{1-\theta}}\, dt < +\infty.$$

(ii) By using the assumption $\gamma(t) \geq \frac{k}{t^\theta}$, we find

$$\int_s^{+\infty} e^{-\int_0^t \gamma(\tau)\, d\tau}\, dt \leq \frac{1}{k} \int_s^{+\infty} t^\theta\, \gamma(t)\, e^{-\int_0^t \gamma(\tau)\, d\tau}\, dt. \tag{64}$$

An integration by parts in the right-hand side then yields

$$\int_s^{+\infty} t^\theta\, \gamma(t)\, e^{-\int_0^t \gamma(\tau)\, d\tau}\, dt = \left[-t^\theta\, e^{-\int_0^t \gamma(\tau)\, d\tau} \right]_s^{+\infty} + \theta \int_s^{+\infty} t^{\theta-1}\, e^{-\int_0^t \gamma(\tau)\, d\tau}\, dt. \tag{65}$$

Remark that $t^\theta\, e^{-\int_0^t \gamma(\tau)\, d\tau} \leq e^{-\alpha} t^\theta\, e^{-\frac{k}{1-\theta} t^{1-\theta}}$, hence $\lim_{t\to+\infty} t^\theta\, e^{-\int_0^t \gamma(\tau)\, d\tau} = 0$. Therefore, we deduce from (64) and (65) that

$$\int_s^{+\infty} e^{-\int_0^t \gamma(\tau)\, d\tau}\, dt \leq \frac{1}{k}\, s^\theta\, e^{-\int_0^s \gamma(\tau)\, d\tau} + \frac{\theta}{k} \int_s^{+\infty} t^{\theta-1}\, e^{-\int_0^t \gamma(\tau)\, d\tau}\, dt.$$

If $\theta = 0$, formula (48) is proved with $c = 1$. Now assume that $\theta \in]0,1[$ and take $c > 1$. The right term in the above inequality is clearly negligible with respect to the left one, hence $\displaystyle \frac{\theta}{k} \int_s^{+\infty} t^{\theta-1}\, e^{-\int_0^t \gamma(\tau)\, d\tau}\, dt \leq \left(1 - \frac{1}{c}\right) \int_s^{+\infty} e^{-\int_0^t \gamma(\tau)\, d\tau}\, dt$ for s large enough. Formula (48) follows immediately. □

References

[1] M. Abramowitz, I. Stegun, Handbook of Mathematical Functions. Dover, New York, 1972.

[2] F. Alvarez, On the minimizing property of a second order dissipative system in Hilbert spaces, *SIAM j.control optim.*, vol. 38, no. 4, pp. 1102-1119.

[3] F. Alvarez, H. Attouch, Convergence and asymptotic stabilization for some damped hyperbolic equations with non-isolated equilibria, *ESAIM*, 6 (2001), pp. 539-552.

[4] H. Attouch, X. Goudou, P. Redont, The heavy ball with friction method: I the continuous dynamical system, *Communications in Contemporary Mathematics*, 2 (2000), pp. 1-34.

[5] F. Bowman, Introduction to Bessel Functions. Dover, New York, 1958.

[6] R.E. Bruck, Asymptotic convergence of nonlinear contraction semigroups in Hilbert space, *Journal of Functional Analysis*, 18 (1975), pp. 15-26.

[7] A. Cabot, Asymptotics for a gradient system with memory term, *Proc. of the Amer. Math. Soc.*, 137 (2009), 3013-3024.

[8] A. Cabot, H. Engler, S. Gadat, On the long time behavior of second order differential equations with asymptotically small dissipation, *Trans. Amer. Math. Soc.*, 361 (2009), no. 11, pp. 5983-6017.

[9] A. Cabot, H. Engler, S. Gadat, Second order differential equations with asymptotically small dissipation and piecewise flat potentials, *Electronic Journal of Differential Equations*, 17 (2009), 33-38.

[10] T. Cazenave, A. Haraux, An introduction to semilinear evolution equations, Oxford Lecture Series in Mathematics and its Applications 13 (1998).

[11] J.K. Hale, Asymptotic behavior of dissipative systems, Mathematical Surveys and Monographs 25, Amer. Math. Soc., Providence (1988).

[12] A. Haraux, Semilinear hyperbolic problems in bounded domains, Mathematical reports, Vol 3, Harwood Academic Publishers, Gordon & Breach, 1987.

[13] A. Haraux, Systèmes dynamiques dissipatifs et applications, Masson, Paris, 1991.

[14] A. Haraux, M. A. Jendoubi, Convergence of solutions of second-order gradient-like systems with analytic nonlinearities, *J. Differential Equations*, 144 (1998), pp. 313-320.

[15] A. Haraux, M. A. Jendoubi, Convergence of bounded weak solutions of the wave equation with dissipation and analytic nonlinearity, *Calc. Var. Partial Differential Equations*, 9 (1999), pp. 95-124.

[16] M. A. Jendoubi, Convergence of global and bounded solutions of the wave equation with linear dissipation and analytic nonlinearity, *J. Differential Equations*, 144 (1998), pp. 302-312.

[17] B. Larrouturou, Modélisation mathématique et numérique pour les sciences de l'ingénieur, Majeure Science de l'ingénieur et calcul scientifique, Tome 1, Ecole Polytechnique, 1988.

[18] Z. Opial, Weak convergence of the sequence of successive approximations for nonexpansive mappings, *Bull. Amer. Math. Soc.*, 73 (1967), pp. 591-597.

[19] L. Simon, Asymptotics for a class of nonlinear evolution equations, with applications to geometric problems, *Ann. Math.*, 118 (1983), pp. 525-571.

[20] R. Temam, Infinite-Dimensional Dynamical Systems in Mechanics and Physics, Applied Mathematical Sciences 68, Springer-Verlag, New York, 1988.

[21] E. Zuazua, Stability and decay for a class of nonlinear hyperbolic problems, *Asympto. Anal.*, 1 (1998), pp. 161-185.

II. COMPORTEMENT ASYMPTOTIQUE DE METHODES PROXIMALES METTANT EN JEU INERTIE ET MEMOIRE

Asymptotics for some proximal-like method involving inertia and memory aspects

A. Cabot[1] and P. Frankel[2]

[1] Université Montpellier II, Place Eugène Bataillon, 34095 Montpellier Cedex 05, France
acabot@math.univ-montp2.fr
[2] Université Montpellier II, Place Eugène Bataillon, 34095 Montpellier Cedex 05, France
p.frankel30@orange.fr

Accepted in *Set-Valued and Variational Analysis*

Summary. Given a Hilbert space H and a closed convex function $\Phi : H \to \mathbb{R} \cup \{+\infty\}$, we consider the inertial proximal algorithm

$$(\mathcal{A}) \qquad x_{n+1} - x_n - \alpha_n(x_n - x_{n-1}) + \beta_n \partial\Phi(x_{n+1}) \ni 0,$$

where (α_n) and (β_n) are nonnegative sequences. The notation $\partial\Phi$ stands for the subdifferential of Φ in the sense of convex analysis. This algorithm can be viewed as the implicit discretization of a continuous gradient system involving a memory term. We give conditions that ensure that a suitable discrete energy decreases to $\inf \Phi$ as $n \to +\infty$. When Φ has a unique minimum, the question of the convergence of (x_n) is solved. In the case of multiple minima, it is proved that if $(\prod_{k=1}^n \alpha_k) \notin l^1$ and if a suitable geometric condition on the set ArgminΦ is fulfilled, then non stationary sequences of $(\mathcal{A})$ cannot converge.

Key words: Proximal point algorithm, averaged gradient method, dissipative dynamical system, memory effect.
Subject classification:65K10, 49M25.

1 Introduction

Let H be a Hilbert space endowed with the scalar product $\langle ., . \rangle$ and the corresponding norm $| . |$. We consider a smooth convex potential function $\Phi : H \to \mathbb{R}$ to be minimized. A classical approach consists in following the orbits of the steepest descent method. In a series of recent papers [6, 7, 8], a special attention was devoted to gradient systems involving memory terms. The model considered in [6] corresponds to the following continuous dynamical system

$$(S) \qquad \dot{x}(t) + \frac{1}{k(t)} \int_0^t h(s)\, \nabla\Phi(x(s))\, ds = 0, \qquad t \geq 0,$$

where $h, k : [0, +\infty) \to \mathbb{R}_+^*$ are continuous maps. If $k(t) \sim \int_0^t h(s)\, ds$ as $t \to +\infty$, this equation can be interpreted as an averaged gradient system. When Φ is convex and has multiple minima, it is proved in [6] that the trajectories of (S)

converge if and only if the weighted memory privileges the recent past enough.

For numerical purposes, it is natural to deal with a discretized version of (S). In this paper, we are interested in the following implicit discretization of (S)

$$x_{n+1} - x_n + \frac{1}{k_{n+1}} \sum_{i=0}^{n} h_i \nabla\Phi(x_{i+1}) = 0, \tag{1}$$

where (h_n) and (k_n) are suitable sequences. A special attention will be devoted to the particular case corresponding to $h_n = n^a$, $k_n = n^b$ for every $n \in \mathbb{N}$. A major goal of this paper is to give a satisfying description of the behavior of the corresponding algorithm, for every a, $b \geq 0$. Iteration (1) can be rewritten as

$$k_{n+1}\,(x_{n+1} - x_n) - k_n\,(x_n - x_{n-1}) + h_n\,\nabla\Phi(x_{n+1}) = 0,$$

which is in turn equivalent to

$$x_{n+1} - x_n - \alpha_n(x_n - x_{n-1}) + \beta_n \nabla\Phi(x_{n+1}) = 0,$$

by setting $\alpha_n = \frac{k_n}{k_{n+1}}$ and $\beta_n = \frac{h_n}{k_{n+1}}$. The extrapolation term $\alpha_n\,(x_n - x_{n-1})$ takes into account a kind of inertia associated with the sequence. If the convex function Φ is not assumed to be smooth and takes its values in $\mathbb{R} \cup \{+\infty\}$, one can easily adapt the previous algorithm as follows

$$(\mathcal{A}) \qquad x_{n+1} - x_n - \alpha_n(x_n - x_{n-1}) + \beta_n \partial\Phi(x_{n+1}) \ni 0,$$

where ∂ denotes the subdifferential in the sense of convex analysis. When $\alpha_n \equiv 0$, we recover the standard proximal point algorithm, for which we refer the reader to the abundant literature on this subject [15, 17, 18, 20]. The inertial proximal algorithm $(\mathcal{A})$ was studied in [1, 2] and various extensions were considered in [3, 13, 16, 19]. It is proved in [1] that the sequence (x_n) generated by $(\mathcal{A})$ weakly converges toward a minimum of Φ, provided that the sequence (α_n) is bounded from above by some $\overline{\alpha} \in [0, 1[$. One of the purposes of this paper is to get rid of this assumption and to examine what happens when $\lim_{n\to+\infty} \alpha_n = 1$.

The paper is organized as follows. In section 2, we exhibit a discrete energy (E_n) for algorithm $(\mathcal{A})$ and we compute the corresponding decay. It is shown in section 3 that the energy (E_n) converges toward $\min \Phi$ as $n \to +\infty$, under suitable conditions on (α_n) and (β_n). This enables us to solve the question of the convergence of (x_n) in the case of a unique minimum. The case of multiple minima is more delicate and is discussed in section 4. We prove that if $(\prod_{k=1}^{n} \alpha_k) \notin l^1$ and if a suitable geometric condition on the set $\mathrm{Argmin}\Phi$ is fulfilled, then non stationary sequences of $(\mathcal{A})$ cannot converge. The question of the convergence under condition $(\prod_{k=1}^{n} \alpha_k) \in l^1$ is difficult and still open in its full generality.

2 General facts. Energy decay

In the entire paper, we assume that $\Phi : H \to \mathbb{R} \cup \{+\infty\}$ is a closed convex function and that the sequences (α_n), (β_n) are positive. Iteration $(\mathcal{A})$ can be equivalently rewritten as

$$x_{n+1} = J_{\beta_n}(x_n + \alpha_n(x_n - x_{n-1}))$$

where $J_{\beta_n} = (I + \beta_n \partial\Phi)^{-1}$ is the resolvent of index β_n of the maximal monotone operator $\partial\Phi$. This shows that for any couple $(x_0, x_1) \in H^2$ of initial data, there exists a unique sequence (x_n) satisfying algorithm $(\mathcal{A})$.

Remark 2.1 *It is worthwhile noticing that algorithm* $(\mathcal{A})$ *can be reformulated as*

$$\frac{x_{n+1} - 2x_n + x_{n-1}}{\beta_n} + \frac{1-\alpha_n}{\beta_n}(x_n - x_{n-1}) + \partial\Phi(x_{n+1}) \ni 0. \tag{2}$$

Hence algorithm $(\mathcal{A})$ *appears as a discretization of the following second-order in time differential inclusion*

$$\ddot{x}(t) + \gamma(t)\dot{x}(t) + \partial\Phi(x(t)) \ni 0, \qquad t \geq 0, \tag{3}$$

where γ *is a time-dependent damping. In the finite difference scheme (2), the step length equals* $\sqrt{\beta_n}$*, while* $\frac{1-\alpha_n}{\sqrt{\beta_n}}$ *corresponds to the value of* $\gamma(.)$ *at time* $t_n = \sum_{k=0}^{n} \sqrt{\beta_k}$*. This interpretation of* $(\mathcal{A})$ *will be used to enlighten some aspects of the paper.*

The result below states the decay property of the energy (E_n) defined, for every $n \in \mathbb{N}$, by[3]

$$E_n = \frac{1}{2\beta_{n-1}}|x_n - x_{n-1}|^2 + \Phi(x_n).$$

Proposition 2.1 *Let* $\Phi : H \to \mathbb{R} \cup \{+\infty\}$ *be a closed convex function and let* (α_n), (β_n) *be two positive sequences such that* $\alpha_n \leq 1$ *and* $\alpha_n \leq \frac{\beta_n}{\beta_{n-1}}$ *for every* $n \geq 1$*. Then any sequence* (x_n) *defined by* $(\mathcal{A})$ *satisfies*[4]

$$E_{n+1} - E_n \leq -\frac{1-\alpha_n}{2\,\beta_n}|x_{n+1} - x_n|^2. \tag{4}$$

If moreover the function Φ *is bounded from below then*

(i) The nonincreasing sequence (E_n) *converges toward some* $E_\infty \in \mathbb{R}$.
(ii) There exists $C > 0$ *such that* $|x_{n+1} - x_n| \leq C\sqrt{\beta_n}$ *for every* $n \geq 0$*. In particular, if* $\left(\sqrt{\beta_n}\right) \in l^1$ *then* $(|x_{n+1} - x_n|) \in l^1$*, hence* $\overline{x} = \lim_{n\to+\infty} x_n$ *exists.*
(iii) The following estimate holds true: $\sum_{n=0}^{+\infty} \frac{1-\alpha_n}{2\,\beta_n}|x_{n+1} - x_n|^2 < +\infty$.

[3] Notice that there is a slight difference with the corresponding energy given in [1].
[4] The expression of the energy decay is clearly related to the damping coefficient $\frac{1-\alpha_n}{\sqrt{\beta_n}}$, see Remark 2.1.

Proof. Let $\xi_{n+1} \in \partial\Phi(x_{n+1})$ be such that $x_{n+1} - x_n - \alpha_n (x_n - x_{n-1}) + \beta_n \xi_{n+1} = 0$. From the definition of the subdifferential of Φ, we have

$$\begin{aligned}\Phi(x_{n+1}) - \Phi(x_n) &\leq \langle \xi_{n+1}, x_{n+1} - x_n \rangle \\ &= -\frac{1}{\beta_n}|x_{n+1} - x_n|^2 + \frac{\alpha_n}{\beta_n}\langle x_n - x_{n-1}, x_{n+1} - x_n \rangle \\ &\leq -\frac{1}{\beta_n}|x_{n+1} - x_n|^2 + \frac{\alpha_n}{2\beta_n}|x_n - x_{n-1}|^2 + \frac{\alpha_n}{2\beta_n}|x_{n+1} - x_n|^2.\end{aligned}$$

We infer that

$$\begin{aligned}E_{n+1} - E_n &= \frac{1}{2\beta_n}|x_{n+1} - x_n|^2 - \frac{1}{2\beta_{n-1}}|x_n - x_{n-1}|^2 + \Phi(x_{n+1}) - \Phi(x_n) \\ &\leq -\frac{1 - \alpha_n}{2\,\beta_n}|x_{n+1} - x_n|^2 + \frac{1}{2}\left(\frac{\alpha_n}{\beta_n} - \frac{1}{\beta_{n-1}}\right)|x_n - x_{n-1}|^2.\end{aligned}$$

Inequality (4) is then a consequence of $\alpha_n \leq \frac{\beta_n}{\beta_{n-1}}$ for every $n \geq 1$.
(i) From the assumption $\alpha_n \leq 1$ for every $n \geq 1$, the sequence (E_n) is nonincreasing. Since (E_n) is minorized by $\inf \Phi$, it is convergent.
(ii) For every $n \geq 1$, we have $E_n \leq E_1$, hence

$$\frac{1}{2\beta_{n-1}}|x_n - x_{n-1}|^2 \leq E_1 - \inf \Phi,$$

and the conclusion immediately follows.
(iii) By summing inequality (4) from $n = 1$ to N, we obtain $\sum_{n=1}^{N} \frac{1-\alpha_n}{2\,\beta_n}|x_{n+1} - x_n|^2 \leq E_1 - E_{N+1} \leq E_1 - \inf \Phi$, which allows to conclude.

Example 2.1 Assume that $\alpha_n = \frac{n^b}{(n+1)^b}$ and $\beta_n = \frac{n^a}{(n+1)^b}$ for every $n \in \mathbb{N}$. It is immediate to check that the assumption $\alpha_n \leq 1$ is equivalent to $b \geq 0$ while the assumption $\alpha_n \leq \frac{\beta_n}{\beta_{n-1}}$ is equivalent to $a \geq 0$. If $b - a > 2$, we have $\left(\sqrt{\beta_n}\right) \in l^1$ and we deduce from Proposition 2.1(ii) that the corresponding sequence (x_n) converges (not in $\mathrm{Argmin}\Phi$ in general, see Remark 2.2 below).

Remark 2.2 If $\left(\sqrt{\beta_n}\right) \in l^1$, the sequence of discrete times $t_n = \sum_{k=0}^{n} \sqrt{\beta_k}$ tends toward $t_\infty < +\infty$. This implies that the asymptotic behavior of $(\mathcal{A})$ as $n \to +\infty$ is not related to the one of the continuous system (3) as $t \to +\infty$. As a consequence, the minimization process of Φ does not hold and in general, the limit point $\overline{x} = \lim_{n\to+\infty} x_n$ is not a minimum point of Φ.

Remark 2.3 In order to deal with numerical applications, it is convenient to authorize at each iteration n an error ε_n in the evaluation of the subdifferential. More precisely, denoting by ∂_ε the ε-approximate subdifferential, we are led to the following algorithm:

$$(\mathcal{A}_\varepsilon) \qquad x_{n+1} - x_n - \alpha_n(x_n - x_{n-1}) + \beta_n\, \partial_{\varepsilon_n}\Phi(x_{n+1}) \ni 0.$$

The sequence (ε_n) of errors is assumed to be summable so as to remain close to the exact subdifferential. If one uses algorithm $(\mathcal{A}_\varepsilon)$ instead of $(\mathcal{A})$, one has to add the quantity ε_n in the right-hand side of inequality (4). The sequence (E_n) is not necessarily nonincreasing, but it is still convergent. The other conclusions of Proposition 2.1 are unchanged.

3 Summability of the energy. Case of a unique minimum

We now show that the sequence $\Big((1-\alpha_n)(E_{n+1}-\min\Phi)\Big)$ is summable. This property implies the convergence of the sequence (E_n) toward $\min\Phi$ provided that the sequence $(1-\alpha_n)$ itself is not summable.

Theorem 3.1 *Let $\Phi : H \to \mathbb{R}\cup\{+\infty\}$ be a closed convex function such that Argmin$\Phi \neq \emptyset$. Let (α_n) be a positive nondecreasing sequence such that $\alpha_n \le 1$ for every $n\in\mathbb{N}$. Let (β_n) be a positive sequence such that $\alpha_n \le \frac{\beta_n}{\beta_{n-1}}$ for every $n \ge 1$. Assume that the sequence $\left(\frac{1-\alpha_n}{\sqrt{\beta_n}}\right)$ is bounded[5]. Defining the sequence (θ_n) by $\theta_n = \frac{1-\alpha_n}{\beta_n}$, suppose that $(\alpha_{n+1}\theta_{n+1}-(1+\alpha_n)\theta_n+\theta_{n-1}) \in l^1$. Then any bounded[6] sequence (x_n) generated by algorithm $(\mathcal{A})$ satisfies*

$$\Big((1-\alpha_n)(E_{n+1}-\min\Phi)\Big) \in l^1.$$

If additionaly $(1-\alpha_n) \notin l^1$, then $\lim_{n\to+\infty} E_n = \min\Phi$. As a consequence, $\lim_{n\to+\infty} \frac{1}{\beta_{n-1}}|x_n-x_{n-1}|^2 = 0$ and $\lim_{n\to+\infty}\Phi(x_n) = \min\Phi$.

Proof. Without loss of generality, we can assume that $\min\Phi = 0$. Given $z \in$ ArgminΦ, let us set $\varphi_n = \frac{1}{2}|x_n - z|^2$. We have for every $n\in\mathbb{N}$

$$\varphi_{n+1}-\varphi_n = \langle x_{n+1}-x_n, x_{n+1}-z\rangle - \frac{1}{2}|x_{n+1}-x_n|^2. \tag{5}$$

Set $\psi_n = \varphi_{n+1}-\varphi_n-\alpha_n(\varphi_n-\varphi_{n-1})$ and let $\xi_{n+1}\in\partial\Phi(x_{n+1})$ be such that $x_{n+1}-x_n-\alpha_n(x_n-x_{n-1})+\beta_n\,\xi_{n+1}=0$. We then have

$$\begin{aligned}\psi_n &= \langle x_{n+1}-x_n-\alpha_n(x_n-x_{n-1}), x_{n+1}-z\rangle + \alpha_n\langle x_n-x_{n-1}, x_{n+1}-x_n\rangle\\ &\quad -\frac{1}{2}|x_{n+1}-x_n|^2 + \frac{\alpha_n}{2}|x_n-x_{n-1}|^2\\ &\le -\beta_n\langle\xi_{n+1}, x_{n+1}-z\rangle + \alpha_n|x_n-x_{n-1}|^2 + \frac{\alpha_n-1}{2}|x_{n+1}-x_n|^2.\end{aligned}$$

Since $\xi_{n+1}\in\partial\Phi(x_{n+1})$ and $\Phi(z)=0$, we have

[5] Recall that from Remark 2.1 the term $\frac{1-\alpha_n}{\sqrt{\beta_n}}$ can be interpreted as a damping coefficient.

[6] Notice that the sequence (x_n) is automatically bounded if the function Φ is coercive, *i.e.* $\Phi(\xi)\to+\infty$ as $|\xi|\to+\infty$.

$$\langle \xi_{n+1}, x_{n+1} - z \rangle \geq \Phi(x_{n+1}) = E_{n+1} - \frac{1}{2\beta_n}|x_{n+1} - x_n|^2.$$

Hence we deduce that

$$\psi_n + \beta_n E_{n+1} \leq \alpha_n |x_n - x_{n-1}|^2 + \frac{\alpha_n}{2}|x_{n+1} - x_n|^2.$$

Let us multiply this inequality by $\frac{1-\alpha_n}{\beta_n}$. Since $\alpha_{n-1} \leq \alpha_n \leq 1$ and $\alpha_n \leq \frac{\beta_n}{\beta_{n-1}}$ for every $n \geq 1$, we derive that

$$\frac{1-\alpha_n}{\beta_n}\psi_n + (1-\alpha_n)\, E_{n+1} \leq \frac{1-\alpha_{n-1}}{\beta_{n-1}}\, |x_n - x_{n-1}|^2 + \frac{1-\alpha_n}{2\,\beta_n}\, |x_{n+1} - x_n|^2. \quad (6)$$

Let us set $\theta_n = \frac{1-\alpha_n}{\beta_n}$ and sum these inequalities from $n = 1$ to N. In view of Proposition 2.1 (iii), we have for every $N \in \mathbb{N}$

$$\sum_{n=1}^{N} \theta_n \psi_n + \sum_{n=1}^{N} (1-\alpha_n) E_{n+1} \leq \sum_{n=1}^{+\infty} \theta_{n-1} |x_n - x_{n-1}|^2 + \frac{1}{2} \sum_{n=1}^{+\infty} \theta_n |x_{n+1} - x_n|^2 < +\infty.$$

It suffices now to prove that the sequence $(\sum_{n=1}^{N} \theta_n \psi_n)$ is bounded with respect to N. Setting $\omega_n = \alpha_{n+1}\theta_{n+1} - (1+\alpha_n)\theta_n + \theta_{n-1}$ and using a technique of Abel transformation, we find

$$\sum_{n=1}^{N} \theta_n \psi_n = \sum_{n=1}^{N} \omega_n \varphi_n + (\varphi_{N+1}\theta_N - \varphi_N \alpha_{N+1}\theta_{N+1}) - \varphi_1\theta_0 + \varphi_0\alpha_1\theta_1. \quad (7)$$

Since the sequence (x_n) is bounded, the sequence (φ_n) is also bounded, say by $\overline{\varphi} > 0$. Since $(\omega_n) \in l^1$ by assumption, we deduce that

$$\left| \sum_{n=1}^{N} \omega_n \varphi_n \right| \leq \overline{\varphi} \sum_{n=1}^{+\infty} |\omega_n| < +\infty. \quad (8)$$

Now observe that

$$\varphi_{N+1}\theta_N - \varphi_N \alpha_{N+1}\theta_{N+1} = (\varphi_{N+1} - \varphi_N)\theta_N + \varphi_N(\theta_N - \alpha_{N+1}\theta_{N+1}). \quad (9)$$

The summability of (ω_n) shows that $\lim_{n\to+\infty} \alpha_{n+1}\theta_{n+1} - \theta_n$ exists. We deduce that

$$\varphi_N(\theta_N - \alpha_{N+1}\theta_{N+1}) \text{ is bounded with respect to } N. \quad (10)$$

Coming back to equality (5) and using the boundedness of the sequence (x_n), we derive the existence of $A > 0$ such that $|\varphi_{N+1} - \varphi_N| \leq A|x_{N+1} - x_N|$ for every $N \geq 0$. Recalling from Proposition 2.1 (ii) that $|x_{N+1} - x_N| \leq C\sqrt{\beta_N}$, we obtain for every $N \geq 0$ that

$$|\varphi_{N+1} - \varphi_N|\theta_N \leq A\,C\sqrt{\beta_N}\theta_N = A\,C\,\frac{1-\alpha_N}{\sqrt{\beta_N}}.$$

By assumption, the sequence $\left(\frac{1-\alpha_n}{\sqrt{\beta_n}}\right)$ is majorized, hence we infer that

$$(\varphi_{N+1} - \varphi_N)\theta_N \text{ is bounded with respect to } N. \tag{11}$$

By combining (7), (8), (9), (10) and (11), we conclude that the quantity $\sum_{n=1}^{N} \theta_n \psi_n$ is bounded with respect to N, which ends the proof of the summability of the sequence $\big((1-\alpha_n)\,E_{n+1}\big)$.
Let us now assume that $(1-\alpha_n) \notin l^1$. If $E_\infty = \lim_{n\to+\infty} E_n > 0$, then $((1-\alpha_n)E_{n+1}) \notin l^1$. Hence the limit E_∞ equals zero. The other assertions are immediate.

Example 3.1 Let $a,\, b \geq 0$ and assume that $\alpha_n = \frac{n^b}{(n+1)^b}$ and $\beta_n = \frac{n^a}{(n+1)^b}$ for every $n \in \mathbb{N}$. It is immediate to check that

$$\alpha_n = 1 - \frac{b}{n} + O(\frac{1}{n^2}), \qquad \alpha_{n+1} - \alpha_n = \frac{b}{n^2} + O(\frac{1}{n^3}) \qquad \text{as } n \to +\infty. \tag{12}$$

We have $\frac{1-\alpha_n}{\sqrt{\beta_n}} \sim b\,n^{\frac{b-a-2}{2}}$ as $n \to +\infty$. Hence the sequence $\frac{1-\alpha_n}{\sqrt{\beta_n}}$ is bounded if and only if $b-a \leq 2$. On the other hand, by setting $\omega_n = \alpha_{n+1}\theta_{n+1} - (1+\alpha_n)\theta_n + \theta_{n-1}$ as in the previous proof, we have

$$\omega_n = (\theta_{n+1} - 2\theta_n + \theta_{n-1}) - (1-\alpha_{n+1})(\theta_{n+1} - \theta_n) + (\alpha_{n+1} - \alpha_n)\theta_n. \tag{13}$$

An easy computation allows to find the following asymptotic expansions as $n \to +\infty$

$$\theta_n = b\,n^{b-a-1} + O(n^{b-a-2}), \tag{14}$$
$$\theta_{n+1} - \theta_n = b\,(b-a-1)n^{b-a-2} + O(n^{b-a-3}), \tag{15}$$
$$\theta_{n+1} - 2\theta_n + \theta_{n-1} = b\,(b-a-1)(b-a-2)n^{b-a-3} + O(n^{b-a-4}). \tag{16}$$

By combining the asymptotic expansions (12) and (14)-(16), we find in view of equality (13)

$$\omega_n = -b\,(b-a-2)(a+1)n^{b-a-3} + O(n^{b-a-4}) \qquad \text{as } n \to +\infty.$$

This sequence is clearly summable if $b-a \leq 2$. We conclude that the assumptions of Theorem 3.1 are satisfied if $a,\, b \geq 0$ and $b-a \leq 2$. If moreover $b > 0$, we have $(1-\alpha_n) \notin l^1$, hence we deduce from Theorem 3.1 that $\lim_{n\to+\infty} E_n = \min \Phi$.

Remark 3.1 Consider the approximate algorithm $(\mathcal{A}_\varepsilon)$ defined in Remark 2.3. The arguments developed in the proof of Theorem 3.1 are still valid for $(\mathcal{A}_\varepsilon)$, we simply have to add the term $(1-\alpha_n)\,\varepsilon_n$ in the right member of inequality (6). Since $(1-\alpha_n)\,\varepsilon_n \leq \varepsilon_n$ and since $(\varepsilon_n) \in l^1$ by assumption, the rest of the proof is the same and the conclusions of the theorem are identical for algorithm $(\mathcal{A}_\varepsilon)$.

We are now able to investigate the question of the convergence of the sequence (x_n) in the case of a unique minimum.

Corollary 3.1 *Under the hypotheses of Theorem 3.1 together with condition* $(1-\alpha_n) \notin l^1$*, assume that* $\mathrm{Argmin}\Phi = \{\overline{x}\}$ *for some* $\overline{x} \in H$*. Then any bounded sequence* (x_n) *generated by* $(\mathcal{A})$ *weakly converges to* $\overline{x}$ *in* H*.*

Proof. Since the sequence (x_n) is bounded, there exist $x_\infty \in H$ and a subsequence (x_{n_k}) of (x_n) such that $\lim_{k\to+\infty} x_{n_k} = x_\infty$ weakly in H. Since Φ is convex and closed for the strong topology, it is closed for the weak topology. Hence, we have $\Phi(x_\infty) \leq \liminf_{k\to+\infty} \Phi(x_{n_k})$. On the other hand, by applying Theorem 3.1, we obtain $\lim_{n\to+\infty} \Phi(x_n) = \min \Phi$. Therefore we deduce that $\Phi(x_\infty) \leq \min \Phi$, *i.e.* $x_\infty \in \mathrm{Argmin}\Phi = \{\overline{x}\}$. Hence $\overline{x}$ is the unique limit point of the sequence (x_n) for the weak topology. It ensues that $\lim_{n\to+\infty} x_n = \overline{x}$ weakly in H.

We say that $\overline{x} \in H$ is a strong minimum for Φ if for every $x \in H$, $\Phi(x) \geq \Phi(\overline{x}) + \delta(|x - \overline{x}|)$, where the map $\delta : \mathbb{R}_+ \to \mathbb{R}_+$ is such that $\delta(t_n) \to 0$ implies $t_n \to 0$ for every sequence $(t_n) \subset \mathbb{R}_+$.

Corollary 3.2 *Under the hypotheses of Theorem 3.1 together with condition* $(1-\alpha_n) \notin l^1$*, assume that* $\overline{x}$ *is a strong minimum for* Φ*. Then any bounded sequence* (x_n) *generated by* $(\mathcal{A})$ *strongly converges to* $\overline{x}$ *in* H*.*

Proof. By applying Theorem 3.1, we obtain $\lim_{n\to+\infty} \Phi(x_n) = \min \Phi = \Phi(\overline{x})$. Since $\overline{x}$ is a strong minimum for Φ, we deduce that $\lim_{n\to+\infty} \delta(|x_n - \overline{x}|) = 0$ and we conclude that $\lim_{n\to+\infty} |x_n - \overline{x}| = 0$.

Remark 3.2 *Condition* $(1-\alpha_n) \notin l^1$ *is equivalent to* $\lim_{n\to+\infty} \prod_{k=1}^n \alpha_k = 0$*. In the case of functions having a unique minimum* $\overline{x}$*, this condition is sufficient to obtain the (weak) convergence of the iterates* x_n *toward* $\overline{x}$*. It will be shown in the next section that the more stringent condition* $(\prod_{k=1}^n \alpha_k) \in l^1$ *is required to ensure the convergence of the sequence* (x_n) *for potentials* Φ *with multiple minima.*

4 The problem of convergence of algorithm $(\mathcal{A})$ for potentials with multiple minima

We are going to investigate the question of convergence of the sequences associated to $(\mathcal{A})$ when the convex potential Φ has multiple minima. Let us first consider the particular case $\Phi \equiv 0$. Algorithm $(\mathcal{A})$ then becomes $x_{n+1} - x_n - \alpha_n(x_n - x_{n-1}) = 0$ and an immediate computation shows that

$$x_{N+1} = x_1 + \left(\sum_{n=1}^{N} \prod_{k=1}^{n} \alpha_k\right)(x_1 - x_0).$$

It ensues that, when $\Phi \equiv 0$, the sequence (x_n) converges if and only if the quantity $\sum_{n=1}^{+\infty} \prod_{k=1}^{n} \alpha_k$ is finite. Therefore it is natural to ask whether for a general potential Φ, the sequence (x_n) is convergent under the condition $(\prod_{k=1}^{n} \alpha_k) \in l^1$. This question in its full generality is difficult and still open. The purpose of this section is to show that, conversely if $(\prod_{k=1}^{n} \alpha_k) \notin l^1$ then non stationary sequences cannot converge. Let us give a preliminary result that emphasizes the role of condition $(\prod_{k=1}^{n} \alpha_k) \notin l^1$.

Lemma 4.1 *Let (α_n) be a nonnegative sequence such that $(\prod_{k=1}^{n} \alpha_k) \notin l^1$.*
(i) Suppose that a sequence $(p_n) \subset \mathbb{R}$ satisfies

$$\forall n \geq n_0, \quad p_{n+1} - p_n - \alpha_n(p_n - p_{n-1}) \leq 0.$$

Then, we have either $\lim_{n\to+\infty} p_n = -\infty$ or $p_n \geq p_{n-1}$ for every $n \geq n_0$.
(ii) Suppose now that a sequence $(x_n) \subset H$ satisfies

$$\forall n \geq n_0, \quad x_{n+1} - x_n - \alpha_n(x_n - x_{n-1}) = 0.$$

Then, either $\lim_{n\to+\infty} |x_n| = +\infty$ or $x_n = x_{n_0}$ for every $n \geq n_0$.

Proof. (i) Assume that there exists $n_1 \geq n_0$ such that $p_{n_1} < p_{n_1-1}$. Then we have

$$\forall n \geq n_1, \qquad p_{n+1} - p_n \leq \left(\prod_{k=n_1}^{n} \alpha_k\right)(p_{n_1} - p_{n_1-1}).$$

By summing from $n = n_1$ to N, we find

$$p_{N+1} - p_{n_1} \leq \sum_{n=n_1}^{N}\left(\prod_{k=n_1}^{n} \alpha_k\right)(p_{n_1} - p_{n_1-1}).$$

Since $p_{n_1} < p_{n_1-1}$ and since $(\prod_{k=1}^{n} \alpha_k) \notin l^1$, we conclude that $\lim_{N\to+\infty} p_N = -\infty$.

(ii) Assume now that there exists $n_1 \geq n_0$ such that $x_{n_1} \neq x_{n_1-1}$. The same computation as above shows that

$$\forall n \geq n_1, \quad x_{N+1} - x_{n_1} = \sum_{n=n_1}^{N}\left(\prod_{k=n_1}^{n} \alpha_k\right)(x_{n_1} - x_{n_1-1}).$$

Since $x_{n_1} \neq x_{n_1-1}$ and since $(\prod_{k=1}^{n} \alpha_k) \notin l^1$, we conclude that $\lim_{N\to+\infty} |x_N| = +\infty$.

Given a closed convex set $S \subset H$ and $\overline{x} \in S$, recall that the normal cone $N_S(\overline{x})$ and the tangent cone $T_S(\overline{x})$ are respectively defined by

$$N_S(\overline{x}) = \{\xi \in H | \quad \forall x \in S, \quad \langle \xi, x - \overline{x}\rangle \leq 0\}$$
$$T_S(\overline{x}) = \mathrm{cl}\left[\cup_{\lambda>0} \lambda\,(S - \overline{x})\right].$$

The polar cone K^* of a cone $K \subset H$ is defined by

$$K^* = \{y \in H| \quad \forall x \in K, \quad \langle x, y\rangle \leq 0\}.$$

The convex cones $N_S(\overline{x})$ and $T_S(\overline{x})$ are polar to each other, *i.e.* $N_S(\overline{x}) = [T_S(\overline{x})]^*$ and $T_S(\overline{x}) = [N_S(\overline{x})]^*$. A cone K is said to be pointed if $K \cap -K = \{0\}$. For further details relative to convex analysis, the reader is referred to classical textbooks [21, 22]. In the sequel, the notation $\mathbb{B}$ (resp. $\mathbb{S}$) stands for the closed unit ball (resp. sphere) of H. Before stating the result of non-convergence for algorithm $(\mathcal{A})$, we need the following lemma.

Lemma 4.2 *Assume that* $\dim H < +\infty$. *Let* $\Phi : H \to \mathbb{R} \cup \{+\infty\}$ *be a closed convex function and let* $\overline{x} \in S = Argmin\Phi$. *Assume that*

$$-N_S(\overline{x}) \subset \mathrm{int}\,(T_S(\overline{x})) \cup \{0\}. \tag{17}$$

Then there exist a scalar $\lambda > 0$, *a convex cone* $K \subset H$ *which is closed and pointed, along with a neighborhood* V *of* $\overline{x}$ *such that*

$$K \cap \mathbb{B} \subset \lambda\,(\mathrm{int}\,(S) - \overline{x}) \cup \{0\} \quad \text{and} \quad -\partial\Phi(x) \subset K \quad \text{for every } x \in V. \tag{18}$$

Proof. If $\overline{x} \in \mathrm{int}\,(S)$, there exists a neighborhood V of $\overline{x}$ such that condition (18) is satisfied with $K = \{0\}$ and any $\lambda > 0$. Now assume that $\overline{x} \in \mathrm{bd}\,(S)$. Let us define the set K by

$$K = \{x \in H, \quad d(x, -N_S(\overline{x})) \leq d(x, H \setminus T_S(\overline{x}))\}.$$

It is immediate to check that the set K is a closed cone satisfying

$$K \subset \mathrm{int}\,(T_S(\overline{x})) \cup \{0\} \tag{19}$$

and

$$-N_S(\overline{x}) \setminus \{0\} \subset \mathrm{int}\,(K). \tag{20}$$

Since $\overline{x} \in \mathrm{bd}\,(S)$, there exists $u \in H \setminus \{0\}$ such that $\mathbb{R}_+ u \subset N_S(\overline{x})$. By polarity, we have $T_S(\overline{x}) \subset \{x \in H, \langle x, u\rangle \leq 0\}$, hence

$$K \subset \{x \in H, \quad \langle x, u\rangle < 0\} \cup \{0\}.$$

It ensues that the cone K is pointed. To prove the convexity of the set K, we resort to the following claim.

Claim 4.1 *Let* $C \subset H$ *be a nonempty convex set. Then we have:*

(i) The function $d(., C)$ *is convex on* H.
(ii)If $C \neq H$, *the function* $d(., H \setminus C)$ *is concave on* C.

The first point is elementary. The second one is given as an exercise by N. Bourbaki [5, Exercise 18, p. 150], see[7] also [12]. We deduce from this claim that the function $\Delta = d(., -N_S(\overline{x})) - d(., H \setminus T_S(\overline{x}))$ is convex on $T_S(\overline{x})$. In view of formula (19), we have $K \subset T_S(\overline{x})$ and we infer that the set $K = \{x \in T_S(\overline{x}), \quad \Delta(x) \leq 0\}$ is convex as a sublevel set of the convex function Δ. Using again inclusion (19) and recalling that $\operatorname{int}(T_S(\overline{x})) = \cup_{\lambda>0}\lambda\,(\operatorname{int}(S) - \overline{x})$, we obtain

$$K \cap \mathbb{S} \subset \cup_{\lambda>0}\lambda\,(\operatorname{int}(S) - \overline{x}).$$

From the compactness property of $K \cap \mathbb{S}$, we can extract a finite cover of $K \cap \mathbb{S}$: there exist $\lambda_1, \ldots, \lambda_n > 0$ such that

$$K \cap \mathbb{S} \subset \cup_{i=1}^{n}\lambda_i\,(\operatorname{int}(S) - \overline{x}). \tag{21}$$

Setting $\overline{\lambda} = \max\{\lambda_1, \ldots, \lambda_n\}$, observe that $\lambda_i\,(S - \overline{x}) \subset \overline{\lambda}\,(S - \overline{x})$ for every $i \in \{1, \ldots, n\}$. Taking the interior of each member, we infer that $\cup_{i=1}^{n}\lambda_i\,(\operatorname{int}(S) - \overline{x}) \subset \overline{\lambda}\,(\operatorname{int}(S) - \overline{x})$, hence $K \cap \mathbb{S} \subset \overline{\lambda}\,(\operatorname{int}(S) - \overline{x})$ in view of (21). It ensues immediately that $K \cap \mathbb{B} \subset \overline{\lambda}\,(\operatorname{int}(S) - \overline{x}) \cup \{0\}$, which proves the first part of assertion (18).

Let us finally prove that there exists a neighborhood V of $\overline{x}$ such that $-\partial\Phi(x) \subset K$ for every $x \in V$. Let us argue by contradiction and assume that there exist a sequence (x_n) tending to $\overline{x}$ as $n \to +\infty$, along with a sequence (ξ_n) such that $\xi_n \in -\partial\Phi(x_n)$ and $\xi_n \in H \setminus K$. Since the sequence $(\xi_n/|\xi_n|)$ is bounded, it has a subsequence, still denoted by $(\xi_n/|\xi_n|)$ such that $\lim_{n\to+\infty} \xi_n/|\xi_n| = \overline{\xi}$, for some $\overline{\xi} \in H$. Recalling that K is a cone, we have $\xi_n/|\xi_n| \in H \setminus K$ for every $n \in \mathbb{N}$, hence

$$\overline{\xi} \in \operatorname{cl}(H \setminus K) = H \setminus \operatorname{int}(K). \tag{22}$$

Let us now fix $x \in S$. From the fact that $-\xi_n \in \partial\Phi(x_n)$, we infer that

$$\langle -\xi_n, x - x_n\rangle \leq \Phi(x) - \Phi(x_n) \leq 0.$$

Dividing by $|\xi_n|$ and taking the limit as $n \to +\infty$, we derive that $\langle -\overline{\xi}, x - \overline{x}\rangle \leq 0$. Since this is true for every $x \in S$, we deduce that $-\overline{\xi} \in N_S(\overline{x})$. Recalling that $\overline{\xi} \neq 0$, we obtain from inclusion (20) that $\overline{\xi} \in \operatorname{int}(K)$, which clearly contradicts (22).

A closed convex cone $K \subset H$ is said to be acute (resp. obtuse) if $K \subset -K^*$ (resp. $K \supset -K^*$). These notions are widely used in the field of optimization, see for example [4, 9, 10, 11, 14]. Condition (17) amounts to saying that the cone $N_S(\overline{x})$ is strictly acute or equivalently that the cone $T_S(\overline{x})$ is strictly obtuse. This condition is satisfied in particular if the set S is smooth[8] at $\overline{x} \in \operatorname{bd}(S)$. When $H = \mathbb{R}$, condition (17) is satisfied if and only if the interval $\operatorname{Argmin}\Phi$ is not a singleton.

[7] The first author is indebted to L. Thibault (U. Montpellier II) for suggesting references [5, 12].

[8] Recall that the set S is smooth at $\overline{x} \in \operatorname{bd}(S)$ if there exists $d \neq 0$ such that $N_S(\overline{x}) = \mathbb{R}_+ d$.

Let us now state the general result of non-convergence for the sequences associated to $(\mathcal{A})$ under the condition $(\prod_{k=1}^n \alpha_k) \notin l^1$.

Theorem 4.1 *Assume that* $\dim H < +\infty$. *Let* $\Phi : H \to \mathbb{R} \cup \{+\infty\}$ *be a closed convex function such that for every* $\overline{x} \in S = Argmin\Phi$,

$$-N_S(\overline{x}) \subset \operatorname{int}(T_S(\overline{x})) \cup \{0\}.$$

Let (α_n), (β_n) *be nonnegative sequences and assume that* $(\prod_{k=1}^n \alpha_k) \notin l^1$. *If the sequence* (x_n) *defined by algorithm* $(\mathcal{A})$ *is non stationary*[9] *then it cannot converge toward* $\overline{x} \in S$.

Proof. Let us prove the contraposition of the previous statement and assume that there exists $\overline{x} \in S$ such that $\lim_{n\to+\infty} x_n = \overline{x}$. We must prove that the sequence (x_n) is stationary. In view of Lemma 4.2, there exist a convex cone $K \subset H$ which is closed and pointed, along with $\lambda > 0$ and $n_0 \geq 0$ such that

$$K \cap \frac{1}{\lambda}\mathbb{B} \subset (\operatorname{int}(S) - \overline{x}) \cup \{0\} \quad \text{and} \quad -\partial\Phi(x_n) \subset K \quad \text{for every } n \geq n_0. \tag{23}$$

Let $v \in K^*$. Observing that for every $n \geq n_0$

$$x_{n+1} - x_n - \alpha_n (x_n - x_{n-1}) \in -\beta_n \partial\Phi(x_{n+1}) \subset K,$$

we deduce that

$$\forall n \geq n_0, \quad \langle x_{n+1} - x_n - \alpha_n (x_n - x_{n-1}), v \rangle \leq 0.$$

Let us apply Lemma 4.1(i) to the sequence (p_n) defined by $p_n = \langle x_n, v\rangle$. From the boundedness of the sequence (x_n), we infer that $\langle x_{n+1}, v\rangle \geq \langle x_n, v\rangle$ for every $n \geq n_0$. Since this is true for every $v \in K^*$, we derive that $x_n - x_{n+1} \in K^{**}$ for every $n \geq n_0$. Recalling that $K^{**} = K$ for every closed convex cone K, we conclude that

$$\forall n \geq n_0, \qquad x_n - x_{n+1} \in K. \tag{24}$$

The cone K is stable under addition and closure operation, hence we deduce by summation from n to $+\infty$ that

$$\forall n \geq n_0, \qquad x_n - \overline{x} \in K. \tag{25}$$

Since $\lim_{n\to+\infty} x_n = \overline{x}$, there exists $n_1 \geq n_0$ such that $x_n - \overline{x} \in \frac{1}{\lambda}\mathbb{B}$ for every $n \geq n_1$. In view of (23), we infer that $x_n \in \operatorname{int}(S) \cup \{\overline{x}\}$ for every $n \geq n_1$. Let us now distinguish the following two cases:

(a)For every $n \geq n_1$, we have $x_n \in \operatorname{int}(S)$.

[9] If the function Φ is differentiable, inclusion $(\mathcal{A})$ holds as an equation and the principle of backward uniqueness applies. Stationary sequences are then characterized by the initial conditions $x_0 \in S$ and $x_1 = x_0$.

(b)There exists $n_2 \geq n_1$ such that $x_{n_2} = \overline{x}$.

Case (a) We then have $\partial\Phi(x_n) = \{0\}$ for every $n \geq n_1$, so that algorithm $(\mathcal{A})$ becomes

$$\forall n \geq n_1, \quad x_{n+1} - x_n - \alpha_n (x_n - x_{n-1}) = 0.$$

From Lemma 4.1(ii) and the fact that the sequence (x_n) is bounded, we deduce that $x_n = x_{n_1}$ for every $n \geq n_1$.

Case (b) Since $x_{n_2} = \overline{x}$ by assumption, we derive from (24) and (25) that

$$x_{n_2+1} - x_{n_2} = x_{n_2+1} - \overline{x} \in (-K) \cap K.$$

Recalling that the cone K is pointed, we have $(-K) \cap K = \{0\}$, hence $x_{n_2+1} = \overline{x}$. An immediate recurrence then shows that $x_n = \overline{x}$ for every $n \geq n_2$.

As a conclusion, we have proved in both cases (a) and (b) that the sequence (x_n) is stationary, which ends the proof.

Example 4.1 Let Φ be as in the previous theorem. Assume that $\alpha_n = \frac{n^b}{(n+1)^b}$ for every $n \in \mathbb{N}$. Since $\prod_{k=1}^n \alpha_k = \frac{1}{(n+1)^b}$, condition $(\prod_{k=1}^n \alpha_k) \notin l^1$ is satisfied if $b \leq 1$. Hence, we deduce from the previous theorem that if $b \leq 1$, then the non stationary sequences of $(\mathcal{A})$ cannot converge in S.

In view of the previous theorem the iterates of $(\mathcal{A})$ cannot converge in S, but they may tend toward some $\overline{x} \in H \setminus S$. To prevent this eventuality, we now give sufficient conditions on (α_n), (β_n), ensuring that any converging sequence (x_n) generated by $(\mathcal{A})$ tends toward a minimum point of Φ.

Proposition 4.1 *Let $\Phi : H \to \mathbb{R}$ be a continuous convex function. Let (α_n), (β_n) be nonnegative sequences satisfying the following assumptions*

$$\begin{cases} (i) \quad |\alpha_n - \alpha_{n-1}| = O(\beta_n) \qquad \text{as } n \to +\infty \\ (ii) \quad (\beta_n) \notin l^1 \qquad \text{or} \qquad \left[(\beta_n) \in l^1,\ \left(\sum_{k=n}^{+\infty} \beta_k\right) \notin l^1,\ \lim_{n\to+\infty} \alpha_n = 1\right]. \end{cases}$$

Let (x_n) be a sequence defined by algorithm $(\mathcal{A})$ and assume that $\lim_{n\to+\infty} x_n = \overline{x}$. Then we have $\overline{x} \in Argmin\Phi$.

Since the proof is a little bit technical, we postpone it to the appendix.

Example 4.2 Assume that $\alpha_n = \frac{n^b}{(n+1)^b}$ and $\beta_n = \frac{n^a}{(n+1)^b}$ for every $n \in \mathbb{N}$. First of all, observe that $\alpha_n = 1 - \frac{b}{n} + O\left(\frac{1}{n^2}\right)$, hence $\alpha_{n+1} - \alpha_n = O\left(\frac{1}{n^2}\right)$ as $n \to +\infty$. It ensues that condition $|\alpha_n - \alpha_{n-1}| = O(\beta_n)$ is realized if $b - a \leq 2$. On the other hand, condition $(\beta_n) \notin l^1$ holds if $b - a \leq 1$, while condition $\left(\sum_{k=n}^{+\infty} \beta_k\right) \notin l^1$ holds if $b - a \in]1, 2]$. Therefore we deduce from the previous proposition that, if $b - a \leq 2$ then any converging sequence (x_n) generated by $(\mathcal{A})$ tends toward a minimum point of Φ.

Conclusion and perspectives. To end this paper, let us come back to the proximal-like iteration

$$x_{n+1} - x_n + \frac{1}{(n+1)^b} \sum_{i=0}^{n} i^a \, \nabla\Phi(x_{i+1}) = 0, \tag{26}$$

where a, $b \geq 0$ and $\Phi : H \to \mathbb{R}$ is a differentiable convex function. As explained in the introduction, an elementary transformation of the above iteration leads to algorithm $(\mathcal{A})$ associated with the coefficients $\alpha_n = \frac{n^b}{(n+1)^b}$, $\beta_n = \frac{n^a}{(n+1)^b}$ for every $n \in \mathbb{N}$. Let us now list our results of convergence for algorithm (26). First of all, the energy sequence (E_n) defined by $E_n = \frac{n^b}{2\,(n-1)^a}|x_n - x_{n-1}|^2 + \Phi(x_n)$ is nonincreasing. Let us distinguish the cases $b - a > 2$ and $b - a \leq 2$.

- If $b - a > 2$, we derive from Example 2.1 that $\overline{x} = \lim_{n\to+\infty} x_n$ exists but $\overline{x}$ is not a minimum of Φ in general.
- If $b - a \leq 2$ and $b > 0$, Example 3.1 shows that $\lim_{n\to+\infty} E_n = \min\Phi$. If $\mathrm{Argmin}\Phi = \{\overline{x}\}$, this implies that $\lim_{n\to+\infty} x_n = \overline{x}$ weakly in H.
  Assume now that Φ has multiple minima.
  - If $b \leq 1$, we deduce from Examples 4.1 and 4.2 that (x_n) does not converge in general.
  - If $b > 1$, the problem of the convergence of (x_n) is open.

It would be interesting to replace the subgradient in algorithm $(\mathcal{A})$ by a maximal monotone operator. The main difficulty lies in the fact that no energy sequence is available in this framework. In view of numerical computations, another interesting problem would consist in studying an explicit version of $(\mathcal{A})$, namely with $\partial\Phi(x_n)$ in place of $\partial\Phi(x_{n+1})$. These remarks certainly indicate directions for future investigation.

Appendix: Proof of Proposition 4.1

Let us argue by contradiction and assume that $0 \notin \partial\Phi(\overline{x})$. It is then possible to strictly separate the convex compact set $\{0\}$ from the nonempty closed convex set $\partial\Phi(\overline{x})$. Therefore, there exist $p \in H$ and $m \in \mathbb{R}_+^*$ such that

$$\forall \xi \in \partial\Phi(\overline{x}), \quad \langle \xi, p\rangle > m. \tag{27}$$

Let us first prove that there exists $n_0 \in \mathbb{N}$ such that

$$\forall n \geq n_0, \, \forall \xi \in \partial\Phi(x_n), \quad \langle \xi, p\rangle > m. \tag{28}$$

If this was not true, there would exist a subsequence (x_{n_k}) of (x_n) along with a sequence (ξ_k) such that $\xi_k \in \partial\Phi(x_{n_k})$ and $\langle \xi_k, p\rangle \leq m$ for every $k \in \mathbb{N}$. Since the convex function Φ is continuous on H, the operator $\partial\Phi : H \rightrightarrows H$ is locally bounded, hence the sequence (ξ_k) is bounded. Therefore, there exist $\overline{\xi} \in H$ and a weakly converging subsequence of (ξ_k), still denoted by (ξ_k) such that w-$\lim_{k\to+\infty} \xi_k = \overline{\xi}$. On the other hand, by using the graph-closedness property of the operator $\partial\Phi$ in $H\times$ w-H, we find $\overline{\xi} \in \partial\Phi(\overline{x})$. By combining this property with the fact that $\langle\overline{\xi}, p\rangle \leq m$, we clearly obtain a contradiction with (27). Hence property (28) is proved. Without loss of generality, we will assume that $n_0 = 0$ in the sequel.

Case $(\beta_n) \notin l^1$. From the definition of algorithm $(\mathcal{A})$, for every $k \geq 1$, there exists $\xi_{k+1} \in \partial\Phi(x_{k+1})$ such that

$$x_{k+1} - x_k - \alpha_k(x_k - x_{k-1}) + \beta_k\xi_{k+1} = 0.$$

By summing from $k = 1$ to n and by using a technique of Abel transformation, we obtain

$$\sum_{k=1}^{n}(\alpha_{k+1} - \alpha_k)\, x_k + [x_{n+1} - \alpha_{n+1}\, x_n - x_1 + \alpha_1\, x_0] + \sum_{k=1}^{n} \beta_k\, \xi_{k+1} = 0$$

or equivalently

$$\sum_{k=1}^{n}(\alpha_{k+1} - \alpha_k)\,(x_k - x_n) + [x_{n+1} - \alpha_1\, x_n - x_1 + \alpha_1\, x_0] + \sum_{k=1}^{n} \beta_k\, \xi_{k+1} = 0. \quad (29)$$

Recalling that $|\alpha_n - \alpha_{n-1}| = O(\beta_n)$ as $n \to +\infty$, that $(\beta_n) \notin l^1$ and that $\lim_{n\to+\infty} x_n = \overline{x}$, we have

$$\sum_{k=1}^{n}(\alpha_{k+1} - \alpha_k)\,(x_k - x_n) = o\left(\sum_{k=1}^{n} \beta_k\right), \qquad \text{as } n \to +\infty. \quad (30)$$

On the other hand, from assertion (28) applied with $\xi = \xi_{k+1}$ for $k = 1, \ldots, n$, we derive that

$$\left\langle \sum_{k=1}^{n} \beta_k\, \xi_{k+1}, p \right\rangle \geq m \sum_{k=1}^{n} \beta_k. \quad (31)$$

Since the sequence (x_n) is bounded, the term between brackets in equality (29) is negligible with respect to $\sum_{k=1}^{n} \beta_k$ as $n \to +\infty$ and we obtain a contradiction in view of (30) and (31). As a conclusion, we have proved that $0 \in \partial\Phi(\overline{x})$ in the case $(\beta_n) \notin l^1$.

Case $(\beta_n) \in l^1$, $\left(\sum_{k=n}^{+\infty} \beta_k\right) \notin l^1$ and $\lim_{n\to+\infty} \alpha_n = 1$. By using the same technique of Abel transformation as above, we obtain

$$\sum_{k=n}^{+\infty}(\alpha_{k+1} - \alpha_k)\, x_k + [\alpha_n\, x_{n-1} - x_n] + \sum_{k=n}^{+\infty} \beta_k\, \xi_{k+1} = 0.$$

Observing that $1-\alpha_n=\sum_{k=n}^{+\infty}(\alpha_{k+1}-\alpha_k)$, this can be equivalently rewritten as

$$\sum_{k=n}^{+\infty}(\alpha_{k+1}-\alpha_k)\,(x_k-x_{n-1})+[x_{n-1}-x_n]+\sum_{k=n}^{+\infty}\beta_k\,\xi_{k+1}=0.$$

By summation from $n=1$ to N we obtain

$$\sum_{n=1}^{N}\sum_{k=n}^{+\infty}(\alpha_{k+1}-\alpha_k)\,(x_k-x_{n-1})+[x_0-x_N]+\sum_{n=1}^{N}\sum_{k=n}^{+\infty}\beta_k\,\xi_{k+1}=0. \tag{32}$$

Recalling that $|\alpha_n-\alpha_{n-1}|=O(\beta_n)$ as $n\to+\infty$, that $\left(\sum_{k=n}^{+\infty}\beta_k\right)\notin l^1$ and that $\lim_{n\to+\infty}x_n=\overline{x}$, we have

$$\sum_{n=1}^{N}\sum_{k=n}^{+\infty}(\alpha_{k+1}-\alpha_k)\,(x_k-x_{n-1})=o\left(\sum_{n=1}^{N}\sum_{k=n}^{+\infty}\beta_k\right),\qquad \text{as } N\to+\infty. \tag{33}$$

On the other hand, from assertion (28) applied with $\xi=\xi_{k+1}$ for $k=n,n+1,\ldots$ we derive that

$$\left\langle\sum_{n=1}^{N}\sum_{k=n}^{+\infty}\beta_k\,\xi_{k+1},p\right\rangle\geq m\sum_{n=1}^{N}\sum_{k=n}^{+\infty}\beta_k. \tag{34}$$

Since the sequence (x_n) is bounded, the term between brackets in equality (32) is negligible with respect to $\sum_{n=1}^{N}\sum_{k=n}^{+\infty}\beta_k$ as $N\to+\infty$ and we obtain a contradiction in view of (33) and (34). This achieves the proof of $0\in\partial\Phi(\overline{x})$ in the second case. □

Acknowledgments

The authors thank the anonymous referee for valuable comments and suggestions.

References

[1] F. Alvarez, On the minimizing property of a second order dissipative system in Hilbert spaces, *SIAM J. on Control and Optimization*, 38 (2000), n° 4, 1102-1119.

[2] F. Alvarez, H. Attouch, An inertial proximal method for maximal monotone operators via discretization of a nonlinear oscillator with damping, *Set-Valued Analysis*, vol. 9 (2001), 3-11.

[3] B. Baji, A. Cabot, An inertial proximal algorithm with dry friction: finite convergence results, *Set-Valued Anal.* 14 (2006), 1-23.

[4] H. H. Bauschke, S. G. Kruk, The method of reflection-projection for convex feasibility problems with an obtuse cone, *Journal of Optimization Theory and Applications*, 120 (2004), 503-531.

[5] N. Bourbaki, Espaces vectoriels topologiques, Fascicule XV, Hermann, Paris (1966).

[6] A. Cabot, Asymptotics for a gradient system with memory term, *Proc. of the Amer. Math. Soc.*, 137 (2009), 3013-3024.

[7] A. Cabot, H. Engler, S. Gadat, On the long time behavior of second order differential equations with asymptotically small dissipation, *Trans. of the Amer. Math. Soc.*, 361 (2009), 5983-6017.

[8] A. Cabot, H. Engler, S. Gadat, Second order differential equations with asymptotically small dissipation and piecewise flat potentials, *Electronic Journal of Differential Equations*, 17 (2009), 33-38.

[9] A. Cegielski, A method of projection onto an acute cone with level control in convex minimization, *Math. Program.*, 85 (1999), 469-490.

[10] A. Cegielski, Obtuse cones and Gram matrices with non-negative inverse, *Linear Algebra Appl.*, 335 (2001), 167-181.

[11] J. L. Goffin, The relaxation method for solving systems of linear inequalities, *Math. Oper. Res.*, 5 (1980), 388-414.

[12] J.-B. Hiriart-Urruty, New concepts in nondifferentiable programming, *Bull. Soc. Math. France*, 60 (1979), 57-85.

[13] F. Jules, P. E. Maingé, Numerical approach to a stationary solution of a second order dissipative dynamical system, *Optimization* 51 (2002), 235-255.

[14] K. C. Kiwiel, The efficiency of subgradient projection methods for convex optimization, part II: implementations and extensions, *SIAM J. Control Optim.*, 34 (1996), 677-697.

[15] B. Lemaire, About the convergence of the proximal method, Advances in optimization, *Lecture Notes in Econom. and Math. Systems*, 382, Springer, Berlin, (1992), 39-51.

[16] P. E. Maingé, A. Moudafi, Convergence of new inertial proximal methods for DC programming, *SIAM J. Optim.* 19 (2008), 397-413.

[17] B. Martinet, Régularisation d'inéquations variationnelles par approximations successives, *Revue Française d'Informatique et Recherche Opérationnelle*, 4 (1970), 154-159.

[18] B. Martinet, Détermination approchée d'un point fixe d'une application pseudo-contractante, *C.R. Acad. Sci.*, Paris, 274 (1972), 163-165.

[19] A. Moudafi, M. Oliny, Convergence of a splitting inertial proximal method for monotone operators, *J. Comput. Appl. Math.*, 155 (2003), 447-454.

[20] R.T. Rockafellar, Monotone operators and the proximal point algorithm, *SIAM Journal on Control and Optimization*, 14 (1976), pp. 877-898.

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

[22] R.T. Rockafellar and R. Wets, Variational analysis, Springer, Berlin (1998).

III. ALGORITHMES PROXIMAUX ALTERNES AVEC COÛTS-AUX-CHANGEMENTS, DESCRIPTION DUALE ET APPLICATION AUX EDP

Alternating proximal algorithms with costs-to-move, dual description and application to PDE's

P. Frankel[1]

Université Montpellier II, Place Eugène Bataillon, 34095 Montpellier Cedex 05, France
p.frankel30@orange.fr

Accepted in *DCDS-S vol. 5, no. 4*

Summary. Given real Hilbert spaces $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$, closed convex functions $f : \mathcal{X} \longrightarrow \mathbb{R} \cup \{+\infty\}$, $g : \mathcal{Y} \longrightarrow \mathbb{R} \cup \{+\infty\}$ and linear continuous operators $A : \mathcal{X} \longrightarrow \mathcal{Z}$, $B : \mathcal{Y} \longrightarrow \mathcal{Z}$, we study the following alternating proximal algorithm

$$(\mathcal{A}) \qquad \begin{cases} x_{n+1} = \text{Argmin}\left\{ f(\zeta) + \dfrac{1}{2\gamma}\|A\zeta - By_n\|^2_{\mathcal{Z}} + \dfrac{\alpha}{2}\|\zeta - x_n\|^2_{\mathcal{X}}; \quad \zeta \in \mathcal{X} \right\} \\ y_{n+1} = \text{Argmin}\left\{ g(\eta) + \dfrac{1}{2\gamma}\|Ax_{n+1} - B\eta\|^2_{\mathcal{Z}} + \dfrac{\nu}{2}\|\eta - y_n\|^2_{\mathcal{Y}}; \quad \eta \in \mathcal{Y} \right\}, \end{cases}$$

where γ, α and ν are positive parameters. Under suitable conditions, we prove that any sequence (x_n, y_n) generated by $(\mathcal{A})$ weakly converges toward a minimum point of the function $(x, y) \mapsto f(x) + g(y) + \frac{1}{2\gamma}\|Ax - By\|^2_{\mathcal{Z}}$ and that the sequence of dual variables $\left(-\frac{1}{\gamma}(Ax_n - By_n)\right)$ strongly converges in $\mathcal{Z}$ toward the unique minimizer of the function $z \mapsto f^*(A^*z) + g^*(-B^*z) + \frac{\gamma}{2}\|z\|^2_{\mathcal{Z}}$. An application is given in variational problems and PDE's.

Key words: Convex minimization, alternating minimization, proximal algorithm, domain decomposition for PDE's.
Subject classification: 65K05, 65K10, 49J40, 90C25.

1 Introduction

Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ be real Hilbert spaces. We note respectively $\langle ., . \rangle_{\mathcal{X}}$, $\langle ., . \rangle_{\mathcal{Y}}$ and $\langle ., . \rangle_{\mathcal{Z}}$ the scalar product of the spaces $\mathcal{X}$, $\mathcal{Y}$ or $\mathcal{Z}$, and $\|.\|_{\mathcal{X}}$, $\|.\|_{\mathcal{Y}}$, $\|.\|_{\mathcal{Z}}$ the corresponding norms. Given closed convex proper functions $f : \mathcal{X} \longrightarrow \mathbb{R} \cup \{+\infty\}$, $g : \mathcal{Y} \longrightarrow \mathbb{R} \cup \{+\infty\}$ and linear continuous operators $A : \mathcal{X} \longrightarrow \mathcal{Z}$, $B : \mathcal{Y} \longrightarrow \mathcal{Z}$, we consider the convex function $\Phi : \mathcal{X} \times \mathcal{Y} \longrightarrow \mathbb{R} \cup \{+\infty\}$ defined by

$$\Phi(x, y) = f(x) + g(y) + \frac{1}{2\gamma}\|Ax - By\|^2_{\mathcal{Z}},$$

where γ is a positive real parameter. We denote by $(\mathcal{P})$ the following minimization problem

$$(\mathcal{P}) \qquad \inf\{\Phi(x,y); \quad x \in \mathcal{X}, y \in \mathcal{Y}\}.$$

The weak coupling term $Q(x,y) = \|Ax - By\|^2_{\mathcal{Z}}$ allows asymmetric and partial relations between the variables x and y, contrary to the strong coupled problem

$$\left\{f(x) + g(y) + \frac{1}{2\gamma}\|x-y\|^2_{\mathcal{H}}; \quad x \in \mathcal{H}, y \in \mathcal{H}\right\}$$

where x and y lie in the same Hilbert space $\mathcal{H}$ and are involved in a symmetric way. We study the alternating algorithm with costs-to-move introduced by Attouch, Redont and Soubeyran [5]:

$$(\mathcal{A}) \quad \begin{cases} x_{n+1} = \operatorname{Argmin}\left\{f(\zeta) + \frac{1}{2\gamma}\|A\zeta - By_n\|^2_{\mathcal{Z}} + \frac{\alpha}{2}\|\zeta - x_n\|^2_{\mathcal{X}}; \quad \zeta \in \mathcal{X}\right\} \\ y_{n+1} = \operatorname{Argmin}\left\{g(\eta) + \frac{1}{2\gamma}\|Ax_{n+1} - B\eta\|^2_{\mathcal{Z}} + \frac{\nu}{2}\|\eta - y_n\|^2_{\mathcal{Y}}; \quad \eta \in \mathcal{Y}\right\}, \end{cases}$$

where α and ν are positive real numbers. This algorithm generates a sequence (x_n, y_n) whose convergence is studied in [2]. In references [1, 6], a particular case of algorithm $(\mathcal{A})$ with $\alpha = \nu = 0$ is studied for the strong coupled problem ($\mathcal{X} = \mathcal{Y}$ and $A = B = \mathcal{I}$). In this paper, we generalize some convergence results of [6] to the weak coupled problem $(\mathcal{P})$. More particularly, we prove that, if Φ is bounded from below, the sequence (x_n, y_n) is a minimizing sequence for Φ which slightly improves the corresponding convergence result of [2]. By a different way, we show that, if $\operatorname{Argmin}\Phi \neq \varnothing$, the sequence (x_n, y_n) weakly converges toward a minimum point of Φ. Moreover, a special attention is devoted to some dual problem $(\mathcal{P}^*)$ associated to problem $(\mathcal{P})$. We prove that the sequence of dual variables $\left(-\frac{1}{\gamma}(Ax_n - By_n)\right)$ strongly converges to the unique minimizer of problem $(\mathcal{P}^*)$. Attouch, Bolte, Redont and Soubeyran have given in [2] an application of algorithm $(\mathcal{A})$ to domain decomposition for PDE's. They have studied a minimization problem with Dirichlet boundary condition associated to problem $(\mathcal{P})$. Here we consider the corresponding problem with Neuman boundary condition.

The paper is organized as follows. We establish the convergence of algorithm $(\mathcal{A})$ in section 2. The sequence of dual variables $\left(-\frac{1}{\gamma}(Ax_n - By_n)\right)$ is studied in section 3. An application to PDE's is given in section 4.

2 Convergence of the algorithm

Let $f : \mathcal{X} \longrightarrow \mathbb{R} \cup \{+\infty\}$, $g : \mathcal{Y} \longrightarrow \mathbb{R} \cup \{+\infty\}$ be closed convex proper functions and let $A : \mathcal{X} \longrightarrow \mathcal{Z}$, $B : \mathcal{Y} \longrightarrow \mathcal{Z}$ be linear continuous operators. We consider the convex function $\Phi : \mathcal{X} \times \mathcal{Y} \longrightarrow \mathbb{R} \cup \{+\infty\}$ defined by

$$\Phi(x,y) = f(x) + g(y) + \frac{1}{2\gamma}\|Ax - By\|^2_{\mathcal{Z}}, \tag{1}$$

where γ is a positive real parameter. Given positive coefficients α, $\nu > 0$ and initial data $(x_0, y_0) \in \mathcal{X} \times \mathcal{Y}$, let us consider the following alternating proximal algorithm

$$(\mathcal{A}) \quad \begin{cases} x_{n+1} = \text{Argmin}\left\{ f(\zeta) + \frac{1}{2\gamma}\|A\zeta - By_n\|^2_{\mathcal{Z}} + \frac{\alpha}{2}\|\zeta - x_n\|^2_{\mathcal{X}}; \quad \zeta \in \mathcal{X} \right\} \\ y_{n+1} = \text{Argmin}\left\{ g(\eta) + \frac{1}{2\gamma}\|Ax_{n+1} - B\eta\|^2_{\mathcal{Z}} + \frac{\nu}{2}\|\eta - y_n\|^2_{\mathcal{Y}}; \quad \eta \in \mathcal{Y} \right\}. \end{cases}$$

By writing down the optimality conditions, it is immediate to check that points x_{n+1} and y_{n+1} are characterized by

$$\begin{cases} -\frac{1}{\gamma}A^*(Ax_{n+1} - By_n) - \alpha(x_{n+1} - x_n) \in \partial f(x_{n+1}) \\ \frac{1}{\gamma}B^*(Ax_{n+1} - By_{n+1}) - \nu(y_{n+1} - y_n) \in \partial g(y_{n+1}), \end{cases} \tag{2}$$

where $A^* \in \mathcal{L}(\mathcal{Z}, \mathcal{X})$ and $B^* \in \mathcal{L}(\mathcal{Z}, \mathcal{Y})$ denote the respective adjoint operators of A and B. It ensues that we have, for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$,

$$\begin{cases} f(x) - f(x_{n+1}) + \frac{1}{\gamma}\langle By_n - Ax_{n+1}, Ax_{n+1} - Ax\rangle_{\mathcal{Z}} - \alpha\langle x_{n+1} - x_n, x_{n+1} - x\rangle_{\mathcal{X}} \geq 0 \\ g(y) - g(y_{n+1}) + \frac{1}{\gamma}\langle Ax_{n+1} - By_{n+1}, By_{n+1} - By\rangle_{\mathcal{Z}} - \nu\langle y_{n+1} - y_n, y_{n+1} - y\rangle_{\mathcal{Y}} \geq 0. \end{cases} \tag{3}$$

These inequalities will be used intensively in the sequel. The next theorem states the main convergence properties of algorithm $(\mathcal{A})$.

Theorem 2.1 *Let α, ν and γ be positive coefficients and let $A \in \mathcal{L}(\mathcal{X}, \mathcal{Z})$, $B \in \mathcal{L}(\mathcal{Y}, \mathcal{Z})$ be linear continuous operators. Let $f : \mathcal{X} \longrightarrow \mathbb{R} \cup \{+\infty\}$ and $g : \mathcal{Y} \longrightarrow \mathbb{R} \cup \{+\infty\}$ be closed convex proper functions. Assume that the function Φ defined by equality (1) is bounded from below. If (x_n, y_n) is a sequence generated by $(\mathcal{A})$, then*

(i) $\forall n \in \mathbb{N}$, $\Phi(x_{n+1}, y_{n+1}) \leq \Phi(x_{n+1}, y_n) \leq \Phi(x_n, y_n)$;
(ii) $\lim_{n\to+\infty} \Phi(x_{n+1}, y_n) = \lim_{n\to+\infty} \Phi(x_n, y_n) = \inf \Phi$;
(iii) the sequences $(\|x_{n+1} - x_n\|^2_{\mathcal{X}})$ *and* $(\|y_{n+1} - y_n\|^2_{\mathcal{Y}})$ *are summable;*
(iv) if Argmin$\Phi \neq \varnothing$, then for all $(x, y) \in$ ArgminΦ,

(a) the sequences $(\|Ax - Ax_n\|^2_{\mathcal{Z}} + \gamma\alpha\|x_n - x\|^2_{\mathcal{X}} + \gamma\nu\|y_{n-1} - y\|^2_{\mathcal{Y}})$ and $(\|By - By_n\|^2_{\mathcal{Z}} + \gamma\alpha\|x_n - x\|^2_{\mathcal{X}} + \gamma\nu\|y_n - y\|^2_{\mathcal{Y}})$ are nonincreasing and convergent;

(b) the sequences $(\|(Ax - By) - (Ax_n - By_n)\|^2_{\mathcal{Z}})$, $(\|(Ax - By) - (Ax_{n+1} - By_n)\|^2_{\mathcal{Z}})$, $(\Phi(x_n, y_n) - \Phi(x, y))$ and $(\Phi(x_{n+1}, y_n) - \Phi(x, y))$ are summable;

(c) the sequence (x_n, y_n) weakly converges in $\mathcal{X} \times \mathcal{Y}$ toward a minimum point $(\overline{x}, \overline{y})$ of Φ. Moreover $f(x_n) \to f(\overline{x})$ and $g(y_n) \to g(\overline{y})$ as $n \to +\infty$;

(v)if $Argmin\Phi = \varnothing$*, then* $\|x_n\|_{\mathcal{X}} + \|y_n\|_{\mathcal{Y}} \to +\infty$ *as* $n \to +\infty$*.*

Proof. The arguments follow the same lines as those of Bauschke, Combettes and Reich [6].
(i) From the definition of algorithm $(\mathcal{A})$, we have

$$f(x_{n+1}) + \frac{1}{2\gamma}\|Ax_{n+1} - By_n\|^2_{\mathcal{Z}} + \frac{\alpha}{2}\|x_{n+1} - x_n\|^2_{\mathcal{X}} \leq f(x_n) + \frac{1}{2\gamma}\|Ax_n - By_n\|^2_{\mathcal{Z}},$$
$$g(y_{n+1}) + \frac{1}{2\gamma}\|Ax_{n+1} - By_{n+1}\|^2_{\mathcal{Z}} + \frac{\nu}{2}\|y_{n+1} - y_n\|^2_{\mathcal{Y}} \leq g(y_n) + \frac{1}{2\gamma}\|Ax_{n+1} - By_n\|^2_{\mathcal{Z}}.$$

We deduce that, for all $n \in \mathbb{N}$

$$\begin{aligned}\Phi(x_{n+1}, y_{n+1}) &= f(x_{n+1}) + g(y_{n+1}) + \frac{1}{2\gamma}\|Ax_{n+1} - By_{n+1}\|^2_{\mathcal{Z}} \\ &\leq f(x_{n+1}) + g(y_{n+1}) + \frac{1}{2\gamma}\|Ax_{n+1} - By_{n+1}\|^2_{\mathcal{Z}} + \frac{\nu}{2}\|y_{n+1} - y_n\|^2_{\mathcal{Y}} \\ &\leq f(x_{n+1}) + g(y_n) + \frac{1}{2\gamma}\|Ax_{n+1} - By_n\|^2_{\mathcal{Z}} = \Phi(x_{n+1}, y_n) \\ &\leq f(x_{n+1}) + g(y_n) + \frac{1}{2\gamma}\|Ax_{n+1} - By_n\|^2_{\mathcal{Z}} + \frac{\alpha}{2}\|x_{n+1} - x_n\|^2_{\mathcal{X}} \\ &\leq f(x_n) + g(y_n) + \frac{1}{2\gamma}\|Ax_n - By_n\|^2_{\mathcal{Z}} = \Phi(x_n, y_n),\end{aligned}$$

which ends the proof of (i).
(ii) The sequence $(\Phi(x_n, y_n))$ is nonincreasing and minorized hence convergent toward $\varphi \geq \inf \Phi$. By item (i), $(\Phi(x_{n+1}, y_n))$ converges toward φ too.
Let us use the following lemma borrowed from [6].

Lemma 2.1 *Let* $(s, t, u, v, w) \in \mathcal{Z}^5$*, then*

$$\|s-u\|^2_{\mathcal{Z}} = \|s-w\|^2_{\mathcal{Z}} + \|w-v\|^2_{\mathcal{Z}} - \|s-t\|^2_{\mathcal{Z}} + \|(s-t)-(u-v)\|^2_{\mathcal{Z}} + 2\langle s-w, w-v\rangle_{\mathcal{Z}} + 2\langle u-v, v-t\rangle_{\mathcal{Z}}.$$

Taking $s = Ax,\ t = By,\ u = Ax_n,\ v = By_n,\ w = Ax_{n+1}$, we obtain

$$\begin{aligned}&\|Ax - Ax_n\|^2_{\mathcal{Z}} - \|Ax - Ax_{n+1}\|^2_{\mathcal{Z}} \\ &\quad = \|Ax_{n+1} - By_n\|^2_{\mathcal{Z}} - \|Ax - By\|^2_{\mathcal{Z}} + \|(Ax - By) - (Ax_n - By_n)\|^2_{\mathcal{Z}} \\ &\qquad + 2\langle Ax - Ax_{n+1}, Ax_{n+1} - By_n\rangle_{\mathcal{Z}} + 2\langle Ax_n - By_n, By_n - By\rangle_{\mathcal{Z}}.\end{aligned} \tag{4}$$

In view of inequalities (3), we have

$$\begin{aligned}
&\|Ax - Ax_n\|_{\mathcal{Z}}^2 - \|Ax - Ax_{n+1}\|_{\mathcal{Z}}^2 \\
&\geq \|Ax_{n+1} - By_n\|_{\mathcal{Z}}^2 - \|Ax - By\|_{\mathcal{Z}}^2 + \|(Ax - By) - (Ax_n - By_n)\|_{\mathcal{Z}}^2 \\
&\quad + 2\left\{\gamma(f(x_{n+1}) - f(x)) + \gamma\alpha\langle x_{n+1} - x_n, x_{n+1} - x\rangle_{\mathcal{X}}\right\} \\
&\quad + 2\left\{\gamma(g(y_n) - g(y)) + \gamma\nu\langle y_n - y_{n-1}, y_n - y\rangle_{\mathcal{Y}}\right\} \\
&= 2\gamma\left\{f(x_{n+1}) + g(y_n) + \frac{1}{2\gamma}\|Ax_{n+1} - By_n\|_{\mathcal{Z}}^2 - f(x) - g(y) - \frac{1}{2\gamma}\|Ax - By\|_{\mathcal{Z}}^2\right\} \\
&\quad + \|(Ax - By) - (Ax_n - By_n)\|_{\mathcal{Z}}^2 \\
&\quad + 2\gamma\alpha\langle x_{n+1} - x_n, x_{n+1} - x\rangle_{\mathcal{X}} + 2\gamma\nu\langle y_n - y_{n-1}, y_n - y\rangle_{\mathcal{Y}} \\
&= 2\gamma\left\{\Phi(x_{n+1}, y_n) - \Phi(x, y)\right\} + \|(Ax - By) - (Ax_n - By_n)\|_{\mathcal{Z}}^2 \\
&\quad + 2\gamma\alpha\langle x_{n+1} - x_n, x_{n+1} - x\rangle_{\mathcal{X}} + 2\gamma\nu\langle y_n - y_{n-1}, y_n - y\rangle_{\mathcal{Y}} \\
&= 2\gamma\left\{\Phi(x_{n+1}, y_n) - \Phi(x, y)\right\} + \|(Ax - By) - (Ax_n - By_n)\|_{\mathcal{Z}}^2 \\
&\quad + \gamma\alpha(\|x_{n+1} - x_n\|_{\mathcal{X}}^2 + \|x_{n+1} - x\|_{\mathcal{X}}^2 - \|x_n - x\|_{\mathcal{X}}^2) \\
&\quad + \gamma\nu(\|y_n - y_{n-1}\|_{\mathcal{Y}}^2 + \|y_n - y\|_{\mathcal{Y}}^2 - \|y_{n-1} - y\|_{\mathcal{Y}}^2).
\end{aligned}$$

Finally, we find

$$\begin{aligned}
&\|Ax - Ax_n\|_{\mathcal{Z}}^2 + \gamma\alpha\|x_n - x\|_{\mathcal{X}}^2 + \gamma\nu\|y_{n-1} - y\|_{\mathcal{Y}}^2 \\
&\qquad - \|Ax - Ax_{n+1}\|_{\mathcal{Z}}^2 - \gamma\alpha\|x_{n+1} - x\|_{\mathcal{X}}^2 - \gamma\nu\|y_n - y\|_{\mathcal{Y}}^2 \\
&\qquad\geq 2\gamma\left\{\Phi(x_{n+1}, y_n) - \Phi(x, y)\right\} + \|(Ax - By) - (Ax_n - By_n)\|_{\mathcal{Z}}^2 \\
&\qquad\qquad + \gamma\alpha\|x_{n+1} - x_n\|_{\mathcal{X}}^2 + \gamma\nu\|y_n - y_{n-1}\|_{\mathcal{Y}}^2. \quad (5)
\end{aligned}$$

Let us prove that $\inf\Phi \geq \varphi$, thus implying $\inf\Phi = \varphi$. Let us argue by contradiction and assume that $\inf\Phi < \varphi$. There exist $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ so that $\inf\Phi \leq \Phi(x, y) < \varphi$. By summing inequality (5), we get

$$2\gamma\sum_{n\geq 1}(\varphi - \Phi(x, y)) \leq \|Ax - Ax_1\|_{\mathcal{Z}}^2 + \gamma\alpha\|x_1 - x\|_{\mathcal{X}}^2 + \gamma\nu\|y_0 - y\|_{\mathcal{Y}}^2 < +\infty,$$

and we obtain a contradiction hence $\inf\Phi \geq \varphi$.
(iii) Taking $x = x_n$ and $y = y_{n-1}$ in inequality (5), we obtain

$$\alpha\|x_{n+1} - x_n\|_{\mathcal{X}}^2 + \nu\|y_n - y_{n-1}\|_{\mathcal{Y}}^2 \leq \Phi(x_n, y_{n-1}) - \Phi(x_{n+1}, y_n).$$

By summing this inequality, we infer

$$\alpha\sum_{n\geq 1}\|x_{n+1} - x_n\|_{\mathcal{X}}^2 + \nu\sum_{n\geq 1}\|y_n - y_{n-1}\|_{\mathcal{Y}}^2 \leq \Phi(x_1, y_0) - \inf\Phi < +\infty,$$

this achieves the proof of item (iii).
(iv)(a) In view of inequality (5), the sequence $(\|Ax - Ax_n\|_{\mathcal{Z}}^2 + \gamma\alpha\|x_n - x\|_{\mathcal{X}}^2 + \gamma\nu\|y_{n-1} - y\|_{\mathcal{Y}}^2)$ is nonincreasing and nonnegative hence convergent. Let us prove the same result for the sequence $(\|By - By_n\|_{\mathcal{Z}}^2 + \gamma\alpha\|x_n - x\|_{\mathcal{X}}^2 + \gamma\nu\|y_n - y\|_{\mathcal{Y}}^2)$.

Using Lemma 2.1 with $s = By,\ t = Ax,\ u = By_n,\ v = Ax_{n+1},\ w = By_{n+1}$, we obtain

$$\begin{aligned}\|By - By_n\|_{\mathcal{Z}}^2 - \|By - By_{n+1}\|_{\mathcal{Z}}^2 = &\|By_{n+1} - Ax_{n+1}\|_{\mathcal{Z}}^2 - \|By - Ax\|_{\mathcal{Z}}^2 \\ &+\|(By - Ax) - (By_n - Ax_{n+1})\|_{\mathcal{Z}}^2 \\ &+2\langle By - By_{n+1}, By_{n+1} - Ax_{n+1}\rangle_{\mathcal{Z}} \\ &+2\langle By_n - Ax_{n+1}, Ax_{n+1} - Ax\rangle_{\mathcal{Z}}. \qquad (6)\end{aligned}$$

Using inequalities (3), we have

$$\begin{aligned}&\|By - By_n\|_{\mathcal{Z}}^2 - \|By - By_{n+1}\|_{\mathcal{Z}}^2 \\ &\geq \|By_{n+1} - Ax_{n+1}\|_{\mathcal{Z}}^2 - \|By - Ax\|_{\mathcal{Z}}^2 + \|(By - Ax) - (By_n - Ax_{n+1}\|_{\mathcal{Z}}^2 \\ &\quad + 2\gamma\left\{g(y_{n+1}) - g(y) + \nu\langle y_{n+1} - y_n, y_{n+1} - y\rangle_{\mathcal{Y}}\right\} \\ &\quad + 2\gamma\left\{f(x_{n+1}) - f(x) + \alpha\langle x_{n+1} - x_n, x_{n+1} - x\rangle_{\mathcal{X}}\right\},\end{aligned}$$

and by using the same arguments as above, we obtain

$$\begin{aligned}&\|By - By_n\|_{\mathcal{Z}}^2 + \gamma\alpha\|x_n - x\|_{\mathcal{X}}^2 + \gamma\nu\|y_n - y\|_{\mathcal{Y}}^2 \\ &\qquad - \|By - By_{n+1}\|_{\mathcal{Z}}^2 - \gamma\alpha\|x_{n+1} - x\|_{\mathcal{X}}^2 - \gamma\nu\|y_{n+1} - y\|_{\mathcal{Y}}^2 \\ &\quad \geq 2\gamma\left\{\Phi(x_{n+1}, y_{n+1}) - \Phi(x, y)\right\} + \|(Ax - By) - (Ax_{n+1} - By_n)\|_{\mathcal{Z}}^2 \\ &\qquad + \gamma\alpha\|x_{n+1} - x_n\|_{\mathcal{X}}^2 + \gamma\nu\|y_{n+1} - y_n\|_{\mathcal{Y}}^2, \qquad (7)\end{aligned}$$

this achieves the proof of (iv)(a).
(iv)(b) This claim follows by summing inequalities (5) and (7).
(iv)(c) Here we adapt an argument borrowed from [2]. Let us use Opial's lemma [8], that we recall below for the sake of completeness.

Lemma 2.2 (Opial) *Let H be a Hilbert space endowed with the norm N. Let (ξ_n) be a sequence of H such that there exists a nonempty set $S \subset H$ which verifies*

(1)For all $\xi \in S$, $\lim_{n\to+\infty} N(\xi_n - \xi)$ exists.
(2)If $(\xi_{n_k}) \rightharpoonup \overline{\xi}$ weakly in H as $k \to +\infty$, we have $\overline{\xi} \in S$.

Then the sequence (ξ_n) weakly converges in H as $n \to +\infty$ toward a point of S.

Let us define the norm $N(x, y) = \left(\|By\|_{\mathcal{Z}}^2 + \gamma\alpha\|x\|_{\mathcal{X}}^2 + \gamma\nu\|y\|_{\mathcal{Y}}^2\right)^{1/2}$ on the space $\mathcal{X} \times \mathcal{Y}$. Since the linear operator B is continuous, the norm N is equivalent to the canonical norm on $\mathcal{X} \times \mathcal{Y}$. Thus, in view of (iv)(a),

$$\forall (x, y) \in \mathrm{Argmin}\Phi, \quad N((x_n, y_n) - (x, y)) \text{ has a limit when } n \to +\infty,$$

which shows point (1). Let (x_{n_k}, y_{n_k}) be a subsequence of (x_n, y_n) which weakly converges toward $(\overline{x}, \overline{y})$. Using the closedness of Φ and item (ii), we can write

$$\Phi(\overline{x}, \overline{y}) \leq \liminf_{k\to+\infty} \Phi(x_{n_k}, y_{n_k}) = \lim_{n\to+\infty} \Phi(x_n, y_n) = \inf \Phi,$$

hence $(\overline{x},\overline{y}) \in \operatorname{Argmin}\Phi$, which shows point (2). Opial's lemma then shows that (x_n, y_n) weakly converges toward a point $(\overline{x},\overline{y})$ in $\operatorname{Argmin}\Phi$.
Let us prove that $f(x_n) \to f(\overline{x})$ as $n \to +\infty$. Using the closedness of f, we have $f(\overline{x}) \leq \liminf_{n\to+\infty} f(x_n)$. By using inequality (3) with $x = \overline{x}$, we obtain

$$f(\overline{x}) \geq f(x_{n+1}) - \frac{1}{\gamma}\langle Ax_{n+1} - By_n, A\overline{x} - Ax_{n+1}\rangle_{\mathcal{Z}} - \alpha\langle x_{n+1} - x_n, \overline{x} - x_{n+1}\rangle_{\mathcal{X}}.$$

Since the linear operator A is continuous and since the sequence (x_n) weakly converges towards $\overline{x}$, we derive that the sequence (Ax_n) weakly converges towards $A\overline{x}$. Moreover, from item (iv)(b), the sequence $(Ax_{n+1} - By_n)$ strongly converges in $\mathcal{Z}$ toward $(A\overline{x} - B\overline{y})$ and, from item (iii), the sequence $(x_{n+1} - x_n)$ strongly converges in $\mathcal{X}$ toward 0. Hence we deduce from the above inequality that $\limsup_{n\to+\infty} f(x_n) \leq f(\overline{x})$ and finally $\lim_{n\to+\infty} f(x_n) = f(\overline{x})$. In a similar way, we easily infer that $\lim_{n\to+\infty} g(y_n) = g(\overline{y})$.
(v) Let us argue by contradiction and assume that the conclusion is false. We can extract a subsequence (x_{n_k}, y_{n_k}) which weakly converges toward a point of $\mathcal{X} \times \mathcal{Y}$. The closedness of Φ implies that this point is a minimizer of Φ, which is a contradiction.

3 Dual problem

Let us define the map $p : \mathcal{Z} \to \mathbb{R}$ by $p(z) = \inf\{f(x) + g(y) + \frac{1}{2\gamma}\|Ax - By - z\|^2_{\mathcal{Z}};\quad x \in \mathcal{X}, y \in \mathcal{Y}\}$. We recover the primal problem $(\mathcal{P})$ for $z = 0$. Since the map $(x, y, z) \in \mathcal{X} \times \mathcal{Y} \times \mathcal{Z} \longrightarrow f(x) + g(y) + \frac{1}{2\gamma}\|Ax - By - z\|^2_{\mathcal{Z}} \in \mathbb{R} \cup \{+\infty\}$ is convex, the map p is convex. Moreover p is locally majorized, hence continuous. By the Fenchel Moreau Rockafellar's theorem, we can assert that

$$p(z) = \sup\{\langle z^*, z\rangle_{\mathcal{Z}} - p^*(z^*);\quad z^* \in \mathcal{Z}\},$$

where $p^* : \mathcal{Z} \to \mathbb{R} \cup \{+\infty\}$ is the conjugate function of p. In particular we obtain $p(0) = -\inf\{p^*(z^*); z^* \in \mathcal{Z}\}$. The map p^* is precised in the following lemma.

Lemma 3.1 $\forall z^* \in \mathcal{Z},\quad p^*(z^*) = f^*(A^*z^*) + g^*(-B^*z^*) + \frac{\gamma}{2}\|z^*\|^2_{\mathcal{Z}}$.

Proof.

$$\begin{aligned}
p^*(z^*) &= \sup\{\langle z^*, z\rangle_{\mathcal{Z}} - p(z); z \in \mathcal{Z}\}\\
&= \sup\left\{\langle z^*, z\rangle_{\mathcal{Z}} - \inf\left\{f(x) + g(y) + \frac{1}{2\gamma}\|Ax - By - z\|^2_{\mathcal{Z}}; x \in \mathcal{X}, y \in \mathcal{Y}\right\}; z \in \mathcal{Z}\right\}\\
&= \sup\left\{\sup\left\{\langle z^*, z\rangle_{\mathcal{Z}} - (f(x) + g(y) + \frac{1}{2\gamma}\|Ax - By - z\|^2_{\mathcal{Z}}); x \in \mathcal{X}, y \in \mathcal{Y}\right\}; z \in \mathcal{Z}\right\}\\
&= \sup\left\{\sup\left\{\langle z^*, z\rangle_{\mathcal{Z}} - (f(x) + g(y) + \frac{1}{2\gamma}\|Ax - By - z\|^2_{\mathcal{Z}}); z \in \mathcal{Z}\right\}; x \in \mathcal{X}, y \in \mathcal{Y}\right\}\\
&= \sup\left\{-f(x) - g(y) + \sup\left\{\langle z^*, z\rangle_{\mathcal{Z}} - \frac{1}{2\gamma}\|Ax - By - z\|^2_{\mathcal{Z}}; z \in \mathcal{Z}\right\}; x \in \mathcal{X}, y \in \mathcal{Y}\right\}.
\end{aligned}$$

By a differential computation, we let the reader check that $\sup\left\{\langle z^*, z\rangle_{\mathcal{Z}} - \frac{1}{2\gamma}\|Ax - By - z\|^2_{\mathcal{Z}}; z \in \mathcal{Z}\right\} = \langle z^*, \gamma z^* + Ax - By\rangle_{\mathcal{Z}} - \frac{\gamma}{2}\|z^*\|^2_{\mathcal{Z}}$. Hence, we deduce that

$$\begin{aligned}
p^*(z^*) &= \sup\left\{-f(x) - g(y) + \langle z^*, \gamma z^* + Ax - By\rangle_{\mathcal{Z}} - \frac{\gamma}{2}\|z^*\|^2_{\mathcal{Z}}; x \in \mathcal{X}, y \in \mathcal{Y}\right\}\\
&= \sup\left\{\frac{\gamma}{2}\|z^*\|^2_{\mathcal{Z}} + \langle A^*z^*, x\rangle_{\mathcal{X}} - f(x) + \langle -B^*z^*, y\rangle_{\mathcal{Y}} - g(y); x \in \mathcal{X}, y \in \mathcal{Y}\right\}\\
&= \frac{\gamma}{2}\|z^*\|^2_{\mathcal{Z}} + \sup\{\langle A^*z^*, x\rangle_{\mathcal{X}} - f(x); x \in \mathcal{X}\} + \sup\{\langle -B^*z^*, y\rangle_{\mathcal{Y}} - g(y); y \in \mathcal{Y}\}\\
&= \frac{\gamma}{2}\|z^*\|^2_{\mathcal{Z}} + f^*(A^*z^*) + g^*(-B^*z^*).
\end{aligned}$$

We denote by $(\mathcal{P}^*)$ the following minimization problem

$$(\mathcal{P}^*) \qquad \inf\left\{f^*(A^*z^*) + g^*(-B^*z^*) + \frac{\gamma}{2}\|z^*\|^2_{\mathcal{Z}}; \quad z^* \in \mathcal{Z}\right\}.$$

Hence problems $(\mathcal{P})$ and $(\mathcal{P}^*)$ are linked by the relation $\inf\mathcal{P} = -\inf\mathcal{P}^*$. Since the function $z^* \mapsto f^*(A^*z^*) + g^*(-B^*z^*) + \frac{\gamma}{2}\|z^*\|^2_{\mathcal{Z}}$ is closed, proper and strongly convex, $(\mathcal{P}^*)$ has a unique solution $\overline{z^*}$.

Proposition 3.1 *Let γ be a positive coefficient and let $A \in \mathcal{L}(\mathcal{X}, \mathcal{Z})$, $B \in \mathcal{L}(\mathcal{Y}, \mathcal{Z})$ be linear continuous operators. Let $f : \mathcal{X} \longrightarrow \mathbb{R} \cup \{+\infty\}$ and $g : \mathcal{Y} \longrightarrow \mathbb{R} \cup \{+\infty\}$ be closed convex proper functions. Assume that the function Φ defined by equality (1) is bounded from below. Let $\overline{z^*}$ be the unique minimizer of $(\mathcal{P}^*)$. If (u_n, v_n) is a sequence in $\mathcal{X} \times \mathcal{Y}$ such that $\lim_{n\to+\infty}\Phi(u_n, v_n) = \inf\Phi$, then $\lim_{n\to+\infty} -\frac{1}{\gamma}(Au_n - Bv_n) = \overline{z^*}$ strongly in $\mathcal{Z}$.*

Proof. Recalling that $\inf\Phi = -\inf\mathcal{P}^* = -\{f^*(A^*\overline{z^*}) + g^*(-B^*\overline{z^*}) + \frac{\gamma}{2}\|\overline{z^*}\|^2_{\mathcal{Z}}\}$, we have

$$\Phi(u_n, v_n) - \inf\Phi = f(u_n) + f^*(A^*\overline{z^*}) + g(v_n) + g^*(-B^*\overline{z^*}) + \frac{1}{2\gamma}\|Au_n - Bv_n\|^2_{\mathcal{Z}} + \frac{\gamma}{2}\|\overline{z^*}\|^2_{\mathcal{Z}}.$$

Using the Fenchel's inequality, we find

$$\begin{aligned}
\Phi(u_n, v_n) - \inf\Phi &\geq \langle A^*\overline{z^*}, u_n\rangle_{\mathcal{X}} + \langle -B^*\overline{z^*}, v_n\rangle_{\mathcal{Y}} + \frac{1}{2\gamma}\|Au_n - Bv_n\|^2_{\mathcal{Z}} + \frac{\gamma}{2}\|\overline{z^*}\|^2_{\mathcal{Z}}\\
&= \frac{1}{2\gamma}\|\gamma\overline{z^*} + Au_n - Bv_n\|^2_{\mathcal{Z}}.
\end{aligned}$$

Since $\lim_{n\to+\infty}\Phi(u_n, v_n) = \inf\Phi$, the conclusion is immediate.

Remark 3.1 *Assume moreover that Argmin$\Phi \neq \varnothing$. Then for every $(x, y) \in$ ArgminΦ, the vector $(Ax - By)$ is constant*[1] *and we have $\overline{z^*} = -\frac{1}{\gamma}(Ax - By)$. This is obvious from the previous proposition by taking $(u_n, v_n) = (x, y)$ for every $n \in \mathbb{N}$.*

[1] We can recover this result by observing that, for every $(x, y) \in$ ArgminΦ, $\lim_{n\to+\infty}\|Ax - By - (Ax_n - By_n)\|^2_{\mathcal{Z}} = 0$ from Theorem 2.1 (iv)(b).

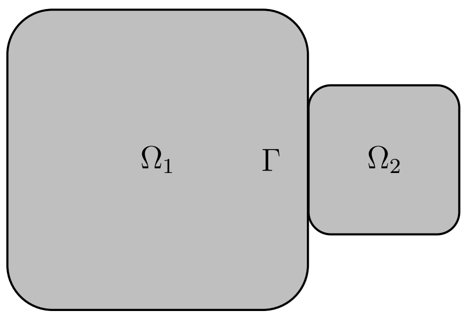
We are now able to deduce the following corollary of Theorem 2.1.

Corollary 3.1 *Under the assumptions of Proposition 3.1, consider a sequence* (x_n, y_n) *generated by* $(\mathcal{A})$*. Then the sequences* $\left(-\frac{1}{\gamma}(Ax_n - By_n)\right)$ *and* $\left(-\frac{1}{\gamma}(Ax_{n+1} - By_n)\right)$ *strongly converge in* $\mathcal{Z}$ *to the unique minimizer* $\overline{z^*}$ *of* $(\mathcal{P}^*)$*.*

Proof. This follows immediately from Proposition 3.1 and Theorem 2.1 (ii).

4 Application to domain decomposition for PDE's

Let us consider a bounded domain $\Omega = \Omega_1 \cup \Omega_2 \cup \Gamma$ of $\mathbb{R}^N$ which can be decomposed in two nonoverlapping subdomains Ω_1 and Ω_2 with a common interface Γ. We assume that the open sets Ω_1 and Ω_2 are of class $\mathcal{C}^1$ and that $\mathcal{H}^{N-1}(\Gamma) > 0$, where $\mathcal{H}^{N-1}$ is the Hausdorff measure of dimension $N-1$.



Given some $h \in L^2(\Omega)$ such that $\int_\Omega h = 0$, we are interested in the following variational problem

$(\mathcal{P})$

$$\min\left\{\frac{1}{2}\int_{\Omega_1}|\nabla u|^2 + \frac{1}{2}\int_{\Omega_2}|\nabla v|^2 + \frac{1}{2\gamma}\int_\Gamma [w]^2 - \int_\Omega hw; \quad u \in H^1(\Omega_1), v \in H^1(\Omega_2)\right\}$$

where $w = \begin{cases} u \text{ on } \Omega_1 \\ v \text{ on } \Omega_2 \end{cases}$ and $[w]$ =jump of w through the interface Γ.

This kind of minimization problems often arises in the description of phenomena where the boundary is free, *i.e.* no external action is exerted on $\partial\Omega$, and involving discontinuities through the interface. Attouch, Bolte, Redont and Soubeyran consider in [2] the corresponding Dirichlet version of problem $(\mathcal{P})$. On the other hand, the companion paper [4] analyses the Neumann problem $(\mathcal{P})$ formally associated with the value $\gamma = 0$. This forces the jump to be equal to zero and the corresponding solutions satisfy a Neumann problem on the whole set Ω. The recent paper [7] studies the opposite situation corresponding formally to $\gamma = \infty$, see the concluding comments for more details.

Let us now show how algorithm $(\mathcal{A})$ can be applied so as to solve problem $(\mathcal{P})$. The space $\mathcal{X} = H^1(\Omega_1)$ is equipped with the scalar product $\langle u_1, u_2\rangle_{\mathcal{X}} = \int_{\Omega_1}(\nabla u_1.\nabla u_2 + u_1 u_2)$ and the corresponding norm. The same holds for $\mathcal{Y} = H^1(\Omega_2)$ by replacing Ω_1 with Ω_2. The space $\mathcal{Z} = L^2(\Gamma)$ is equipped with the scalar product $\langle z_1, z_2\rangle_{\mathcal{Z}} = \int_\Gamma z_1 z_2$ and the associate norm. Problem $(\mathcal{P})$ can be reformulated as

$$\begin{aligned} &\min\{\Phi(u,v);\quad u\in\mathcal{X}, v\in\mathcal{Y}\}\\ =&\min\left\{f(u)+g(v)+\frac{1}{2\gamma}\|Au-Bv\|^2_{\mathcal{Z}};\quad u\in\mathcal{X}, v\in\mathcal{Y}\right\},\end{aligned}$$

where

$$f(u)=\frac{1}{2}\int_{\Omega_1}|\nabla u|^2-\int_{\Omega_1}hu \quad\text{and}\quad g(v)=\frac{1}{2}\int_{\Omega_2}|\nabla v|^2-\int_{\Omega_2}hv,$$

and the operators $A:\mathcal{X}\longrightarrow\mathcal{Z}$ and $B:\mathcal{Y}\longrightarrow\mathcal{Z}$ are respectively the trace operators on Γ. Algorithm $(\mathcal{A})$ runs as follows

$$\begin{cases} u_{n+1}=\operatorname{Argmin}\left\{f(u)+\frac{1}{2\gamma}\|Au-Bv_n\|^2_{\mathcal{Z}}+\frac{\alpha}{2}\|u-u_n\|^2_{\mathcal{X}};\quad u\in\mathcal{X}\right\}\\ v_{n+1}=\operatorname{Argmin}\left\{g(v)+\frac{1}{2\gamma}\|Au_{n+1}-Bv\|^2_{\mathcal{Z}}+\frac{\nu}{2}\|v-v_n\|^2_{\mathcal{Y}};\quad v\in\mathcal{Y}\right\},\end{cases}$$

where α and ν are fixed positive parameters. An elementary directional derivative computation shows that the weak variational formulation of algorithm $(\mathcal{A})$ is given by

$$\begin{aligned}\forall u\in\mathcal{X},\quad &\int_{\Omega_1}\nabla u_{n+1}.\nabla u+\frac{1}{\gamma}\int_\Gamma(Au_{n+1}-Bv_n)Au\\ &+\alpha\int_{\Omega_1}(\nabla u_{n+1}-\nabla u_n).\nabla u+\alpha\int_{\Omega_1}(u_{n+1}-u_n)u=\int_{\Omega_1}hu, \qquad (8)\\ \forall v\in\mathcal{Y},\quad &\int_{\Omega_2}\nabla v_{n+1}.\nabla v+\frac{1}{\gamma}\int_\Gamma(Bv_{n+1}-Au_{n+1})Bv\\ &+\nu\int_{\Omega_2}(\nabla v_{n+1}-\nabla v_n).\nabla v+\nu\int_{\Omega_2}(v_{n+1}-v_n)v=\int_{\Omega_2}hv. \qquad (9)\end{aligned}$$

Equality (8) is the variational weak formulation of the following mixed Dirichlet-Neumann boundary value problem on Ω_1

$$\begin{cases} -(1+\alpha)\Delta u_{n+1}+\alpha u_{n+1}=h-\alpha\Delta u_n+\alpha u_n & \text{on }\ \Omega_1\\ (1+\alpha)\frac{\partial u_{n+1}}{\partial n}=\alpha\frac{\partial u_n}{\partial n} & \text{on }\ \partial\Omega_1\cap\partial\Omega\\ (1+\alpha)\frac{\partial u_{n+1}}{\partial n}+\frac{1}{\gamma}u_{n+1}=\alpha\frac{\partial u_n}{\partial n}+\frac{1}{\gamma}v_n & \text{on }\ \Gamma,\end{cases}$$

where, for $u\in\mathcal{X}$, $\frac{\partial u}{\partial n}=\nabla u.\nabla\overrightarrow{n}$ and $\overrightarrow{n}$ is the unit outward normal to $\partial\Omega_1$. In the same way, equality (9) gives

$$\begin{cases} -(1+\nu)\Delta v_{n+1} + \nu v_{n+1} = h - \nu\Delta v_n + \nu v_n & \text{on } \Omega_2 \\ (1+\nu)\frac{\partial v_{n+1}}{\partial n} = \nu\frac{\partial v_n}{\partial n} & \text{on } \partial\Omega_2 \cap \partial\Omega \\ (1+\nu)\frac{\partial v_{n+1}}{\partial n} + \frac{1}{\gamma}v_{n+1} = \nu\frac{\partial v_n}{\partial n} + \frac{1}{\gamma}u_{n+1} & \text{on } \Gamma. \end{cases}$$

In order to apply Theorem 2.1, let us describe the set ArgminΦ.

Claim 4.1 *If* $(\overline{u}, \overline{v}) \in Argmin\Phi$*, then* $Argmin\Phi = \{(\overline{u}+C, \overline{v}+C); \quad C \in \mathbb{R}\}$.

Proof. In view of hypothesis $\int_\Omega h = 0$, it is immediate to check that, for all $C \in \mathbb{R}$, $\Phi(\overline{u}+C, \overline{v}+C) = \Phi(\overline{u}, \overline{v})$ hence $(\overline{u}+C, \overline{v}+C) \in \text{Argmin}\Phi$. Let us prove the reverse inclusion. By a differential computation, $(\overline{u}, \overline{v}) \in \text{Argmin}\Phi$ if and only if

$$\forall k \in \mathcal{X}, \quad \int_{\Omega_1} \nabla\overline{u}.\nabla k = \int_{\Omega_1} hk - \frac{1}{\gamma}\int_\Gamma (A\overline{u} - B\overline{v})Ak, \tag{10}$$

$$\forall l \in \mathcal{Y}, \quad \int_{\Omega_2} \nabla\overline{v}.\nabla l = \int_{\Omega_2} hl + \frac{1}{\gamma}\int_\Gamma (A\overline{u} - B\overline{v})Bl. \tag{11}$$

Let $(\overline{u}_1, \overline{v}_1)$,$(\overline{u}_2, \overline{v}_2)$ be in ArgminΦ. Using equality (10) respectively with $(\overline{u}_1, \overline{v}_1)$ and $(\overline{u}_2, \overline{v}_2)$, then subtracting, we find

$$\forall k \in \mathcal{X}, \quad \int_{\Omega_1} \nabla(\overline{u}_1 - \overline{u}_2).\nabla k = -\frac{1}{\gamma}\int_\Gamma (A(\overline{u}_1 - \overline{u}_2) - B(\overline{v}_1 - \overline{v}_2))Ak.$$

Taking $k = \overline{u}_1 - \overline{u}_2$, we infer that

$$\int_{\Omega_1} |\nabla(\overline{u}_1 - \overline{u}_2)|^2 + \frac{1}{\gamma}\int_\Gamma (A(\overline{u}_1 - \overline{u}_2) - B(\overline{v}_1 - \overline{v}_2))A(\overline{u}_1 - \overline{u}_2) = 0.$$

In the same way, using equality (11), we have

$$\int_{\Omega_2} |\nabla(\overline{v}_1 - \overline{v}_2)|^2 - \frac{1}{\gamma}\int_\Gamma (A(\overline{u}_1 - \overline{u}_2) - B(\overline{v}_1 - \overline{v}_2))B(\overline{v}_1 - \overline{v}_2) = 0.$$

Finally we find

$$\int_{\Omega_1} |\nabla(\overline{u}_1 - \overline{u}_2)|^2 + \int_{\Omega_2} |\nabla(\overline{v}_1 - \overline{v}_2)|^2 + \frac{1}{\gamma}\int_\Gamma (A(\overline{u}_1 - \overline{u}_2) - B(\overline{v}_1 - \overline{v}_2))^2 = 0.$$

We deduce that $\nabla(\overline{u}_1 - \overline{u}_2) = \nabla(\overline{v}_1 - \overline{v}_2) = 0$, hence $\overline{u}_1 - \overline{u}_2 = C_1$ and $\overline{v}_1 - \overline{v}_2 = C_2$ for some $C_1, C_2 \in \mathbb{R}$. Using $\mathcal{H}^{N-1}(\Gamma) > 0$ and $\int_\Gamma (A(\overline{u}_1 - \overline{u}_2) - B(\overline{v}_1 - \overline{v}_2))^2 = 0$, we conclude that $C_1 = C_2$.

Remark 4.1 *Equalities (10) and (11) are the variational weak formulations of the following mixed Dirichlet-Neumann boundary value problems respectively on* Ω_1

$$\begin{cases} -\Delta\overline{u} = h & \textit{on } \Omega_1 \\ \frac{\partial\overline{u}}{\partial n} = 0 & \textit{on } \partial\Omega_1 \cap \partial\Omega \\ \frac{\partial\overline{u}}{\partial n} = -\frac{1}{\gamma}(\overline{u} - \overline{v}) & \textit{on } \Gamma, \end{cases}$$

and Ω_2

$$\begin{cases} -\Delta\overline{v} = h & \textit{on}\ \ \Omega_2 \\ \frac{\partial \overline{v}}{\partial n} = 0 & \textit{on}\ \ \partial\Omega_2 \cap \partial\Omega \\ \frac{\partial \overline{v}}{\partial n} = \frac{1}{\gamma}(\overline{u} - \overline{v}) & \textit{on}\ \ \Gamma. \end{cases}$$

We must prove that the set ArgminΦ is nonempty. Let us note the symmetric continuous bilinear form $Q : (\mathcal{X} \times \mathcal{Y}) \times (\mathcal{X} \times \mathcal{Y}) \to \mathbb{R}$ and the linear continuous form $L : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ respectively defined, for all $(u, v), (k, l) \in \mathcal{X} \times \mathcal{Y}$, by

$$Q((u,v),(k,l)) = \int_{\Omega_1} \nabla u.\nabla k + \int_{\Omega_2} \nabla v.\nabla l + \frac{1}{\gamma}\int_{\Gamma}(Au - Bv)(Ak - Bl),$$
$$L(u,v) = \int_{\Omega_1} hu + \int_{\Omega_2} hv.$$

First observe that the bilinear form Q is not coercive on the space $(\mathcal{X}\times\mathcal{Y})\times(\mathcal{X}\times\mathcal{Y})$. To remedy this lack of coercivity, we have to consider the following suitable closed hyperplane[2] $\mathcal{F}$ of $\mathcal{X} \times \mathcal{Y}$ defined by[3]

$$\mathcal{F} = \left\{(u,v) \in \mathcal{X} \times \mathcal{Y};\quad \int_{\Omega_1} u + \int_{\Omega_2} v = 0\right\}.$$

Claim 4.2 *There exists* $\epsilon > 0$ *such that, for every* $(u, v) \in \mathcal{F}$,

$$Q((u,v),(u,v)) \geq \epsilon \|(u,v)\|^2_{L^2(\Omega_1)\times L^2(\Omega_2)}.$$

Proof. The proof uses the same arguments as those of [3, Theorem 5.4.3]. Let us argue by contradiction and assume that there exists a sequence (u_n, v_n) in $\mathcal{F}$, $(u_n, v_n) \neq 0$, such that

$$\int_{\Omega_1} |\nabla u_n|^2 + \int_{\Omega_2} |\nabla v_n|^2 + \frac{1}{\gamma}\int_{\Gamma}(Au_n - Bv_n)^2 < \frac{1}{n}\left(\int_{\Omega_1} u_n^2 + \int_{\Omega_2} v_n^2\right).$$

Let us take $(\tilde{u}_n, \tilde{v}_n) = (u_n, v_n)/\|(u_n, v_n)\|_{L^2(\Omega_1)\times L^2(\Omega_2)}$. We have $(\tilde{u}_n, \tilde{v}_n) \in \mathcal{F}$, $\|(\tilde{u}_n, \tilde{v}_n)\|_{L^2(\Omega_1)\times L^2(\Omega_2)} = 1$ and $\int_{\Omega_1} |\nabla \tilde{u}_n|^2 + \int_{\Omega_2} |\nabla \tilde{v}_n|^2 + \frac{1}{\gamma}\int_\Gamma (A\tilde{u}_n - B\tilde{v}_n)^2 < \frac{1}{n}$, hence $\lim_{n\to+\infty} \int_{\Omega_1} |\nabla \tilde{u}_n|^2 + \int_{\Omega_2} |\nabla \tilde{v}_n|^2 = 0$. We deduce that the sequence $(\tilde{u}_n, \tilde{v}_n)$ is bounded in $\mathcal{X} \times \mathcal{Y}$ and, by the Rellich-Kondrakov theorem, we can extract a subsequence $(\tilde{u}_{n_k}, \tilde{v}_{n_k})$ which strongly converges in $L^2(\Omega_1) \times L^2(\Omega_2)$ toward $(\tilde{u}, \tilde{v})$. Since $\lim_{k\to+\infty} \int_{\Omega_1} |\nabla \tilde{u}_{n_k}|^2 + \int_{\Omega_2} |\nabla \tilde{v}_{n_k}|^2 = 0$, we infer that the sequence $(\tilde{u}_{n_k}, \tilde{v}_{n_k})$ strongly converges in $\mathcal{X} \times \mathcal{Y}$ and $\int_{\Omega_1} |\nabla \tilde{u}|^2 + \int_{\Omega_2} |\nabla \tilde{v}|^2 + \frac{1}{\gamma}\int_\Gamma (A\tilde{u} - B\tilde{v})^2 = 0$. Hence, as in the proof of Claim 4.1, we deduce that $(\tilde{u}, \tilde{v}) = (C, C)$ for some $C \in \mathbb{R}$. Since the hyperplane $\mathcal{F}$ is closed, we have $(C, C) \in \mathcal{F}$ thus implying that $C = 0$ which is a contradiction with $\|(C, C)\|_{L^2(\Omega_1)\times L^2(\Omega_2)} = 1$.

[2] The author is indebted to A. Cabot for this pertinent remark.

[3] Notice that the hyperplane $\mathcal{F}$ is the orthogonal space of the one-dimensional closed subspace $\{(C, C);\quad C \in \mathbb{R}\}$ of $\mathcal{X} \times \mathcal{Y}$ which naturally appears in Claim 4.1.

We deduce from Claim 4.2 that the bilinear form Q is coercive on the space $\mathcal{F}\times\mathcal{F}$. We can now state the following result.

Claim 4.3 *There exists* $(\overline{u},\overline{v}) \in \mathcal{X}\times\mathcal{Y}$ *which verifies* $\int_{\Omega_1}\overline{u}+\int_{\Omega_2}\overline{v}=0$*, such that* $Argmin\Phi = \{(\overline{u}+C,\overline{v}+C); \quad C\in\mathbb{R}\}$.

Proof. The bilinear form Q is symmetric, continuous and coercive on $\mathcal{F}\times\mathcal{F}$ and the linear form L is continuous on $\mathcal{F}$. By applying the Lax-Milgram theorem, there exists a unique $(\overline{u},\overline{v}) \in \mathcal{F}$ such that, for all $(k,l)\in\mathcal{F}$, $Q((\overline{u},\overline{v}),(k,l)) = L(k,l)$, *i.e.* ,

$$\int_{\Omega_1}\nabla\overline{u}.\nabla k+\int_{\Omega_2}\nabla\overline{v}.\nabla l=\int_{\Omega_1}hk+\int_{\Omega_2}hl-\frac{1}{\gamma}\int_\Gamma(A\overline{u}-B\overline{v})(Ak-Bl). \quad (12)$$

Let $(k,l)\in\mathcal{X}\times\mathcal{Y}$ and let us note $C=\frac{1}{|\Omega_1|+|\Omega_2|}\left(\int_{\Omega_1}k+\int_{\Omega_2}l\right)$, we can verify that $(k-C,l-C)\in\mathcal{F}$. Thus, using equality (12) with $(k-C,l-C)$, we obtain that $(\overline{u},\overline{v})$ verifies equalities (10) and (11) hence $(\overline{u},\overline{v})$ is a minimizer of Φ on $\mathcal{X}\times\mathcal{Y}$. It suffices to apply Claim 4.1 to achieve the proof.

We conclude from Theorem 2.1 (iv)(c) and the above analysis that any sequence (u_n,v_n) generated by $(\mathcal{A})$ weakly converges in $H^1(\Omega_1)\times H^1(\Omega_2)$ toward a minimum point $(\overline{u}+C,\overline{v}+C)$, $(C\in\mathbb{R})$ of problem $(\mathcal{P})$. Without loss of generality, we can assume that $C=0$. Since the injections $H^1(\Omega_1)\hookrightarrow L^2(\Omega_1)$ and $H^1(\Omega_2)\hookrightarrow L^2(\Omega_2)$ are compact, the convergence is strong in $L^2(\Omega_1)\times L^2(\Omega_2)$. Moreover, from Theorem 2.1 (iv)(c), we have $\lim_{n\to+\infty} f(u_n)=f(\overline{u})$ and $\lim_{n\to+\infty} g(v_n)=g(\overline{v})$, hence $\lim_{n\to+\infty}\int_{\Omega_1}|\nabla u_n|^2=\int_{\Omega_1}|\nabla\overline{u}|^2$ and $\lim_{n\to+\infty}\int_{\Omega_2}|\nabla v_n|^2=\int_{\Omega_2}|\nabla\overline{v}|^2$. As a consequence, we have $\lim_{n\to+\infty}\|(u_n,v_n)\|_{H^1(\Omega_1)\times H^1(\Omega_2)}=\|(\overline{u},\overline{v})\|_{H^1(\Omega_1)\times H^1(\Omega_2)}$. Since (u_n,v_n) weakly converges in $H^1(\Omega_1)\times H^1(\Omega_2)$ toward $(\overline{u},\overline{v})$, the convergence is strong in $H^1(\Omega_1)\times H^1(\Omega_2)$. We can state the following theorem.

Theorem 4.1 *Let* $\Omega=\Omega_1\cup\Omega_2\cup\Gamma$ *be a bounded domain of* $\mathbb{R}^N$ *which can be decomposed in two nonoverlapping subdomains* Ω_1 *and* Ω_2 *of class* C^1 *with a commun interface* Γ*. Assume that* $\mathcal{H}^{N-1}(\Gamma)>0$*. Let* $h\in L^2(\Omega)$ *be such that* $\int_\Omega h=0$*. Then any sequence* (u_n,v_n) *generated by* $(\mathcal{A})$ *strongly converges in* $H^1(\Omega_1)\times H^1(\Omega_2)$ *toward a minimum point of problem* $(\mathcal{P})$.

Notice that the algorithm $(\mathcal{A})$ allows to solve the initial problem $(\mathcal{P})$ on Ω by solving separately Dirichlet-Neumann problems on Ω_1 and Ω_2. Let us describe the dual problem $(\mathcal{P}^*)$ associated to $(\mathcal{P})$ studied in section 3,

$$(\mathcal{P}^*) \qquad \inf\left\{f^*(A^*z^*)+g^*(-B^*z^*)+\frac{\gamma}{2}\|z^*\|^2_{\mathcal{Z}};\quad z^*\in\mathcal{Z}\right\}.$$

In view of the symmetry on f and g, let us focus on f^*.

$$\begin{aligned} f^*(A^*z^*) &= \sup\left\{\langle A^*z^*,u\rangle_{\mathcal{X}}-\frac{1}{2}\int_{\Omega_1}|\nabla u|^2+\int_{\Omega_1}hu;\quad u\in\mathcal{X}\right\} \\ &= -\inf\left\{\frac{1}{2}\int_{\Omega_1}|\nabla u|^2-\int_{\Omega_1}hu-\int_\Gamma z^*Au;\quad u\in\mathcal{X}\right\}. \end{aligned} \quad (13)$$

If $\int_{\Omega_1} h + \int_\Gamma z^* \neq 0$, we easily deduce by using constant functions that $f^*(A^*z^*) = +\infty$. Let us suppose that $\int_{\Omega_1} h + \int_\Gamma z^* = 0$. Let us introduce the bilinear symmetric and continuous form $a : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ defined by $a(u_1, u_2) = \int_{\Omega_1} \nabla u_1 . \nabla u_2$ and the linear continuous form $l : \mathcal{X} \to \mathbb{R}$ defined by $l(u) = \int_{\Omega_1} hu + \int_\Gamma z^* Au$. Let us note $\mathcal{U}$ the closed hyperplane of $\mathcal{X}$ defined by $\mathcal{U} = \left\{ u \in \mathcal{X}; \quad \int_{\Omega_1} u = 0 \right\}$. From the Poincaré-Wirtinger inequality, we deduce that a is coercive on $\mathcal{U} \times \mathcal{U}$ and from the Lax-Milgram theorem that the following minimization problem

$$\inf \left\{ \frac{1}{2} \int_{\Omega_1} |\nabla u|^2 - \int_{\Omega_1} hu - \int_\Gamma z^* Au; \quad u \in \mathcal{U} \right\}$$

has a unique solution $u_{z^*} \in \mathcal{U}$. Moreover u_{z^*} is characterized by

$$\forall k \in \mathcal{U}, \quad \int_{\Omega_1} \nabla u_{z^*} . \nabla k = \int_{\Omega_1} hk + \int_\Gamma z^* Ak.$$

Notice that, for all $k \in \mathcal{X}$, $\left(k - \frac{1}{|\Omega_1|} \int_{\Omega_1} k \right) \in \mathcal{U}$. Hence, using the hypothesis $\int_{\Omega_1} h + \int_\Gamma z^* = 0$, the map u_{z^*} is a solution of the minimization problem (13). Moreover u_{z^*} satisfies

$$\forall k \in \mathcal{X}, \quad \int_{\Omega_1} \nabla u_{z^*} . \nabla k = \int_{\Omega_1} hk + \int_\Gamma z^* Ak. \tag{14}$$

This is a variational weak formulation of the following mixed Dirichlet-Neumann boundary value problem on Ω_1

$$\begin{cases} -\Delta u_{z^*} = h & \text{on } \Omega_1 \\ \frac{\partial u_{z^*}}{\partial n} = 0 & \text{on } \partial\Omega_1 \cap \partial\Omega \\ \frac{\partial u_{z^*}}{\partial n} = z^* & \text{on } \Gamma. \end{cases}$$

From equalities (13)-(14), we infer that

$$\begin{aligned} f^*(A^*z^*) &= -\frac{1}{2} \int_{\Omega_1} |\nabla u_{z^*}|^2 + \int_{\Omega_1} hu_{z^*} + \int_\Gamma z^* Au_{z^*} = \frac{1}{2} \int_{\Omega_1} hu_{z^*} + \frac{1}{2} \int_{\Omega_1} z^* Au_{z^*} \\ &= \frac{1}{2} \int_{\Omega_1} |\nabla u_{z^*}|^2. \end{aligned}$$

In the same way, if $\int_{\Omega_2} h - \int_\Gamma z^* = 0$, we have

$$g^*(-B^*z^*) = \frac{1}{2} \int_{\Omega_2} hv_{z^*} - \frac{1}{2} \int_\Gamma z^* Bv_{z^*} = \frac{1}{2} \int_{\Omega_2} |\nabla v_{z^*}|^2,$$

where v_{z^*} is solution of the following mixed Dirichlet-Neumann boundary value problem on Ω_2

$$\begin{cases} -\Delta v_{z^*} = h & \text{on } \Omega_2 \\ \frac{\partial v_{z^*}}{\partial n} = 0 & \text{on } \partial\Omega_2 \cap \partial\Omega \\ \frac{\partial v_{z^*}}{\partial n} = -z^* & \text{on } \Gamma. \end{cases}$$

Notice that, since $\int_\Omega h = 0$, we have $\int_{\Omega_1} h = -\int_{\Omega_2} h$. Finally problem $(\mathcal{P}^*)$ has the following expression

$$\inf \left\{ \frac{1}{2} \int_\Omega h w_{z^*} + \frac{1}{2} \int_\Gamma z^*[w_{z^*}] + \frac{\gamma}{2} \int_\Gamma z^{*2}; \quad z^* \in \mathcal{Z} \text{ such that } \int_\Gamma z^* + \int_{\Omega_1} h = 0 \right\}$$

where $w_{z^*} = \begin{cases} u_{z^*} \text{ on } \Omega_1 \\ v_{z^*} \text{ on } \Omega_2 \end{cases}$. From Corollary 3.1, the sequence $\left(-\frac{1}{\gamma}(Au_n - Bv_n)\right)$ strongly converges in $\mathcal{Z}$ toward the unique minimizer $\overline{z^*}$ of $(\mathcal{P}^*)$.

Concluding comments and related papers. In this paper, the parameter γ is supposed to be constant in algorithm $(\mathcal{A})$. In the companion paper [4], the authors consider the case of a sequence (γ_n) which decreases toward zero. Under suitable conditions on the decay rate of (γ_n), the associated algorithm minimizes the function $(x, y) \mapsto f(x)+g(y)$ over the space $\mathcal{V} = \{(x, y) \in \mathcal{X}\times\mathcal{Y}, \quad Ax = By\}$. Another situation of interest corresponds to an increasing sequence (γ_n) which tends toward infinity. It is shown in [7] that the associated algorithm minimizes the function $(x, y) \mapsto \frac{1}{2}\|Ax - By\|^2_{\mathcal{Z}}$ over the set $\text{Argmin} f \times \text{Argmin} g$. We refer the reader to [4, 7] for the corresponding applications in domain decomposition for PDE's.

Acknowledgments

The author expresses his gratitude to H. Attouch and A. Cabot for their valuable suggestions and critical comments which made numerous improvements of the paper.

References

[1] F. Acker, M.A. Prestel, Convergence d'un schéma de minimisation alternée, *Annales de la Faculté des Sciences de Toulouse V, série Mathématiques*, 2, (1980), pp. 1-9.

[2] H. Attouch, J. Bolte, P. Redont, A. Soubeyran, Alternating proximal algorithms for weakly coupled convex minimization problems. Applications to dynamical games and PDE's, *Journal of Convex Analysis*, 15 (2008), No 3, pp. 485-506.

[3] H. Attouch, G. Buttazzo, G. Michaille, Variational analysis in Sobolev and BV spaces. Applications to PDE's and optimization, *MPS/SIAM Series on Optimization*, 6. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2006.

[4] H. Attouch, A. Cabot, P. Frankel, J. Peypouquet, Alternating proximal algorithms for constrained variational inequalities. Applications to domain decomposition for PDE's, working paper.

[5] H. Attouch, P. Redont, A. Soubeyran, A new class of alternating proximal minimization algorithms with costs-to-move, *SIAM Journal on Optimization*, 18 (2007), pp. 1061-1081.

[6] H.H. Bauschke, P.-L. Combettes, S. Reich, The asymptotic behavior of the composition of two resolvents, *Nonlinear analysis*, 60 (2005), pp. 283-301.

[7] A. Cabot, P. Frankel, Alternating proximal algorithms with costs-to-move and asymptotically vanishing coupling, submitted to Optimization.

[8] Z. Opial, Weak convergence of the sequence of successive approximations for nonexpansive mappings, *Bull. Amer. Math. Soc.*, 73 (1967), pp. 591-597.

IV. ALGORITHMES PROXIMAUX ALTERNES POUR INEGALITES VARIATIONNELLES AVEC CONTRAINTES LINEAIRES. APPLICATION A LA DECOMPOSITION DE DOMAINES POUR LES EDP

Alternating proximal algorithms for linearly constrained variational inequalities. Application to domain decomposition for PDE's

H. Attouch[1], A. Cabot[2], P. Frankel[3], and J. Peypouquet[4]

[1] Université Montpellier II, Place Eugène Bataillon, 34095 Montpellier Cedex 05, France
attouch@math.univ-montp2.fr
[2] Université Montpellier II, Place Eugène Bataillon, 34095 Montpellier Cedex 05, France
acabot@math.univ-montp2.fr
[3] Université Montpellier II, Place Eugène Bataillon, 34095 Montpellier Cedex 05, France
p.frankel30@orange.fr
[4] Departamento de Matemática, Universidad Técnica Federico Santa María, Avenida España 1680, Valparaíso, Chile
juan.peypouquet@usm.cl

Accepted in ***Nonlinear Analysis***

Summary. Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ be real Hilbert spaces, let $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$, $g : \mathcal{Y} \to \mathbb{R} \cup \{+\infty\}$ be closed convex functions and let $A : \mathcal{X} \to \mathcal{Z}$, $B : \mathcal{Y} \to \mathcal{Z}$ be linear continuous operators. Let us consider the constrained minimization problem

$$(\mathcal{P}) \qquad \min\{f(x) + g(y) : \quad Ax = By\}.$$

Given a sequence (γ_n) which tends toward 0 as $n \to +\infty$, we study the following alternating proximal algorithm

$$(\mathcal{A}) \qquad \begin{cases} x_{n+1} = \text{Argmin}\left\{\gamma_{n+1}\, f(\zeta) + \dfrac{1}{2}\|A\zeta - By_n\|^2_{\mathcal{Z}} + \dfrac{\alpha}{2}\|\zeta - x_n\|^2_{\mathcal{X}};\ \ \zeta \in \mathcal{X}\right\} \\ y_{n+1} = \text{Argmin}\left\{\gamma_{n+1}\, g(\eta) + \dfrac{1}{2}\|Ax_{n+1} - B\eta\|^2_{\mathcal{Z}} + \dfrac{\nu}{2}\|\eta - y_n\|^2_{\mathcal{Y}};\ \ \eta \in \mathcal{Y}\right\}, \end{cases}$$

where α and ν are positive parameters. It is shown that if the sequence (γ_n) tends *moderately slowly* toward 0, then the iterates of $(\mathcal{A})$ weakly converge toward a solution of $(\mathcal{P})$. The study is extended to the setting of maximal monotone operators, for which a general ergodic convergence result is obtained. Applications are given in the area of domain decomposition for PDE's.

Key words: Convex minimization, alternating minimization, proximal algorithm, variational inequalities, monotone inclusions, domain decomposition for PDE's.

Subject classification: 65K05, 65K10, 49J40, 90C25.

1 Introduction

Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ be real Hilbert spaces respectively endowed with the scalar products $\langle ., . \rangle_{\mathcal{X}}$, $\langle ., . \rangle_{\mathcal{Y}}$ and $\langle ., . \rangle_{\mathcal{Z}}$ and the corresponding norms. Let $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$,

$g : \mathcal{Y} \to \mathbb{R} \cup \{+\infty\}$ be closed convex proper functions and let $A : \mathcal{X} \to \mathcal{Z}$, $B : \mathcal{Y} \to \mathcal{Z}$ be linear continuous operators. In this study, our aim is to solve convex structured minimization problems of the form

$$(\mathcal{P}) \qquad \min\{f(x) + g(y) : \quad Ax = By\}.$$

In order to find a point that minimizes the map $(x, y) \mapsto \Phi(x, y) = f(x)+g(y)$ on the subspace $\{(x, y) \in \mathcal{X} \times \mathcal{Y}, \quad Ax = By\}$ we propose the following alternating algorithm:

$$(\mathcal{A}) \quad \begin{cases} x_{n+1} = \text{Argmin}\Big\{\gamma_{n+1}\, f(\zeta) + \dfrac{1}{2}\|A\zeta - By_n\|^2_{\mathcal{Z}} + \dfrac{\alpha}{2}\|\zeta - x_n\|^2_{\mathcal{X}};\ \zeta \in \mathcal{X}\Big\} \\ y_{n+1} = \text{Argmin}\Big\{\gamma_{n+1}\, g(\eta) + \dfrac{1}{2}\|Ax_{n+1} - B\eta\|^2_{\mathcal{Z}} + \dfrac{\nu}{2}\|\eta - y_n\|^2_{\mathcal{Y}};\ \eta \in \mathcal{Y}\Big\}, \end{cases}$$

where α, ν are positive real numbers and (γ_n) is a positive sequence that tends[5] toward 0 as $n \to +\infty$. Due to the structured character of the objective function $\Phi(x, y) = f(x)+g(y)$, alternating algorithms imply a reduction on the size of the subproblems to be solved at each iteration. Our particular choice of $(\mathcal{A})$ is based on the following ideas:

a) Alternating algorithms with costs-to-move. Consider the convex function $\Phi_\gamma : \mathcal{X} \times \mathcal{Y} \to \mathbb{R} \cup \{+\infty\}$ defined by

$$\Phi_\gamma(x, y) = f(x) + g(y) + \frac{1}{2\gamma}\|Ax - By\|^2_{\mathcal{Z}},$$

where γ is a positive real parameter. The minimization of the function Φ_γ is studied in [13], where the authors introduce the alternating algorithm with costs-to-move

$$\begin{cases} x_{n+1} = \text{Argmin}\{f(\zeta) + \dfrac{1}{2\gamma}\|A\zeta - By_n\|^2_{\mathcal{Z}} + \dfrac{\alpha}{2}\|\zeta - x_n\|^2_{\mathcal{X}}; \quad \zeta \in \mathcal{X}\} \\ y_{n+1} = \text{Argmin}\{g(\eta) + \dfrac{1}{2\gamma}\|Ax_{n+1} - B\eta\|^2_{\mathcal{Z}} + \dfrac{\nu}{2}\|\eta - y_n\|^2_{\mathcal{Y}}; \quad \eta \in \mathcal{Y}\}, \end{cases}$$

α and ν being positive coefficients. If $\text{Argmin}\Phi_\gamma \neq \varnothing$, it is shown in [7] that the sequence (x_n, y_n) converges weakly toward a minimum of Φ_γ. The framework of [7, 13] extends the one of [1, 19] from the strong coupled problem to the weak coupled problem with costs-to-change. More precisely, $Q(x, y) = \|x - y\|^2_{\mathcal{Z}}$ is a strong coupling function with $\mathcal{X} = \mathcal{Y} = \mathcal{Z}$ and $A = B = \mathcal{I}$ while $Q(x, y) = \|Ax - By\|^2_{\mathcal{Z}}$ is now a weak coupling function which allows for asymmetric and partial relations between the variables x and y. The interest of the weak coupling term is to cover many situations, ranging from decomposition methods for PDE's

[5] In another direction, algorithm $(\mathcal{A})$ has been recently studied in [23] in the case of a sequence (γ_n) increasing toward $+\infty$ as $n \to +\infty$.

to applications in game theory. In decision sciences, the term $Q(x,y) = \|Ax - By\|_{\mathcal{Z}}^2$ allows to consider agents who interplay, only via some components of their decision variables. For further details, the interested reader is referred to [7].

Observing that problem $(\mathcal{P})$ corresponds formally to the minimization of the function Φ_γ with $\gamma \to 0$, it is natural to consider a vanishing sequence (γ_n) in algorithm $(\mathcal{A})$.

b) Prox-penalization methods. Setting $\Psi(x,y) = \frac{1}{2}\|Ax - By\|_{\mathcal{Z}}^2$ and $\mathbf{x} = (x,y) \in \mathcal{X} \times \mathcal{Y} = \boldsymbol{\mathcal{X}}$, we can rewrite problem $(\mathcal{P})$ as

$$\min\{\ \Phi(\mathbf{x}) :\ \mathbf{x} \in \operatorname{Argmin}\Psi\ \}.$$

This situation is studied in [12, 22], where the authors use a diagonal proximal point algorithm combined with a penalization scheme. This kind of technique can be traced back to the pioneering work [16]. The algorithm of [12, 22] applied to our setting reads as

$$(\mathcal{A}') \qquad \mathbf{x}_{n+1} \in \operatorname{Argmin}\left\{\gamma_n\, \Phi(\mathbf{x}) + \Psi(\mathbf{x}) + \frac{1}{2}\|\mathbf{x} - \mathbf{x}_n\|_{\boldsymbol{\mathcal{X}}}^2\right\}.$$

Under suitable conditions on the sequence (γ_n), it is shown in [12, 22] that the iterates of algorithm $(\mathcal{A}')$ converge weakly to a solution of $(\mathcal{P})$.

These ideas lead us to the formulation of algorithm $(\mathcal{A})$, which has the following distinctive marks. First, it uses the structured character of the objective function to reduce the size of the subproblem solved at each iteration. Second, it combines proximal iterations with a penalization scheme in a simple way, meaning that no new nonlinearities are introduced by the latter, unlike most penalization procedures available in the literature. Consider, for instance, the functions θ described in [24] (see also the references therein).

The main result of the paper asserts that if the solution set is nonempty and if (γ_n) tends *moderately slowly* toward 0, then the iterates of $(\mathcal{A})$ weakly converge toward a solution of $(\mathcal{P})$. When the space $R(A) + R(B)$ is closed in $\mathcal{Z}$, the above condition on (γ_n) is satisfied if the sequence $\left(\frac{1}{\gamma_{n+1}} - \frac{1}{\gamma_n}\right)$ is bounded from above and if $(\gamma_n) \in l^2$.

We apply our abstract results to the framework of splitting methods for PDE's. For that purpose, we consider a domain $\Omega \subset \mathbb{R}^N$ that can be decomposed into two non overlapping subdomains Ω_1, Ω_2 with a common interface Γ. The functional spaces are $\mathcal{X} = H^1(\Omega_1)$, $\mathcal{Y} = H^1(\Omega_2)$ and $\mathcal{Z} = L^2(\Gamma)$, the operators $A : \mathcal{X} \to \mathcal{Z}$ and $B : \mathcal{Y} \to \mathcal{Z}$ being respectively the trace operators on Γ. The term $Au - Bv$ corresponds to the jump of the map $w = \begin{cases} u \text{ on } \Omega_1 \\ v \text{ on } \Omega_2 \end{cases}$ through the interface Γ.

It is shown that algorithm $(\mathcal{A})$ allows to solve some given boundary value problem on Ω by solving separately mixed Dirichlet-Neumann problems on Ω_1 and Ω_2.

Finally observe that by writing down the optimality conditions satisfied by the iterates of algorithm $(\mathcal{A})$, we obtain

$$\begin{cases} 0 \in \gamma_{n+1}\,\partial f(x_{n+1}) + \ \ A^*(Ax_{n+1} - By_n) \ \ + \alpha(x_{n+1} - x_n) \\ 0 \in \gamma_{n+1}\,\partial g(y_{n+1}) - B^*(Ax_{n+1} - By_{n+1}) + \nu(y_{n+1} - y_n). \end{cases}$$

This suggests to extend the previous study to the framework of maximal monotone operators, by replacing respectively the subdifferential operators ∂f and ∂g with two maximal monotone operators M and N. Indeed, in this more general setting we are able to prove the convergence of the sequence of weighted averages.

The paper is organized as follows. Section 2 is devoted to fix the general setting and notations that are used throughout the paper. In section 3, we prove a general result of weak ergodic convergence for the iterates of $(\mathcal{A})$ in a maximal monotone setting. The key conditions are the closedness of the space $R(A) + R(B)$ and the assumption $(\gamma_n) \in l^2 \setminus l^1$. The subdifferential case is analyzed in section 4, where we establish a result of weak convergence toward a solution of $(\mathcal{P})$. We also discuss on the robustness of the algorithm with respect to computational errors. Section 5 presents further convergence results for the strongly coupled problem without cost-to-move. Finally, the applications to domain decomposition for PDE's are illustrated in section 6.

2 General setting and notations

We recall that $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ are real Hilbert spaces respectively endowed with the scalar products $\langle ., . \rangle_{\mathcal{X}}$, $\langle ., . \rangle_{\mathcal{Y}}$, $\langle ., . \rangle_{\mathcal{Z}}$ and the corresponding norms. Let $M : \mathcal{X} \rightrightarrows \mathcal{X}$, $N : \mathcal{Y} \rightrightarrows \mathcal{Y}$ be maximal monotone operators such that $\mathrm{dom} M \neq \emptyset$, $\mathrm{dom} N \neq \emptyset$. Let $A : \mathcal{X} \to \mathcal{Z}$, $B : \mathcal{Y} \to \mathcal{Z}$ be linear continuous operators with adjoints $A^* : \mathcal{Z} \to \mathcal{X}$ and $B^* : \mathcal{Z} \to \mathcal{Y}$. Let (γ_n) be a positive sequence such that $\lim_{n\to+\infty} \gamma_n = 0$. Given positive coefficients $\alpha, \nu > 0$ and initial data $(x_0, y_0) \in \mathcal{X} \times \mathcal{Y}$, let us consider the alternating proximal algorithm defined implicitly by

$$(\mathcal{A}) \qquad \begin{cases} 0 \in \gamma_{n+1}\, Mx_{n+1} + \ \ A^*(Ax_{n+1} - By_n) \ \ + \alpha(x_{n+1} - x_n) \\ 0 \in \gamma_{n+1}\, Ny_{n+1} - B^*(Ax_{n+1} - By_{n+1}) + \nu(y_{n+1} - y_n). \end{cases}$$

Observe that the linear continuous operator A^*A is maximal monotone, hence the operator $\gamma_{n+1}\, M + A^*A$ is also maximal monotone, see for example [20]. Therefore the iterate x_{n+1} is uniquely defined by Minty's theorem. The same holds true for the iterate y_{n+1}.

Remark 2.1 (Strong coupling without cost-to-move) *Assume that $\mathcal{X} = \mathcal{Y} = \mathcal{Z}$ and that $A = B = \mathcal{I}$, which corresponds to a situation of strong coupling. In this case, algorithm $(\mathcal{A})$ is well-defined even if $\alpha = \nu = 0$. We denote by $(\mathcal{A}_0)$ the corresponding algorithm*

$$(\mathcal{A}_0) \qquad \begin{cases} 0 \in \gamma_{n+1}\, M x_{n+1} + x_{n+1} - y_n \\ 0 \in \gamma_{n+1}\, N y_{n+1} + y_{n+1} - x_{n+1}, \end{cases}$$

that can be equivalently rewritten as

$$\begin{cases} x_{n+1} = (I + \gamma_{n+1}\, M)^{-1}\, y_n \\ y_{n+1} = (I + \gamma_{n+1}\, N)^{-1}\, x_{n+1}. \end{cases}$$

It ensues that the sequences (x_n) and (y_n) satisfy the following recurrence formulae

$$x_{n+1} = (I+\gamma_{n+1}\, M)^{-1}(I+\gamma_n\, N)^{-1}\, x_n, \quad y_{n+1} = (I+\gamma_{n+1}\, N)^{-1}(I+\gamma_{n+1}\, M)^{-1}\, y_n.$$

This scheme consisting of a double backward step has been previously studied by Passty [31]. Algorithm $(\mathcal{A})$ can be viewed as an extension of iteration $(\mathcal{A}_0)$, so that our present paper appears as a continuation of the seminal work [31].

Let $\boldsymbol{\mathcal{X}} = \mathcal{X} \times \mathcal{Y}$ and denote by $\boldsymbol{\mathcal{V}}$ the closed subspace $\{(x, y) \in \boldsymbol{\mathcal{X}},\ Ax = By\}$. The normal cone operator $N_{\boldsymbol{\mathcal{V}}}$ takes the constant value $N_{\boldsymbol{\mathcal{V}}} \equiv \boldsymbol{\mathcal{V}}^\perp$ on its domain $\boldsymbol{\mathcal{V}}$. Setting $\mathbf{x} = (x, y)$, define the monotone operators $\mathbf{M} : \boldsymbol{\mathcal{X}} \rightrightarrows \boldsymbol{\mathcal{X}}$ and $\mathbf{T} : \boldsymbol{\mathcal{X}} \rightrightarrows \boldsymbol{\mathcal{X}}$ respectively by

$$\mathbf{Mx} = (Mx, Ny)$$

and

$$\mathbf{Tx} = \mathbf{Mx} + N_{\boldsymbol{\mathcal{V}}}(\mathbf{x}) = \begin{cases} \mathbf{Mx} + \boldsymbol{\mathcal{V}}^\perp & \text{if } \mathbf{x} \in \boldsymbol{\mathcal{V}} \\ \quad \emptyset & \text{if } \mathbf{x} \notin \boldsymbol{\mathcal{V}}. \end{cases}$$

We denote by $\boldsymbol{\mathcal{S}} = \mathbf{T}^{-1}0$ the null set of $\mathbf{T}$. It is also convenient to define the bounded linear operator

$$\begin{aligned} \mathbf{A} :\ \ \boldsymbol{\mathcal{X}} \ &\rightarrow \ \ \mathcal{Z} \\ (x, y) &\mapsto Ax - By, \end{aligned}$$

and the map

$$\begin{aligned} \Psi :\ \ \boldsymbol{\mathcal{X}} \ &\rightarrow \ \ \mathbb{R} \\ (x, y) &\mapsto \tfrac{1}{2}\|Ax - By\|^2_{\mathcal{Z}}. \end{aligned}$$

Recall that the Fenchel conjugate $\Psi^* : \boldsymbol{\mathcal{X}} \to \mathbb{R} \cup \{+\infty\}$ of the map Ψ is defined by $\Psi^*(\mathbf{p}) = \sup_{\mathbf{x} \in \boldsymbol{\mathcal{X}}} \{\langle \mathbf{p}, \mathbf{x} \rangle_{\boldsymbol{\mathcal{X}}} - \Psi(\mathbf{x})\}$ for every $\mathbf{p} \in \boldsymbol{\mathcal{X}}$. The next proposition shows that $\mathrm{dom}\Psi^* = R(\mathbf{A}^*)$ and gives the expression of the function Ψ^* on its domain.

Proposition 2.1 *With the same notations as above, we have $dom\Psi^* = R(\mathbf{A}^*)$ and $\Psi^*(\mathbf{A}^* z) = \frac{1}{2} d^2_{\mathcal{Z}}(z,\ Ker(\mathbf{A}^*))$ for every $z \in \mathcal{Z}$.*

Proof. Let us fix $\mathbf{p} \in \boldsymbol{\mathcal{X}}$. From the definition of Ψ and Ψ^*, we have $\Psi^*(\mathbf{p}) = \sup_{\mathbf{x}\in\boldsymbol{\mathcal{X}}} \left\{\langle \mathbf{p}, \mathbf{x}\rangle_{\boldsymbol{\mathcal{X}}} - \frac{1}{2}\|\mathbf{A}\mathbf{x}\|^2_{\mathcal{Z}}\right\}$. This maximization problem can be reformulated as

$$- \inf_{\mathbf{x}\in\boldsymbol{\mathcal{X}}} \{F(\mathbf{x}) + G(\mathbf{A}\mathbf{x})\}, \tag{1}$$

where $F : \boldsymbol{\mathcal{X}} \to \mathbb{R}$ and $G : \mathcal{Z} \to \mathbb{R}$ are respectively defined by $F(\mathbf{x}) = -\langle \mathbf{p}, \mathbf{x}\rangle_{\boldsymbol{\mathcal{X}}}$ and $G(z) = \frac{1}{2}\|z\|^2_{\mathcal{Z}}$ for every $\mathbf{x} \in \boldsymbol{\mathcal{X}}$, $z \in \mathcal{Z}$. Let us introduce the following minimization problem

$$\inf_{z^*\in\mathcal{Z}} \{F^*(-\mathbf{A}^*z^*) + G^*(z^*)\} = \inf_{z^*\in\mathcal{Z}} \left\{\delta_{\{-\mathbf{p}\}}(-\mathbf{A}^*z^*) + \frac{1}{2}\|z^*\|^2_{\mathcal{Z}}\right\} \tag{2}$$

$$= \inf_{\substack{z^* \in \mathcal{Z} \\ \mathbf{A}^*z^* = \mathbf{p}}} \frac{1}{2}\|z^*\|^2_{\mathcal{Z}}. \tag{3}$$

Since the functions F and G are convex and continuous, problems (1)-(2) are dual each to other, see for example [25, Chap. III]. Observing that the Moreau-Rockafellar qualification condition is satisfied, we derive from [25, Theorem 4.1, p. 59] that the infimum values of problems (1)-(2) are simultaneously finite and in this case they coincide. Expression (3) shows that the infimum in (2) is finite if and only if $\mathbf{p} \in R(\mathbf{A}^*)$. Coming back to problem (1), we deduce that $\mathbf{p} \in \mathrm{dom}\Psi^*$ if and only if $\mathbf{p} \in R(\mathbf{A}^*)$. Now assume that $\mathbf{p} = \mathbf{A}^*z$ for some $z \in \mathcal{Z}$. Then we have

$$\inf_{\substack{z^* \in \mathcal{Z} \\ \mathbf{A}^*z^* = \mathbf{p}}} \frac{1}{2}\|z^*\|^2_{\mathcal{Z}} = \inf_{\substack{z^* \in \mathcal{Z} \\ z^* - z \in \mathrm{Ker}(\mathbf{A}^*)}} \frac{1}{2}\|z^*\|^2_{\mathcal{Z}} = \frac{1}{2} d^2_{\mathcal{Z}}\left(z,\ \mathrm{Ker}(\mathbf{A}^*)\right),$$

which ends the proof.

3 Maximal monotone framework: ergodic convergence results

The notations and hypotheses are the same as in the previous section. Given any initial point $(x_0, y_0) \in \boldsymbol{\mathcal{X}}$, the iterates generated by algorithm $(\mathcal{A})$ are denoted by (x_n, y_n), $n \in \mathbb{N}$.

3.1 Preliminary results

Let us start with an estimation that is at the core of the convergence analysis. For $(x, y) \in \boldsymbol{\mathcal{X}}$ set

$$h_n(x, y) = \alpha\|x_n - x\|^2_{\mathcal{X}} + \nu\|y_n - y\|^2_{\mathcal{Y}} + \|By_n - By\|^2_{\mathcal{Z}}. \tag{4}$$

Then we have the following:

Lemma 3.1 *For every* $(x,y)\in\boldsymbol{\mathcal{V}}$ *and* $(\zeta,\eta)\in\mathbf{T}(x,y)$*, there exists* $\mathbf{p}\in\boldsymbol{\mathcal{V}}^\perp$ *such that*

$$h_{n+1}(x,y)-h_n(x,y)+2\gamma_{n+1}\Big[\langle\zeta,x_{n+1}-x\rangle_{\mathcal{X}}+\langle\eta,y_{n+1}-y\rangle_{\mathcal{Y}}\Big]$$
$$+\,\alpha\|x_{n+1}-x_n\|^2_{\mathcal{X}}+\nu\|y_{n+1}-y_n\|^2_{\mathcal{Y}}+\|Ax_{n+1}-By_n\|^2_{\mathcal{Z}}\le 2\,\gamma^2_{n+1}\Psi^*(\mathbf{p}). \qquad (5)$$

Proof. To simplify the notation we set $h_n=h_n(x,y)$. The definition of (x_{n+1}) gives

$$\frac{\alpha}{\gamma_{n+1}}(x_{n+1}-x_n)+\frac{1}{\gamma_{n+1}}A^*(Ax_{n+1}-By_n)\in -Mx_{n+1}.$$

On the other hand since $(\zeta,\eta)\in\mathbf{T}(x,y)$, there exists $\mathbf{p}=(p,q)\in\boldsymbol{\mathcal{V}}^\perp$ such that

$$\zeta\in Mx+p \quad\text{ and }\quad \eta\in Ny+q.$$

In particular, we have $p-\zeta\in -Mx$, which by the monotonicity of M implies

$$\frac{\alpha}{\gamma_{n+1}}\langle x_{n+1}-x_n,x_{n+1}-x\rangle_{\mathcal{X}}+\frac{1}{\gamma_{n+1}}\langle A^*(Ax_{n+1}-By_n),x_{n+1}-x\rangle_{\mathcal{X}}\le\langle p-\zeta,x_{n+1}-x\rangle_{\mathcal{X}}.$$

This is equivalent to

$$\alpha\|x_{n+1}-x\|^2_{\mathcal{X}}+\alpha\|x_{n+1}-x_n\|^2_{\mathcal{X}}\le\alpha\|x_n-x\|^2_{\mathcal{X}}-2\langle Ax_{n+1}-By_n,Ax_{n+1}-Ax\rangle_{\mathcal{Z}}$$
$$+\,2\gamma_{n+1}\langle p,x_{n+1}-x\rangle_{\mathcal{X}}-2\gamma_{n+1}\langle\zeta,x_{n+1}-x\rangle_{\mathcal{X}}.$$

In a similar way we obtain

$$\nu\|y_{n+1}-y\|^2_{\mathcal{Y}}+\nu\|y_{n+1}-y_n\|^2_{\mathcal{Y}}\le\nu\|y_n-y\|^2_{\mathcal{Y}}-2\langle By_{n+1}-Ax_{n+1},By_{n+1}-By\rangle_{\mathcal{Z}}$$
$$+\,2\gamma_{n+1}\langle q,y_{n+1}-y\rangle_{\mathcal{Y}}-2\gamma_{n+1}\langle\eta,y_{n+1}-y\rangle_{\mathcal{Y}}.$$

Using the properties of the inner product and the fact that $Ax=By$, we let the reader check that

$$-2\langle Ax_{n+1}-By_n,Ax_{n+1}-Ax\rangle_{\mathcal{Z}}-2\langle By_{n+1}-Ax_{n+1},By_{n+1}-By\rangle_{\mathcal{Z}}=$$
$$\|By_n-By\|^2_{\mathcal{Z}}-\|By_{n+1}-By\|^2_{\mathcal{Z}}-\|Ax_{n+1}-By_n\|^2_{\mathcal{Z}}-\|Ax_{n+1}-By_{n+1}\|^2_{\mathcal{Z}}.$$

Since $(x,y)\in\boldsymbol{\mathcal{V}}$ and $\mathbf{p}=(p,q)\in\boldsymbol{\mathcal{V}}^\perp$, we have

$$\langle p,x_{n+1}-x\rangle_{\mathcal{X}}+\langle q,y_{n+1}-y\rangle_{\mathcal{Y}}=\langle p,x_{n+1}\rangle_{\mathcal{X}}+\langle q,y_{n+1}\rangle_{\mathcal{Y}}=\langle\mathbf{p},(x_{n+1},y_{n+1})\rangle_{\boldsymbol{\mathcal{X}}}.$$

Gathering all this information and writing

$$c_n=h_{n+1}-h_n+2\gamma_{n+1}\left[\langle\zeta,x_{n+1}-x\rangle_{\mathcal{X}}+\langle\eta,y_{n+1}-y\rangle_{\mathcal{Y}}\right]$$
$$+\;\alpha\|x_{n+1}-x_n\|^2_{\mathcal{X}}+\nu\|y_{n+1}-y_n\|^2_{\mathcal{Y}}+\|Ax_{n+1}-By_n\|^2_{\mathcal{Z}}$$

we deduce that

$$c_n\le 2\gamma_{n+1}\langle\mathbf{p},(x_{n+1},y_{n+1})\rangle_{\boldsymbol{\mathcal{X}}}-\|Ax_{n+1}-By_{n+1}\|^2_{\mathcal{Z}}$$
$$=2\left[\langle\gamma_{n+1}\mathbf{p},(x_{n+1},y_{n+1})\rangle_{\boldsymbol{\mathcal{X}}}-\Psi(x_{n+1},y_{n+1})\right].$$

By definition of Ψ^*, the term between brackets is majorized by $\Psi^*(\gamma_{n+1}\mathbf{p})$. Since $\Psi^*(\gamma_{n+1}\mathbf{p})=\gamma^2_{n+1}\Psi^*(\mathbf{p})$, inequality (5) immediately follows.

In order to exploit inequality (5), we may assume that $\Psi^*(\mathbf{p}) < +\infty$ for every $\mathbf{p} \in \boldsymbol{\mathcal{V}}^\perp$. In view of Proposition 2.1, this amounts to saying that $\boldsymbol{\mathcal{V}}^\perp \subset \mathrm{dom}\Psi^* = R(\mathbf{A}^*)$. Since $\boldsymbol{\mathcal{V}}^\perp = \mathrm{Ker}(\mathbf{A})^\perp = \overline{R(\mathbf{A}^*)}$, this condition is equivalent to the closedness of the space $R(\mathbf{A}^*)$, which is in turn equivalent to the closedness of $R(\mathbf{A})$ in $\mathcal{Z}$. From now on, we assume in this section that

$$R(\mathbf{A}) = R(A) + R(B) \quad \text{ is closed in } \mathcal{Z}.$$

By using Lemma 3.1, we now prove the boundedness of the sequence (x_n, y_n) along with the summability of the sequences $(\|x_{n+1} - x_n\|^2_{\mathcal{X}})$, $(\|y_{n+1} - y_n\|^2_{\mathcal{Y}})$ and $(\|Ax_n - By_n\|^2_{\mathcal{Z}})$.

Proposition 3.1 *Assume that the space $R(\mathbf{A})$ is closed in $\mathcal{Z}$ and that $(\gamma_n) \in l^2$. Suppose that the set $\boldsymbol{\mathcal{S}}$ is nonempty and let $(\overline{x}, \overline{y}) \in \boldsymbol{\mathcal{S}}$. We have the following*

(i) $\lim_{n\to+\infty} h_n(\overline{x}, \overline{y})$ exists, hence the sequence (x_n, y_n) is bounded.
(ii)The sequences $(\|x_{n+1} - x_n\|^2_{\mathcal{X}})$, $(\|y_{n+1} - y_n\|^2_{\mathcal{Y}})$ and $(\|Ax_n - By_n\|^2_{\mathcal{Z}})$ are summable. In particular,

$$\lim_{n\to+\infty} \|x_{n+1} - x_n\|_{\mathcal{X}} = \lim_{n\to+\infty} \|y_{n+1} - y_n\|_{\mathcal{Y}} = \lim_{n\to+\infty} \|Ax_n - By_n\|_{\mathcal{Z}} = 0 \quad (6)$$

and every weak cluster point of the sequence (x_n, y_n) lies in $\boldsymbol{\mathcal{V}}$.

Proof. (i) Taking $(\zeta, \eta) = (0, 0)$ in inequality (5) and setting $h_n = h_n(\overline{x}, \overline{y})$, we obtain

$$h_{n+1} - h_n + \alpha\|x_{n+1} - x_n\|^2_{\mathcal{X}} + \nu\|y_{n+1} - y_n\|^2_{\mathcal{Y}} + \|Ax_{n+1} - By_n\|^2_{\mathcal{Z}} \leq 2\gamma^2_{n+1}\, \Psi^*(\mathbf{p}). \quad (7)$$

In particular, $h_{n+1} - h_n \leq 2\gamma^2_{n+1}\, \Psi^*(\mathbf{p})$. Since $(\gamma_n) \in l^2$ and $\Psi^*(\mathbf{p}) < +\infty$, the following lemma shows that (h_n) converges, which in turn implies that the sequence (x_n, y_n) is bounded.

Lemma 3.2 *Let (a_n) and (ε_n) be two real sequences. Assume that (a_n) is minorized, that $(\varepsilon_n) \in l^1$ and that $a_{n+1} \leq a_n + \varepsilon_n$ for every $n \in \mathbb{N}$. Then (a_n) converges.*

(ii) Let us sum up inequality (7) from $n = 0$ to $+\infty$. Recalling that $(\gamma_n) \in l^2$, that $\Psi^*(\mathbf{p}) < +\infty$ and that $h_n \geq 0$, we immediately deduce the summability of the sequences $(\|x_{n+1} - x_n\|^2_{\mathcal{X}})$, $(\|y_{n+1} - y_n\|^2_{\mathcal{Y}})$ and $(\|Ax_{n+1} - By_n\|^2_{\mathcal{Z}})$. Since $\|Ax_n - By_n\|^2_{\mathcal{Z}} \leq 2\|Ax_{n+1} - By_n\|^2_{\mathcal{Z}} + 2\|Ax_{n+1} - Ax_n\|^2_{\mathcal{Z}}$, the sequence $(\|Ax_n - By_n\|^2_{\mathcal{Z}})$ is also summable. For the last part use the fact that $\lim_{n\to+\infty} \|Ax_n - By_n\|^2_{\mathcal{Z}} = 0$ and the weak lower-semicontinuity of the function $(x, y) \mapsto \|Ax - By\|^2_{\mathcal{Z}}$.

Remark 3.1 *Proposition 3.1 still holds if one assumes only that $\overline{R(\mathbf{A}^*)} \cap R(\mathbf{M}) \subset R(\mathbf{A}^*)$, a condition that is weaker than the closedness of $R(\mathbf{A})$. The reason is that one uses Lemma 3.1 for $(x, y) = (\overline{x}, \overline{y}) \in \boldsymbol{\mathcal{S}}$.*

3.2 Ergodic convergence

From now on we assume that $(\gamma_n) \in l^2 \setminus l^1$. Condition $(\gamma_n) \notin l^1$ is standard and common to most proximal-type algorithms[6]. For practical purposes it states that the sequence (γ_n) does not tend to 0 too fast as $n \to +\infty$. The condition $(\gamma_n) \in l^2 \setminus l^1$ expresses that the sequence (γ_n) tends *moderately* slowly toward 0. This kind of assumption appears in several works related to proximal algorithms involving maximal monotone operators with alternating features. See for example the seminal work [31] (or also [12] and [22]). In some particular cases it is possible to obtain convergence of our algorithm under the sole assumption that $(\gamma_n) \notin l^1$ (see Remark 5.1) but this relies strongly on the geometry of the problem. Finding the general conditions for convergence when $(\gamma_n) \notin l^2$ is an interesting and challenging open question.

Let us define the averages

$$\widetilde{x}_n = \frac{1}{\sigma_n} \sum_{k=1}^{n} \gamma_k x_k \quad \text{and} \quad \widetilde{y}_n = \frac{1}{\sigma_n} \sum_{k=1}^{n} \gamma_k y_k, \tag{8}$$

and prove that the sequence $(\widetilde{x}_n, \widetilde{y}_n)$ converges weakly to a point in $\boldsymbol{\mathcal{S}}$.

Theorem 3.1 *Assume that the space $R(\mathbf{A})$ is closed in $\mathcal{Z}$ and that $(\gamma_n) \in l^2 \setminus l^1$. Assume moreover that the operator $\mathbf{T}$ is maximal monotone with $\boldsymbol{\mathcal{S}} = \mathbf{T}^{-1}0 \neq \emptyset$. Then the sequence $(\widetilde{x}_n, \widetilde{y}_n)$ of averages converges weakly as $n \to +\infty$ to a point in $\boldsymbol{\mathcal{S}}$.*

Proof. Let us first prove that every weak cluster point of the sequence $(\widetilde{x}_n, \widetilde{y}_n)$ is in $\boldsymbol{\mathcal{S}}$. Fix $(\zeta, \eta) \in \mathbf{T}(x, y)$. By summing up inequality (5) of Lemma 3.1 for $k = 0, \dots, n-1$, we obtain

$$\langle \zeta, \widetilde{x}_n - x \rangle_{\mathcal{X}} + \langle \eta, \widetilde{y}_n - y \rangle_{\mathcal{Y}} \le \frac{1}{2\sigma_n} \left[h_0(x, y) + 2\,\Psi^*(\mathbf{p}) \sum_{k=1}^{n} \gamma_k^2 \right].$$

Let $(\widetilde{x}_\infty, \widetilde{y}_\infty)$ be a weak cluster point of $(\widetilde{x}_n, \widetilde{y}_n)$ as $n \to +\infty$ and let n tend to $+\infty$ in the above inequality. By using the fact that $\Psi^*(\mathbf{p}) < +\infty$ and that $(\gamma_n) \in l^2 \setminus l^1$, we deduce that

$$\langle \zeta, \widetilde{x}_\infty - x \rangle_{\mathcal{X}} + \langle \eta, \widetilde{y}_\infty - y \rangle_{\mathcal{Y}} \le 0. \tag{9}$$

Since this holds whenever $(\zeta, \eta) \in \mathbf{T}(x, y)$ we conclude $(\widetilde{x}_\infty, \widetilde{y}_\infty) \in \boldsymbol{\mathcal{S}}$ by maximality of the monotone operator $\mathbf{T}$.

Now observe that the sequence $(\widetilde{x}_n, \widetilde{y}_n)$ is bounded by Proposition 3.1. In order to establish the weak convergence of the sequence $(\widetilde{x}_n, \widetilde{y}_n)$ it suffices to

[6] When the proximal point algorithm is seen as a discretization of the differential inclusion $-\dot{x}(t) \in Ax(t)$, the partial sums $\sigma_n = \sum_{k=1}^{n} \gamma_k$ have a natural interpretation as *discrete times*. In this setting, the condition $(\gamma_n) \notin l^1$ is an analogue for $t \to +\infty$.

prove that it has at most one weak cluster point[7]. Indeed, let (x,y) and (x',y') be two such points, which must belong to $\boldsymbol{\mathcal{S}}$. Define the quantity $Q(u,v) = \alpha\|u\|^2_{\mathcal{X}} + \nu\|v\|^2_{\mathcal{Y}} + \|Bv\|^2_{\mathcal{Z}}$ for every $(u,v) \in \boldsymbol{\mathcal{X}}$. From Proposition 3.1 (i), the limits

$$\ell(x,y) = \lim_{n\to+\infty} Q(x_n - x, y_n - y) \qquad \text{and} \qquad \ell(x',y') = \lim_{n\to+\infty} Q(x_n - x', y_n - y')$$

exist. Observe that

$$\begin{aligned} Q(x_n - x, y_n - y) &= Q(x_n - x', y_n - y') + Q(x - x', y - y') \\ &\quad + 2\,\alpha\langle x_n - x', x' - x\rangle_{\mathcal{X}} + 2\,\nu\langle y_n - y', y' - y\rangle_{\mathcal{Y}} \\ &\quad + 2\,\langle B(y_n - y'), B(y' - y)\rangle_{\mathcal{Z}}. \end{aligned} \tag{10}$$

Taking the average and letting $(\widetilde{x}_{n_k}, \widetilde{y}_{n_k}) \rightharpoonup (x',y')$ as $k \to +\infty$ we obtain

$$\ell(x,y) = \ell(x',y') + Q(x - x', y - y').$$

In a similar fashion we deduce that

$$\ell(x',y') = \ell(x,y) + Q(x - x', y - y')$$

and hence $Q(x - x', y - y') = 0$ which implies $(x,y) = (x',y')$.

3.3 Links with Passty theorem.

Assume that $\mathcal{X} = \mathcal{Y} = \mathcal{Z}$ and that $A = B = \mathcal{I}$, along with $\alpha = \nu = 0$. This induces a situation of strong coupling without cost-to-move. The corresponding algorithm is denoted by $(\mathcal{A}_0)$, see Remark 2.1. Since $R(\mathbf{A}) = \mathcal{X}$, the closedness of $R(\mathbf{A})$ is automatically satisfied. It is immediate that $0 \in \mathbf{T}(x,y)$ if and only if $x = y$ and $Mx + Nx \ni 0$. Therefore we have

$$\boldsymbol{\mathcal{S}} = \mathbf{T}^{-1}0 = \left\{(x,x) \in \mathcal{X}^2, \quad x \in (M+N)^{-1}0\right\}.$$

Assume that the operator $M + N$ is maximal monotone with $(M+N)^{-1}0 \neq \emptyset$. Let (γ_n) be a positive sequence such that $(\gamma_n) \in l^2 \setminus l^1$. By arguing as in the proof of Proposition 3.1, we obtain that

(a) $\lim_{n\to+\infty} \|y_n - \overline{y}\|^2_{\mathcal{X}}$ exists for every $\overline{y} \in (M+N)^{-1}0$.
(b) The sequence $(\|x_{n+1} - y_n\|^2_{\mathcal{X}})$ is summable, hence $\lim_{n\to+\infty} \|x_{n+1} - y_n\|_{\mathcal{X}} = 0$.

Take $\zeta = 0$ and $x = y$ in the proof of Theorem 3.1. Observe that $(0,\eta) \in T(y,y)$ holds if and only if $\eta \in (M+N)y$. Hence formula (9) implies that $\langle \eta, \widetilde{y}_\infty - y\rangle_{\mathcal{Y}} \leq 0$ for every $\eta \in (M+N)y$. We deduce that $\widetilde{y}_\infty \in (M+N)^{-1}0$ by maximality of the

[7] This idea, inspired by the Opial lemma [30] (see Lemma 4.2 below), can also be found in [31].

monotone operator $M+N$. This proves that every weak cluster point of the sequence $(\widetilde{y}_n)$ lies in $(M+N)^{-1}0$. Then we prove that the sequence $(\widetilde{y}_n)$ has at most one weak cluster point. It suffices to adapt the proof of Theorem 3.1 by invoking point (a) above and by using the quantity $Q(v) = \|v\|^2_{\mathcal{X}}$ (instead of $Q(u,v)$). We obtain that the sequence $(\widetilde{y}_n)$ weakly converges toward some $\widetilde{y}_\infty \in (M+N)^{-1}0$. By using point (b) above, we infer that the sequence $(\widetilde{x}_n)$ weakly converges toward $\widetilde{x}_\infty = \widetilde{y}_\infty$. As a conclusion, we recover the following result

Theorem (Passty [31]) *Assume that the operator $M+N$ is maximal monotone with $(M+N)^{-1}0 \neq \emptyset$. Let (γ_n) be a positive sequence such that $(\gamma_n) \in l^2 \setminus l^1$ and let (x_n, y_n) be any sequence generated by algorithm $(\mathcal{A}_0)$. Then there exists $\widetilde{x}_\infty \in (M+N)^{-1}0$ such that both sequences of averages $(\widetilde{x}_n)$ and $(\widetilde{y}_n)$ converge weakly toward $\widetilde{x}_\infty$.*

3.4 Strong monotonicity.

Under strong monotonicity assumptions, we are able to prove the strong convergence of the sequence (x_n, y_n) itself (not only in average). Let us recall that the operator M is said to be strongly monotone with parameter a if, for every x_1, $x_2 \in \mathrm{dom} M$ and every $\xi_1 \in Mx_1$, $\xi_2 \in Mx_2$, we have

$$\langle \xi_2 - \xi_1, x_2 - x_1 \rangle_{\mathcal{X}} \geq a\, \|x_2 - x_1\|^2_{\mathcal{X}}.$$

Assuming in the same way that the operator N is strongly monotone, we obtain that the operators $\mathbf{M}$ and $\mathbf{T} = \mathbf{M} + N_{\boldsymbol{\mathcal{V}}}$ are strongly monotone. Hence if the set $\boldsymbol{\mathcal{S}} = \mathbf{T}^{-1}0$ is nonempty it must be reduced to a single point, say $\boldsymbol{\mathcal{S}} = \{(\overline{x}, \overline{y})\}$.

Proposition 3.2 *Assume that the space $R(\mathbf{A})$ is closed in $\mathcal{Z}$ and that $(\gamma_n) \in l^2 \setminus l^1$. If the operators M and N are strongly monotone and if $\boldsymbol{\mathcal{S}} \neq \emptyset$ then the sequence (x_n, y_n) converges strongly to the unique $(\overline{x}, \overline{y}) \in \boldsymbol{\mathcal{S}}$.*

Proof. Let us suppose that the operators M and N are strongly monotone, respectively with parameters a, $b > 0$. We let the reader check that this assumption leads to a stronger form of inequality (5), which in turn implies

$$h_{n+1}(\overline{x}, \overline{y}) - h_n(\overline{x}, \overline{y}) + 2a\gamma_{n+1}\|x_{n+1} - \overline{x}\|^2_{\mathcal{X}} + 2b\gamma_{n+1}\|y_{n+1} - \overline{y}\|^2_{\mathcal{Y}} \leq \; 2\gamma^2_{n+1}\, \Psi^*(\mathbf{p}).$$

Since $(\gamma_n) \in l^2$ and $\Psi^*(\mathbf{p}) < +\infty$, and recalling that $h_n(\overline{x}, \overline{y}) \geq 0$, the summation of the above inequality implies

$$\sum_{n=1}^{+\infty} \gamma_n \left[\|x_n - \overline{x}\|^2_{\mathcal{X}} + \|y_n - \overline{y}\|^2_{\mathcal{Y}}\right] < +\infty$$

and hence

$$\sum_{n=1}^{+\infty} \gamma_n h_n(\overline{x},\overline{y}) \leq \alpha \sum_{n=1}^{+\infty} \gamma_n \|x_n - \overline{x}\|_{\mathcal{X}}^2 + (\nu + \|B\|^2) \sum_{n=1}^{+\infty} \gamma_n \|y_n - \overline{y}\|_{\mathcal{Y}}^2 < +\infty.$$

Since $(\gamma_n) \notin l^1$ and since $\lim_{n\to+\infty} h_n(\overline{x},\overline{y})$ exists, this limit must be equal to 0 and we deduce that $\lim_{n\to+\infty}(x_n, y_n) = (\overline{x},\overline{y})$.

Remark 3.2 *Observe that the maximality of the operator* $\mathbf{T}$ *does not come into play in the previous proof. Notice also that if the operator* $\mathbf{T}$ *is both maximal and strongly monotone, then condition* $\boldsymbol{\mathcal{S}} \neq \emptyset$ *is automatically satisfied, see for example [20, Cor. 2.4] or [36, Prop. 12. 54].*

4 The subdifferential case: weak convergence results

4.1 Preliminaries

Let $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$, $g : \mathcal{Y} \to \mathbb{R} \cup \{+\infty\}$ be closed convex proper functions. Define the maximal monotone operators M and N respectively by $M = \partial f$ and $N = \partial g$. The operator $\mathbf{M}$ coincides with the subdifferential of the function Φ defined by $\Phi(x, y) = f(x)+g(y)$ for every $(x, y) \in \boldsymbol{\mathcal{X}}$. Observe that the monotone operator $\mathbf{T} = \partial\Phi + N_{\boldsymbol{\mathcal{V}}} = \partial\Phi + \partial\delta_{\boldsymbol{\mathcal{V}}}$ is maximal if, and only if,

$$\partial\Phi + \partial\delta_{\boldsymbol{\mathcal{V}}} = \partial\left(\Phi + \delta_{\boldsymbol{\mathcal{V}}}\right).$$

Maximality is guaranteed if one assumes some qualification condition such as the Moreau-Rockafellar one [29, 34] or the Attouch-Brézis one [8]. In order to cover various applications to PDE's (see paragraph 6.2), we assume the following Attouch-Brézis qualification condition

$$(QC) \qquad \bigcup_{\lambda>0} \lambda(\mathrm{dom} f \times \mathrm{dom} g - \boldsymbol{\mathcal{V}}) \quad \text{is a closed subspace of } \mathcal{X} \times \mathcal{Y}.$$

Under (QC) the following claim shows that the set $\boldsymbol{\mathcal{S}} = \mathbf{T}^{-1}0$ can be interpreted as the set of minima of a suitable function.

Claim 4.1 *We have*

$$\boldsymbol{\mathcal{S}} \subset Argmin_{\boldsymbol{\mathcal{V}}}\Phi = Argmin\{f(x) + g(y) : Ax = By\}.$$

If condition (QC) *is satisfied, the above inclusion holds true as an equality.*

Proof. First recall that the inclusion $\partial\Phi + \partial\delta_{\boldsymbol{\mathcal{V}}} \subset \partial\left(\Phi + \delta_{\boldsymbol{\mathcal{V}}}\right)$ is always satisfied. It ensues immediately that

$$\boldsymbol{\mathcal{S}} = \mathbf{T}^{-1}0 = \left[\partial\Phi + \partial\delta_{\boldsymbol{\mathcal{V}}}\right]^{-1}0 \subset \left[\partial\left(\Phi + \delta_{\boldsymbol{\mathcal{V}}}\right)\right]^{-1}0 = \mathrm{Argmin}_{\boldsymbol{\mathcal{V}}}\Phi.$$

If condition (QC) is satisfied, the set $\bigcup_{\lambda>0}\lambda(\mathrm{dom}\Phi - \mathrm{dom}\delta_{\boldsymbol{\mathcal{V}}})$ is a closed subspace of $\boldsymbol{\mathcal{X}}$. This classically implies that $\partial\Phi + \partial\delta_{\boldsymbol{\mathcal{V}}} = \partial(\Phi + \delta_{\boldsymbol{\mathcal{V}}})$ and the conclusion follows.

Recall that the iterate (x_{n+1}, y_{n+1}) of algorithm $(\mathcal{A})$ is implicitly defined by

$$\begin{cases} 0 \in \gamma_{n+1}\, \partial f(x_{n+1}) + A^*(Ax_{n+1} - By_n) + \alpha(x_{n+1} - x_n) \\ 0 \in \gamma_{n+1}\, \partial g(y_{n+1}) - B^*(Ax_{n+1} - By_{n+1}) + \nu(y_{n+1} - y_n). \end{cases} \tag{11}$$

These are the optimality conditions associated to the following minimization problems

$$(\mathcal{A}) \quad \begin{cases} x_{n+1} = \text{Argmin}\{\gamma_{n+1} f(\zeta) + \frac{1}{2}\|A\zeta - By_n\|^2_{\mathcal{Z}} + \frac{\alpha}{2}\|\zeta - x_n\|^2_{\mathcal{X}}; \quad \zeta \in \mathcal{X}\} \\ y_{n+1} = \text{Argmin}\{\gamma_{n+1} g(\eta) + \frac{1}{2}\|Ax_{n+1} - B\eta\|^2_{\mathcal{Z}} + \frac{\nu}{2}\|\eta - y_n\|^2_{\mathcal{Y}}; \quad \eta \in \mathcal{Y}\}. \end{cases}$$

For each $(x, y) \in \boldsymbol{\mathcal{X}}$, define

$$h_n(x, y) = \alpha\|x_n - x\|^2_{\mathcal{X}} + \nu\|y_n - y\|^2_{\mathcal{Y}} + \|By_n - By\|^2_{\mathcal{Z}} \tag{12}$$

as in section 3. Define also the sequence (φ_n) by

$$\varphi_n = f(x_n) + g(y_n) + \frac{1}{2\gamma_n}\|Ax_n - By_n\|^2_{\mathcal{Z}}. \tag{13}$$

Lemma 4.1 *With the above notations and hypotheses, we have the following*[8]

(i) For every $(x, y) \in Argmin_{\boldsymbol{\mathcal{V}}}\Phi$ and for every $n \geq 0$,

$$h_{n+1}(x, y) - h_n(x, y) + 2\gamma_{n+1}\Big(f(x_{n+1}) + g(y_{n+1}) - \min_{\boldsymbol{\mathcal{V}}} \Phi\Big) + \|Ax_{n+1} - By_{n+1}\|^2_{\mathcal{Z}} + \|Ax_{n+1} - By_n\|^2_{\mathcal{Z}} + \alpha\|x_{n+1} - x_n\|^2_{\mathcal{X}} + \nu\|y_{n+1} - y_n\|^2_{\mathcal{Y}} \leq 0. \tag{14}$$

(ii) For every $n \geq 0$,

$$\varphi_{n+1} - \varphi_n \leq \frac{1}{2}\left(\frac{1}{\gamma_{n+1}} - \frac{1}{\gamma_n}\right)\|Ax_n - By_n\|^2_{\mathcal{Z}}. \tag{15}$$

Proof. In view of the optimality conditions (11), for all $(x, y) \in \mathcal{X} \times \mathcal{Y}$ we can write the subdifferential inequalities

$$\gamma_{n+1}\,(f(x) - f(x_{n+1})) \geq -\langle Ax_{n+1} - By_n, Ax - Ax_{n+1}\rangle_{\mathcal{Z}} - \alpha\langle x_{n+1} - x_n, x - x_{n+1}\rangle_{\mathcal{X}} \tag{16}$$

and

$$\gamma_{n+1}\,(g(y) - g(y_{n+1})) \geq \langle Ax_{n+1} - By_{n+1}, By - By_{n+1}\rangle_{\mathcal{Z}} - \nu\langle y_{n+1} - y_n, y - y_{n+1}\rangle_{\mathcal{Y}}. \tag{17}$$

Using the properties of the inner product the reader can check that

[8] Inequalities (5) and (14) are closely related, even if they rely on different techniques (monotonicity in the first case and subdifferential inequalities in the second one).

$$\begin{aligned}\|By - By_n\|_{\mathcal{Z}}^2 - \|By - By_{n+1}\|_{\mathcal{Z}}^2 = & \|By_{n+1} - Ax_{n+1}\|_{\mathcal{Z}}^2 - \|By - Ax\|_{\mathcal{Z}}^2 \\ & + \|By - Ax - (By_n - Ax_{n+1})\|_{\mathcal{Z}}^2 \\ & + 2\langle By - By_{n+1}, By_{n+1} - Ax_{n+1}\rangle_{\mathcal{Z}} \\ & + 2\langle By_n - Ax_{n+1}, Ax_{n+1} - Ax\rangle_{\mathcal{Z}}.\end{aligned}$$

Combining (16) and (17) we deduce that

$$\begin{aligned}&\|By - By_n\|_{\mathcal{Z}}^2 - \|By - By_{n+1}\|_{\mathcal{Z}}^2 \\ &\quad \geq \|By_{n+1} - Ax_{n+1}\|_{\mathcal{Z}}^2 - \|By - Ax\|_{\mathcal{Z}}^2 + \|By - Ax - (By_n - Ax_{n+1})\|_{\mathcal{Z}}^2 \\ &\qquad + 2\gamma_{n+1}\left[f(x_{n+1}) - f(x) + g(y_{n+1}) - g(y)\right] \\ &\qquad + 2\alpha\langle x_{n+1} - x_n, x_{n+1} - x\rangle_{\mathcal{X}} + 2\nu\langle y_{n+1} - y_n, y_{n+1} - y\rangle_{\mathcal{Y}} \\ &\quad = \|Ax_{n+1} - By_{n+1}\|_{\mathcal{Z}}^2 - \|Ax - By\|_{\mathcal{Z}}^2 + \|By - Ax - (By_n - Ax_{n+1})\|_{\mathcal{Z}}^2 \\ &\qquad + 2\gamma_{n+1}(f(x_{n+1}) + g(y_{n+1}) - f(x) - g(y)) \\ &\qquad + \alpha(\|x_{n+1} - x_n\|_{\mathcal{X}}^2 + \|x_{n+1} - x\|_{\mathcal{X}}^2 - \|x_n - x\|_{\mathcal{X}}^2) \\ &\qquad + \nu(\|y_{n+1} - y_n\|_{\mathcal{Y}}^2 + \|y_{n+1} - y\|_{\mathcal{Y}}^2 - \|y_n - y\|_{\mathcal{Y}}^2).\end{aligned}$$

We infer that for all $(x, y) \in \mathcal{X} \times \mathcal{Y}$,

$$\begin{aligned}h_n(x,y) - h_{n+1}(x,y) \geq & \, 2\gamma_{n+1}(f(x_{n+1}) + g(y_{n+1}) - f(x) - g(y)) - \|Ax - By\|_{\mathcal{Z}}^2 \\ & + \|Ax_{n+1} - By_{n+1}\|_{\mathcal{Z}}^2 + \|By - Ax - (By_n - Ax_{n+1})\|_{\mathcal{Z}}^2 \\ & + \alpha\|x_{n+1} - x_n\|_{\mathcal{X}}^2 + \nu\|y_{n+1} - y_n\|_{\mathcal{Y}}^2. \qquad (18)\end{aligned}$$

Now let $(x, y) \in \mathrm{Argmin}_{\boldsymbol{\mathcal{V}}}\Phi$. Then $Ax = By$ and $f(x) + g(y) = \min_{\boldsymbol{\mathcal{V}}} \Phi$ so that inequality (18) becomes (14). On the other hand, by using inequality (18) with $x = x_n$ and $y = y_n$, we infer that

$$2\gamma_{n+1}(f(x_{n+1}) + g(y_{n+1}) - f(x_n) - g(y_n)) + \|Ax_{n+1} - By_{n+1}\|_{\mathcal{Z}}^2 \leq \|Ax_n - By_n\|_{\mathcal{Z}}^2. \qquad (19)$$

We finally divide by $2\gamma_{n+1}$ and rearrange the terms to obtain (15).

4.2 Weak convergence

Assuming that $\mathrm{Argmin}_{\boldsymbol{\mathcal{V}}}\Phi \neq \emptyset$, let us set

$$\begin{aligned}\omega_n &= \inf_{(x,y)\in\mathcal{X}\times\mathcal{Y}} \left\{\frac{1}{2}\|Ax - By\|_{\mathcal{Z}}^2 + \gamma_n\left(f(x) + g(y) - \min_{\boldsymbol{\mathcal{V}}}\Phi\right)\right\} \\ &= \inf_{\mathbf{x}\in\boldsymbol{\mathcal{X}}} \left\{\Psi(\mathbf{x}) + \gamma_n\left(\Phi(\mathbf{x}) - \min_{\boldsymbol{\mathcal{V}}}\Phi\right)\right\}. \qquad (20)\end{aligned}$$

Denote by (ω_n^-) the negative part of (ω_n). In the sequel, we will assume the key condition

$$(\omega_n^-) \in l^1.$$

This kind of hypothesis was introduced by the second author in [22].

Proposition 4.1 *Assuming that* $\mathrm{Argmin}_{\boldsymbol{\mathcal{V}}}\Phi \neq \emptyset$*, consider the following assertions:*

(i) $(\gamma_n) \in l^2$, *the space* $R(\mathbf{A})$ *is closed in* $\mathcal{Z}$ *and condition* (QC) *is satisfied.*
(ii) $(\gamma_n) \in l^2$ *and there exists* $\overline{\mathbf{x}} \in \boldsymbol{\mathcal{V}}$ *and* $\mathbf{p} \in R(\mathbf{A}^*)$ *such that* $-\mathbf{p} \in \partial\Phi(\overline{\mathbf{x}})$[9].
(iii) $(\omega_n^-) \in l^1$.

Then we have the implications *(i)* $\Longrightarrow$ *(ii)* $\Longrightarrow$ *(iii).*

Proof. (i) $\Longrightarrow$ (ii) Let $\overline{\mathbf{x}} \in \mathrm{Argmin}_{\boldsymbol{\mathcal{V}}}\Phi$. Since condition (QC) is satisfied, we deduce from Claim 4.1 that $\overline{\mathbf{x}} \in \boldsymbol{\mathcal{S}} = \left[\partial\Phi + \boldsymbol{\mathcal{V}}^\perp\right]^{-1} 0$. Hence there exists $\mathbf{p} \in \boldsymbol{\mathcal{V}}^\perp$ such that $-\mathbf{p} \in \partial\Phi(\overline{\mathbf{x}})$. The closedness of $R(\mathbf{A})$ implies the closedness of $R(\mathbf{A}^*)$, hence we have $\boldsymbol{\mathcal{V}}^\perp = \mathrm{Ker}(\mathbf{A})^\perp = R(\mathbf{A}^*)$ and finally $\mathbf{p} \in R(\mathbf{A}^*)$.
(ii) $\Longrightarrow$ (iii) The subdifferential inequality gives for every $\mathbf{x} \in \boldsymbol{\mathcal{X}}$,

$$\Phi(\mathbf{x}) - \Phi(\overline{\mathbf{x}}) \geq \langle -\mathbf{p}, \mathbf{x} - \overline{\mathbf{x}} \rangle_{\boldsymbol{\mathcal{X}}} = \langle -\mathbf{p}, \mathbf{x} \rangle_{\boldsymbol{\mathcal{X}}},$$

where the last equality is a consequence of $\mathbf{p} \in R(\mathbf{A}^*) \subset \boldsymbol{\mathcal{V}}^\perp$ and $\overline{\mathbf{x}} \in \boldsymbol{\mathcal{V}}$. Since $\Phi(\overline{\mathbf{x}}) = \min_{\boldsymbol{\mathcal{V}}} \Phi$, we deduce that

$$\Psi(\mathbf{x}) + \gamma_n \left(\Phi(\mathbf{x}) - \min_{\boldsymbol{\mathcal{V}}} \Phi \right) \geq \Psi(\mathbf{x}) - \gamma_n \langle \mathbf{p}, \mathbf{x} \rangle_{\boldsymbol{\mathcal{X}}}.$$

Taking the infimum over $\mathbf{x} \in \boldsymbol{\mathcal{X}}$, we find

$$\omega_n \geq - \sup_{\mathbf{x} \in \boldsymbol{\mathcal{X}}} \{ \gamma_n \langle \mathbf{p}, \mathbf{x} \rangle_{\boldsymbol{\mathcal{X}}} - \Psi(\mathbf{x}) \} = -\Psi^*(\gamma_n \mathbf{p}) = -\gamma_n^2\, \Psi^*(\mathbf{p}).$$

It ensues that $\omega_n^- \leq \gamma_n^2\, \Psi^*(\mathbf{p})$. Since $\mathbf{p} \in R(\mathbf{A}^*) = \mathrm{dom}\Psi^*$ (see Proposition 2.1), the conclusion follows from the summability of (γ_n^2).

Notice that in infinite dimensional spaces, conditions (ii) or (iii) can be satisfied even if the space $R(\mathbf{A})$ is not closed. An example will be provided in the last section.

Let us now state the main result of the paper.

Theorem 4.1 *Let* $f : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ *and* $g : \mathcal{Y} \to \mathbb{R} \cup \{+\infty\}$ *be closed convex proper functions. Let* $A : \mathcal{X} \to \mathcal{Z}$ *and* $B : \mathcal{Y} \to \mathcal{Z}$ *be linear continuous operators. Assume that the qualification condition* (QC) *holds and that* $\mathrm{Argmin}\{f(x) + g(y) : Ax = By\} \neq \emptyset$*. Let* (γ_n) *be a positive sequence such that* $\left(\frac{1}{\gamma_{n+1}} - \frac{1}{\gamma_n}\right)$ *is majorized by some* $M > 0$*. Finally suppose that condition* $(\omega_n^-) \in l^1$ *holds, where the sequence* (ω_n) *is defined by (20). Then we have*

(i) (x_n, y_n) *converges weakly to a point* $(x_\infty, y_\infty) \in \mathrm{Argmin}\{f(x)+g(y) : Ax = By\}$.
(ii) $\lim\limits_{n \to +\infty} f(x_n) = f(x_\infty)$ *and* $\lim\limits_{n \to +\infty} g(y_n) = g(y_\infty)$.

[9] The corresponding $\mathbf{z} \in \mathcal{Z}$ such that $\mathbf{p} = \mathbf{A}^*\mathbf{z}$ plays the role of a Lagrange multiplier.

Proof. Let us start with several preliminary claims.

Claim 4.2 *For every* $(x,y) \in Argmin_{\boldsymbol{\mathcal{V}}}\Phi$,

$$\lim_{n\to+\infty} \alpha\|x_n - x\|^2_{\mathcal{X}} + \nu\|y_n - y\|^2_{\mathcal{Y}} + \|By_n - By\|^2_{\mathcal{Z}} \quad \textit{exists in } \mathbb{R}.$$

Proof of Claim 4.2. Fix $(x,y) \in \text{Argmin}_{\boldsymbol{\mathcal{V}}}\Phi$ and set $h_n = \alpha\|x_n - x\|^2_{\mathcal{X}} + \nu\|y_n - y\|^2_{\mathcal{Y}} + \|By_n - By\|^2_{\mathcal{Z}}$ as in (12). From inequality (14) we deduce that

$$h_{n+1} - h_n + 2\omega_{n+1} + \|Ax_{n+1} - By_n\|^2_{\mathcal{Z}} + \alpha\|x_{n+1} - x_n\|^2_{\mathcal{X}} + \nu\|y_{n+1} - y_n\|^2_{\mathcal{Y}} \le 0.$$

This implies

$$h_{n+1} - h_n + \|Ax_{n+1} - By_n\|^2_{\mathcal{Z}} + \alpha\|x_{n+1} - x_n\|^2_{\mathcal{X}} + \nu\|y_{n+1} - y_n\|^2_{\mathcal{Y}} \le 2\omega^-_{n+1}. \tag{21}$$

It ensues that $h_{n+1} - h_n \le 2\omega^-_{n+1}$. Since $(\omega^-_n) \in l^1$, owing to Lemma 3.2, we conclude that $\lim_{n\to+\infty} h_n$ exists. □

Claim 4.3 *The sequence* $(\|Ax_n - By_n\|^2_{\mathcal{Z}})$ *is summable, and therefore* $\lim_{n\to+\infty}\|Ax_n - By_n\|_{\mathcal{Z}} = 0$.

Proof of Claim 4.3. Let us sum up inequalities (21) which are obtained for $n = 0$ to $+\infty$. Recalling that $(\omega^-_n) \in l^1$ and that $h_n \ge 0$, we immediately deduce the summability of the sequences $(\|x_{n+1} - x_n\|^2_{\mathcal{X}})$, $(\|y_{n+1} - y_n\|^2_{\mathcal{Y}})$ and $(\|Ax_{n+1} - By_n\|^2_{\mathcal{Z}})$. Since $\|Ax_n - By_n\|^2_{\mathcal{Z}} \le 2\|Ax_{n+1} - By_n\|^2_{\mathcal{Z}} + 2\|Ax_{n+1} - Ax_n\|^2_{\mathcal{Z}}$, the sequence $(\|Ax_n - By_n\|^2_{\mathcal{Z}})$ is also summable. □

Claim 4.4 *Setting* $\varphi_n = f(x_n) + g(y_n) + \frac{1}{2\gamma_n}\|Ax_n - By_n\|^2_{\mathcal{Z}}$ *as in (13), we have* $\lim_{n\to+\infty}\varphi_n = \min_{\boldsymbol{\mathcal{V}}}\Phi$.

Proof of Claim 4.4. Since $\left(\frac{1}{\gamma_{n+1}} - \frac{1}{\gamma_n}\right) \le M$, we derive from inequality (15) that

$$\varphi_{n+1} - \varphi_n \le \frac{M}{2}\|Ax_n - By_n\|^2_{\mathcal{Z}}. \tag{22}$$

From the previous claim the sequence $(\|Ax_n - By_n\|^2_{\mathcal{Z}})$ is summable. By applying Lemma 3.2 we deduce that the sequence (φ_n) converges. Let us now set

$$a_N = 2\sum_{n=0}^{N}\left\{\gamma_n\left(f(x_n) + g(y_n) - \min_{\boldsymbol{\mathcal{V}}}\Phi\right) + \frac{1}{2}\|Ax_n - By_n\|^2_{\mathcal{Z}}\right\}.$$

From inequality (14), the sequence $(h_N + a_N)$ is nonincreasing. Moreover the assumption $(\omega^-_n) \in l^1$ allows us to assert that, for all $n \in \mathbb{N}$,

$$a_N \ge -2\sum_{n=0}^{+\infty}\omega^-_n > -\infty.$$

Thus the sequence $(h_N + a_N)$ is bounded from below, hence convergent. As a consequence,

$$\lim_{N\to+\infty} a_N = \lim_{N\to+\infty} 2\sum_{n=0}^{N} \gamma_n \left(\varphi_n - \min_{\boldsymbol{\mathcal{V}}} \Phi\right) \text{ exists in } \mathbb{R}. \tag{23}$$

Since $\frac{1}{\gamma_{n+1}} - \frac{1}{\gamma_n} \leq M$ for every $n \geq 0$, we deduce that $\gamma_n \geq \frac{1}{Mn+\frac{1}{\gamma_0}}$, hence $(\gamma_n) \notin l^1$. Recalling that $\lim\limits_{n\to+\infty} \varphi_n$ exists in $\mathbb{R}$, we infer from (23) that $\lim\limits_{n\to+\infty} \varphi_n = \min_{\boldsymbol{\mathcal{V}}} \Phi$. □

Claim 4.5 $\lim\limits_{n\to+\infty} \Phi(x_n, y_n) = \min_{\boldsymbol{\mathcal{V}}} \Phi$.

Proof of Claim 4.5. Let $(x, y) \in \operatorname{Argmin}_{\boldsymbol{\mathcal{V}}}\Phi$. Since condition (QC) holds, we deduce from Claim 4.1 that $(x, y) \in \boldsymbol{\mathcal{S}} = \mathbf{T}^{-1}0$. Hence there exists $(p, q) \in \boldsymbol{\mathcal{V}}^\perp$ such that $-(p, q) \in \partial\Phi(x, y)$. The convex subdifferential inequality then gives

$$\begin{aligned}\Phi(x_n, y_n) &\geq \Phi(x, y) + \langle -(p, q), (x_n, y_n) - (x, y)\rangle_{\mathcal{X}\times\mathcal{Y}} \\ &= \min_{\boldsymbol{\mathcal{V}}} \Phi - \langle (p, q), (x_n, y_n)\rangle_{\mathcal{X}\times\mathcal{Y}}.\end{aligned} \tag{24}$$

Let us prove that $\lim\limits_{n\to+\infty} \langle (p, q), (x_n, y_n)\rangle_{\mathcal{X}\times\mathcal{Y}} = 0$. From Claim 4.2 the sequence (x_n, y_n) is bounded, hence it suffices to prove that 0 is the unique limit point of $\left(\langle (p, q), (x_n, y_n)\rangle_{\mathcal{X}\times\mathcal{Y}}\right)$. Let $\left(\langle (p, q), (x_{n_k}, y_{n_k})\rangle_{\mathcal{X}\times\mathcal{Y}}\right)$ be a convergent subsequence. We can extract a subsequence of (x_{n_k}, y_{n_k}), still denoted by (x_{n_k}, y_{n_k}), which weakly converges toward $(\overline{x}, \overline{y})$. The weak lower semicontinuity of the function $(x, y) \mapsto \|Ax - By\|^2_{\mathcal{Z}}$ combined with Claim 4.3 implies that

$$\|A\overline{x} - B\overline{y}\|^2_{\mathcal{Z}} \leq \liminf_{k\to+\infty} \|Ax_{n_k} - By_{n_k}\|^2_{\mathcal{Z}} = \lim_{n\to+\infty} \|Ax_n - By_n\|^2_{\mathcal{Z}} = 0,$$

hence $(\overline{x}, \overline{y}) \in \boldsymbol{\mathcal{V}}$. Recalling that $(p, q) \in \boldsymbol{\mathcal{V}}^\perp$, we infer that

$$\lim_{k\to+\infty} \langle (p, q), (x_{n_k}, y_{n_k})\rangle_{\mathcal{X}\times\mathcal{Y}} = \langle (p, q), (\overline{x}, \overline{y})\rangle_{\mathcal{X}\times\mathcal{Y}} = 0.$$

We immediately deduce that the whole sequence $\left(\langle (p, q), (x_n, y_n)\rangle_{\mathcal{X}\times\mathcal{Y}}\right)$ converges toward 0. Hence from (24) we obtain that $\liminf\limits_{n\to+\infty} \Phi(x_n, y_n) \geq \min_{\boldsymbol{\mathcal{V}}} \Phi$. On the other hand, since $\Phi(x_n, y_n) \leq \varphi_n$, we have in view of Claim 4.4

$$\limsup_{n\to+\infty} \Phi(x_n, y_n) \leq \lim_{n\to+\infty} \varphi_n = \min_{\boldsymbol{\mathcal{V}}} \Phi.$$

We conclude that $\lim\limits_{n\to+\infty} \Phi(x_n, y_n) = \min_{\boldsymbol{\mathcal{V}}} \Phi$.

The proof of (i) relies on the Opial's lemma [30], that we recall for the sake of completeness.

Lemma 4.2 (Opial) *Let $\mathcal{H}$ be a Hilbert space endowed with the norm N. Let (ξ_n) be a sequence of $\mathcal{H}$ such that there exists a nonempty set $\Xi \subset \mathcal{H}$ which verifies*

(a)for all $\xi \in \Xi$, $\lim_{n\to+\infty} N(\xi_n - \xi)$ exists,

(b)if $(\xi_{n_k}) \rightharpoonup \overline{\xi}$ weakly in $\mathcal{H}$ as $k \to +\infty$, we have $\overline{\xi} \in \Xi$.

Then the sequence (ξ_n) weakly converges in $\mathcal{H}$ as $n \to +\infty$ toward a point of Ξ.

Let us define the norm $N(u,v) = \left[\alpha\|u\|^2_{\mathcal{X}} + \nu\|v\|^2_{\mathcal{Y}} + \|Bv\|^2_{\mathcal{Z}}\right]^{1/2}$ on the space $\mathcal{X}\times\mathcal{Y}$. Since the linear operator B is continuous, the norm N is equivalent to the canonical norm on $\mathcal{X}\times\mathcal{Y}$. In view of Claim 4.2, the quantity $N(x_n-x, y_n-y)$ does have a limit as $n \to +\infty$ for every $(x,y) \in \mathrm{Argmin}_{\boldsymbol{\mathcal{V}}}\Phi$, which shows point (a). Let (x_{n_k}, y_{n_k}) be a subsequence of (x_n, y_n) which weakly converges towards $(\overline{x}, \overline{y})$. The weak lower semicontinuity of the function $(x,y) \mapsto \|Ax - By\|^2_{\mathcal{Z}}$ combined with Claim 4.3 implies that

$$\|A\overline{x} - B\overline{y}\|^2_{\mathcal{Z}} \le \liminf_{k\to+\infty} \|Ax_{n_k} - By_{n_k}\|^2_{\mathcal{Z}} = \lim_{n\to+\infty} \|Ax_n - By_n\|^2_{\mathcal{Z}} = 0,$$

hence $(\overline{x}, \overline{y}) \in \boldsymbol{\mathcal{V}}$. In the same way, using Claim 4.5 and the weak lower semicontinuity of Φ, we infer that $(\overline{x}, \overline{y}) \in \mathrm{Argmin}_{\boldsymbol{\mathcal{V}}}\Phi$. This shows point (b) of Opial's lemma and ends the proof of (i).

Let us now prove that $\lim_{n\to+\infty} f(x_n) = f(x_\infty)$. Using the weak lower semicontinuity of f, we have $f(x_\infty) \le \liminf_{n\to+\infty} f(x_n)$. On the other hand, we deduce from Claim 4.5 that

$$\begin{aligned}\limsup_{n\to+\infty} f(x_n) &= \limsup_{n\to+\infty}(f(x_n) + g(y_n) - g(y_n)) \\ &= f(x_\infty) + g(y_\infty) - \liminf_{n\to+\infty} g(y_n).\end{aligned}$$

By the weak lower semicontinuity of g, we have $g(y_\infty) \le \liminf_{n\to+\infty} g(y_n)$. We infer that $\limsup_{n\to+\infty} f(x_n) \le f(x_\infty)$, and finally $\lim_{n\to+\infty} f(x_n) = f(x_\infty)$. In the same way, we have $\lim_{n\to+\infty} g(y_n) = g(y_\infty)$, which ends the proof of (ii).

4.3 Inexact computation of the iterates

Due to the implicit character of the iterations it is important to account for possible computation errors in their implementation. The optimality conditions (11) defining x_{n+1} and y_{n+1} can be relaxed without losing the convergence properties of the algorithm. Suppose the sequences $(\widetilde{x}_n)$ and $(\widetilde{y}_n)$ satisfy

$$\begin{cases} 0 \in \gamma_{n+1}\,\partial_{\varepsilon_{n+1}} f(\widetilde{x}_{n+1}) + A^*(A\widetilde{x}_{n+1} - B\widetilde{y}_n) + \alpha(\widetilde{x}_{n+1} - \widetilde{x}_n) \\ 0 \in \gamma_{n+1}\,\partial_{\varepsilon_{n+1}} g(\widetilde{y}_{n+1}) - B^*(A\widetilde{x}_{n+1} - B\widetilde{y}_{n+1}) + \nu(\widetilde{y}_{n+1} - \widetilde{y}_n), \end{cases} \tag{25}$$

where ∂_ε denotes the approximate ε-subdifferential[10].

The arguments in the proof of Lemma 4.1 give an additional term $4\gamma_{n+1}\varepsilon_{n+1}$ on the right-hand side of inequality (14) and an additional $2\varepsilon_{n+1}$ on the right-hand side of (15). As a consequence, Claims 4.2 and 4.3 remain true if $\sum_{n=1}^{\infty}\gamma_n\varepsilon_n < +\infty$. Claims 4.4 and 4.5 also hold if $\sum_{n=1}^{\infty}\varepsilon_n < +\infty$.

Corollary 4.1 *Theorem 4.1 holds under the same hypotheses if the iterates satisfy* (25) *instead of* (11) *provided* $(\varepsilon_n) \in l^1$.

There is a rich literature regarding the treatment of errors in proximal-type algorithms. The interested reader may consult [35], [15] or [37]. See also [6] for an alternative approach to computational errors.

5 Further convergence results for strongly coupled problems

In this section, we assume that $\mathcal{X} = \mathcal{Y} = \mathcal{Z}$ and that $A = B = \mathcal{I}$, along with $\alpha = \nu = 0$. Given closed convex functions $f,\ g : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$, consider the following particular case[11] of algorithm $(\mathcal{A})$

$$(\mathcal{A}_0) \qquad \begin{cases} x_{n+1} = \operatorname{Argmin}\Big\{\gamma_{n+1}\, f(\zeta) + \dfrac{1}{2}\|\zeta - y_n\|_{\mathcal{X}}^2;\ \zeta \in \mathcal{X}\Big\} \\ y_{n+1} = \operatorname{Argmin}\Big\{\gamma_{n+1}\, g(\eta) + \dfrac{1}{2}\|x_{n+1} - \eta\|_{\mathcal{X}}^2;\ \eta \in \mathcal{X}\Big\}. \end{cases}$$

Using the same notations as in the previous sections, we have

$$\boldsymbol{\mathcal{V}} = \{(x,x); x \in \mathcal{X}\} \quad \text{and} \quad \operatorname{Argmin}_{\boldsymbol{\mathcal{V}}}\Phi = \{(x,x); x \in \operatorname{Argmin}(f+g)\}.$$

Let us first start with an example.

Example 5.1 Take $\mathcal{X} = \mathbb{R}$ and define the functions $f,\ g : \mathbb{R} \to \mathbb{R}$ respectively by $f(x) = \frac{1}{2}(x-1)^2$ and $g(y) = \frac{1}{2}(y+1)^2$. We then have $\operatorname{Argmin}(f+g) = \{0\}$. By writing down the optimality conditions for algorithm $(\mathcal{A}_0)$, we immediately obtain the following recurrence formulae (see also Remark 2.1)

$$\begin{cases} \gamma_{n+1}(x_{n+1} - 1) + x_{n+1} - y_n = 0 \\ \gamma_{n+1}(y_{n+1} + 1) + y_{n+1} - x_{n+1} = 0. \end{cases}$$

We infer that

$$y_{n+1} = \frac{1}{(1+\gamma_{n+1})^2}y_n - \frac{\gamma_{n+1}^2}{(1+\gamma_{n+1})^2}.$$

[10] In Hilbert space $\mathcal{H}$, given $\varepsilon > 0$ and $F : \mathcal{H} \to \mathbb{R} \cup \{+\infty\}$, the approximate ε-subdifferential of F at ξ is defined by $\partial_\varepsilon F(\xi) = \{\xi^* \in \mathcal{H}\ :\ F(\zeta) \geq F(\xi) + \langle \xi^*, \zeta - \xi\rangle - \varepsilon \quad \forall \zeta \in \mathcal{H}\}$.

[11] See Remark 2.1, where algorithm $(\mathcal{A}_0)$ has been introduced in the framework of maximal monotone operators.

Let us set $a_n = \frac{1}{(1+\gamma_{n+1})^2}$ and $b_n = \frac{\gamma_{n+1}^2}{(1+\gamma_{n+1})^2}$. We deduce from the above equality that $|y_{n+1}| \leq a_n|y_n| + b_n$. To prove the convergence of the sequence (y_n), we use the following lemma borrowed from [32, Lemma 3, p. 45].

Lemma 5.1 *Let* (a_n) *and* (b_n) *be real sequences such that* $0 \leq a_n < 1$ *and* $b_n \geq 0$ *for every* $n \in \mathbb{N}$. *Assume moreover that* $(1-a_n) \notin l^1$ *and that* $\lim_{n\to+\infty} \frac{b_n}{1-a_n} = 0$. *Let* (u_n) *be a real sequence such that* $u_{n+1} \leq a_n u_n + b_n$ *for every* $n \in \mathbb{N}$. *Then we have* $\limsup_{n\to+\infty} u_n \leq 0$.

It is easy to check that, if the sequence (γ_n) is not summable, then the sequence $(1-a_n)$ is not summable. Moreover we have $\lim_{n\to+\infty} \frac{b_n}{1-a_n} = \lim_{n\to+\infty} \frac{\gamma_{n+1}}{2+\gamma_{n+1}} = 0$. Thus the previous lemma implies that $\limsup_{n\to+\infty} |y_n| \leq 0$, hence $\lim_{n\to+\infty} y_n = 0$. Finally we have proved that if $(\gamma_n) \notin l^1$ then $\lim_{n\to+\infty}(x_n, y_n) = (0,0)$.

It is worthwhile noticing that the assumption $(\gamma_n) \in l^2$ does not come into play in the above example. This is in fact a consequence of a general result that will be brought to light by Theorem 5.1 (i), see also Remark 5.1. Before stating Theorem 5.1, we need the following preliminary result.

Proposition 5.1 *Let* $f, g : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ *be closed convex functions which are bounded from below and such that* $dom f \cap dom g \neq \emptyset$. *Let* (γ_n) *be a positive nonincreasing sequence such that* $\lim_{n\to\infty} \gamma_n = 0$. *Then any sequence* (x_n, y_n) *generated by* $(\mathcal{A}_0)$ *satisfies* $\lim_{n\to+\infty} \|x_n - y_n\|_{\mathcal{X}} = 0$.

Proof. Let us define the sequence (ψ_n) by

$$\psi_n = \gamma_n\big(f(x_n) + g(y_n)\big) + \frac{1}{2}\|x_n - y_n\|_{\mathcal{X}}^2. \tag{26}$$

We have

$$\psi_n \geq \gamma_n \inf \Phi + \frac{1}{2}\|x_n - y_n\|_{\mathcal{X}}^2, \tag{27}$$

hence the sequence $(\psi_n - \gamma_n \inf \Phi)$ is nonnegative. By using inequality (19) with $A = B = \mathcal{I}$, we deduce that, for every $n \in \mathbb{N}$,

$$\psi_{n+1} - \psi_n \leq (\gamma_{n+1} - \gamma_n) \inf \Phi.$$

This shows that the sequence $(\psi_n - \gamma_n \inf \Phi)$ is nonincreasing, hence convergent. Since $\lim_{n\to+\infty} \gamma_n = 0$, the sequence (ψ_n) also converges. Let us apply inequality (18) with $A = B = \mathcal{I}$, $\alpha = \nu = 0$ and $x = y$; we find for all $x \in \mathrm{dom} f \cap \mathrm{dom} g \neq \emptyset$,

$$2\psi_{n+1} - 2\gamma_{n+1}(f(x) + g(x)) \leq \|y_n - x\|_{\mathcal{X}}^2 - \|y_{n+1} - x\|_{\mathcal{X}}^2.$$

By summing the above inequalities for $n = 0, ..., N$, we obtain

$$2\sum_{n=0}^{N} [\psi_{n+1} - \gamma_{n+1}(f(x) + g(x))] \leq \|y_0 - x\|_{\mathcal{X}}^2.$$

Since this is true for every $N \in \mathbb{N}$, we derive that

$$\liminf_{n\to+\infty} \left[\psi_{n+1} - \gamma_{n+1}(f(x) + g(x))\right] \leq 0.$$

But both terms are convergent, so we have $\lim_{n\to+\infty} \psi_n \leq 0$. From (27) we immediately deduce that $\lim_{n\to+\infty} \|x_n - y_n\|^2_{\mathcal{X}} = 0$.

The approach that we now develop relies on topological ingredients that can already be found in [5, 10, 18, 22]. The result below shows that if $(\gamma_n) \notin l^1$ and $\lim_{n\to+\infty} \gamma_n = 0$, the iterates x_n, y_n of algorithm $(\mathcal{A}_0)$ approach the set $\mathrm{Argmin}(f+g)$ as $n \to +\infty$. Weak convergence is obtained under the extra assumption $(\gamma_n) \in l^2$. In the next statement, we denote by $d\big(\cdot, \mathrm{Argmin}(f+g)\big)$ the distance function to the set $\mathrm{Argmin}(f+g)$.

Theorem 5.1 *Let $f, g : \mathcal{X} \to \mathbb{R} \cup \{+\infty\}$ be closed convex functions which are bounded from below. Assume that either f or g is inf-compact*[12]. *Let (γ_n) be a positive nonincreasing sequence such that $(\gamma_n) \notin l^1$ and $\lim_{n\to+\infty} \gamma_n = 0$. Finally, let (x_n, y_n) be a sequence generated by $(\mathcal{A}_0)$. Then*

(i) $\lim_{n\to+\infty} d_{\mathcal{X}}\big(x_n, Argmin(f+g)\big) = \lim_{n\to+\infty} d_{\mathcal{X}}\big(y_n, Argmin(f+g)\big) = 0.$

(ii) If $(\gamma_n) \in l^2$, and if condition (QC) is satisfied[13], *then the sequence (x_n, y_n) converges weakly to a point $(\overline{x}, \overline{x})$ with $\overline{x} \in Argmin(f+g)$.*

(iii) If moreover the sequence $\left(\frac{1}{\gamma_{n+1}} - \frac{1}{\gamma_n}\right)$ is majorized by some $M > 0$, then the sequence (x_n, y_n) converges strongly in $\mathcal{X}$.

Proof. First, if $\mathrm{dom} f \cap \mathrm{dom} g = \emptyset$ then $\mathrm{Argmin}(f+g) = \mathcal{X}$, condition (QC) cannot hold and the result is trivial. Hence we assume $\mathrm{dom} f \cap \mathrm{dom} g \neq \emptyset$. We can also assume that the function f is inf-compact. Since the function g is bounded from below, we derive that the function $f+g$ is inf-compact, hence $\mathrm{Argmin}(f+g) \neq \emptyset$.
(i) In view of Proposition 5.1, it suffices to prove that $\lim_{n\to+\infty} d_{\mathcal{X}}\big(y_n, \mathrm{Argmin}(f+g)\big) = 0$. Set $A = B = \mathcal{I}$ and $\alpha = \nu = 0$ in inequality (14) to deduce that for every $y \in \mathrm{Argmin}(f+g)$ and every $n \in \mathbb{N}$ we have

$$\|y_{n+1} - y\|^2_{\mathcal{X}} - \|y_n - y\|^2_{\mathcal{X}} + 2\gamma_{n+1}\left(\Phi(x_{n+1}, y_{n+1}) - \min_{\mathcal{V}} \Phi\right) + \|x_{n+1} - y_{n+1}\|^2_{\mathcal{X}} \leq 0. \tag{28}$$

[12] Recall that a function is said to be inf-compact if its sublevel sets are relatively compact.

[13] In our present setting, it is easy to check that condition (QC) is satisfied if and only if

$$\bigcup_{\lambda>0} \lambda(\mathrm{dom} f - \mathrm{dom} g) \text{ is a closed subspace of } \mathcal{X}.$$

This is precisely the Attouch-Brézis condition, which ensures that $\partial f + \partial g = \partial(f+g)$ and hence $(\partial f + \partial g)^{-1}0 = \mathrm{Argmin}(f+g)$.

Let P denote the projection operator onto the closed convex set $\mathrm{Argmin}(f+g)$ and take $y = P(y_n)$. Setting $u_n = d^2_{\mathcal{X}}\big(y_n, \mathrm{Argmin}(f+g)\big)$, we derive from (28) that

$$u_{n+1} - u_n + 2\gamma_{n+1}\Big(\Phi(x_{n+1}, y_{n+1}) - \min_{\boldsymbol{\mathcal{V}}} \Phi\Big) \leq 0. \tag{29}$$

We now follow the same arguments as those used by the second author in [22, Theorem 3.1]. We distinguish two cases:

(a)There exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, $\Phi(x_n, y_n) > \min_{\boldsymbol{\mathcal{V}}} \Phi$.
(b)For all $n_0 \in \mathbb{N}$, there exists $n \geq n_0$ such that $\Phi(x_n, y_n) \leq \min_{\boldsymbol{\mathcal{V}}} \Phi$.

Case (a). Assume there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, $\Phi(x_n, y_n) > \min_{\boldsymbol{\mathcal{V}}} \Phi$. From inequality (29), we deduce that the sequence $(u_n)_{n \geq n_0}$ is nonincreasing and convergent. We must prove that $\lim_{n\to+\infty} u_n = 0$. Using again inequality (29), we can assert that the sequence $(\gamma_n(\Phi(x_n, y_n) - \min_{\boldsymbol{\mathcal{V}}} \Phi))$ is summable. Moreover, since $(\gamma_n) \notin l^1$, we have $\liminf_{n\to+\infty} \Phi(x_n, y_n) = \min_{\boldsymbol{\mathcal{V}}} \Phi$. Consider a subsequence of (x_n, y_n), still denoted by (x_n, y_n), such that $\lim\limits_{n\to+\infty} \Phi(x_n, y_n) = \min_{\boldsymbol{\mathcal{V}}} \Phi$. Since the function g is bounded from below, the sequence $(f(x_n))$ is majorized. Using the inf-compactness of the map f, we obtain that the sequence (x_n) is relatively compact in $\mathcal{X}$. Thus there exist a subsequence (x_{n_k}) along with $\overline{x} \in \mathcal{X}$ such that $\lim\limits_{k\to+\infty} x_{n_k} = \overline{x}$ strongly in $\mathcal{X}$. In view of Proposition 5.1 we also have $\lim\limits_{k\to+\infty} y_{n_k} = \overline{x}$ strongly in $\mathcal{X}$. The closedness of the function Φ allows to assert that $\Phi(\overline{x}, \overline{x}) \leq \liminf_{k\to+\infty} \Phi(x_{n_k}, y_{n_k}) = \min_{\boldsymbol{\mathcal{V}}} \Phi$. Hence $(\overline{x}, \overline{x}) \in \mathrm{Argmin}_{\boldsymbol{\mathcal{V}}}\Phi$, *i.e.* $\overline{x} \in \mathrm{Argmin}(f+g)$. Thus

$$\lim_{k\to+\infty} u_{n_k} = \lim_{n\to+\infty} d^2_{\mathcal{X}}(y_{n_k}, \mathrm{Argmin}(f+g)) = d^2_{\mathcal{X}}(\overline{x}, \mathrm{Argmin}(f+g)) = 0.$$

Recalling that the sequence (u_n) is convergent, we conclude that $\lim\limits_{n\to+\infty} u_n = 0$.

Case (b). We assume that for all $n_0 \in \mathbb{N}$ there exists $n \geq n_0$ such that $\Phi(x_n, y_n) \leq \min_{\boldsymbol{\mathcal{V}}} \Phi$. Let us define

$$\tau_N = \max\{n \in \mathbb{N}, n \leq N \text{ and } \Phi(x_n, y_n) \leq \min_{\boldsymbol{\mathcal{V}}} \Phi\}.$$

The integer τ_N is well-defined for N large enough and $\lim_{N\to+\infty} \tau_N = \infty$. If $\tau_N < N$ inequality (29) implies $u_{n+1} \leq u_n$ whenever $\tau_N \leq n \leq N-1$. In particular,

$$u_N \leq u_{\tau_N}. \tag{30}$$

Notice that if $\tau_N = N$ this inequality is still true. Therefore it suffices to prove that $\lim\limits_{n\to+\infty} u_{\tau_n} = 0$. First observe that $\Phi(x_{\tau_n}, y_{\tau_n}) \leq \min_{\boldsymbol{\mathcal{V}}} \Phi$ for all sufficiently large n by definition. We deduce, as before, that the sequence (x_{τ_n}) is relatively compact, hence bounded in $\mathcal{X}$. In view of Proposition 5.1, the sequence (y_{τ_n}) is also bounded in $\mathcal{X}$, whence the boundedness of the real sequence (u_{τ_n}). The proof will be complete if we verify that every convergent subsequence of (u_{τ_n}) must vanish. Indeed, assume that $\lim_{k\to+\infty} u_{\tau_{n_k}}$ exists. We may assume, upon passing to a subsequence if necessary, that $\lim_{k\to+\infty} x_{\tau_{n_k}} = \lim_{k\to+\infty} y_{\tau_{n_k}} = \overline{x}$ for some $\overline{x} \in \mathcal{X}$. The closedness of Φ then gives

$$\Phi(\overline{x}, \overline{x}) \le \liminf_{k\to\infty} \Phi(x_{\tau_{n_k}}, y_{\tau_{n_k}}) \le \min_{\mathcal{V}} \Phi,$$

which implies $(\overline{x}, \overline{x}) \in \operatorname{Argmin}_{\mathcal{V}} \Phi$. As before, this implies $\lim_{k\to+\infty} u_{\tau_{n_k}} = 0$ and we deduce that the whole sequence (u_{τ_n}) converges toward 0. Then inequality (30) shows that $\lim_{n\to+\infty} u_n = 0$.

(ii) Let us assume that $(\gamma_n) \in l^2$ and that condition (QC) is satisfied. Observe that $R(A) + R(B) = R(\mathcal{I}) = \mathcal{X}$, so the closedness of the space $R(A) + R(B)$ is fulfilled. By Proposition 4.1, the sequence (ω_n) defined by

$$\omega_n = \inf_{(x,y)\in\mathcal{X}^2} \left\{ \frac{1}{2}\|x-y\|_{\mathcal{X}}^2 + \gamma_n \left(f(x) + g(y) - \min_{\mathcal{V}} \Phi \right) \right\}$$

satisfies $(\omega_n^-) \in l^1$. Let $y \in \operatorname{Argmin}(f+g)$. From inequality (28), we obtain

$$\|y_{n+1} - y\|_{\mathcal{X}}^2 - \|y_n - y\|_{\mathcal{X}}^2 \le 2\omega_{n+1}^-.$$

Since $(\omega_n^-) \in l^1$ this implies in view of Lemma 3.2 that

$$\forall y \in \operatorname{Argmin}(f+g), \quad \lim_{n\to+\infty} \|y_n - y\|_{\mathcal{X}}^2 \text{ exists.} \tag{31}$$

On the other hand, recalling that $\lim_{n\to+\infty} d_{\mathcal{X}}(y_n, \operatorname{Argmin}(f+g)) = 0$, every weak cluster point of the sequence (y_n) lies in $\operatorname{Argmin}(f+g)$. We infer from Lemma 4.2 that the sequence (y_n) weakly converges toward some point in $\operatorname{Argmin}(f+g)$. Finally Proposition 5.1 shows that the sequences (x_n) and (y_n) tend weakly toward the same limit.

(iii) Let us first prove that the sequence (φ_n) defined by formula (13) is bounded. By applying inequality (18) with $A = B = \mathcal{I}$, $\alpha = \nu = 0$ and $(x, y) = (x_n, y_n)$, we easily find

$$\varphi_{n+1} - \varphi_n + \frac{1}{2\,\gamma_{n+1}} \left(\|x_{n+1} - x_n\|_{\mathcal{X}}^2 + \|y_{n+1} - y_n\|_{\mathcal{X}}^2 \right) \le \frac{1}{2} \left(\frac{1}{\gamma_{n+1}} - \frac{1}{\gamma_n} \right) \|x_n - y_n\|_{\mathcal{X}}^2. \tag{32}$$

Observe that this inequality is slightly more precise than (15), where two terms were omitted. Since $\frac{1}{\gamma_{n+1}} - \frac{1}{\gamma_n} \le M$ by assumption and since $\|x_n - y_n\|_{\mathcal{X}}^2 \le 2\,\|x_{n+1} - y_n\|_{\mathcal{X}}^2 + 2\,\|x_{n+1} - x_n\|_{\mathcal{X}}^2$, inequality (32) implies

$$\varphi_{n+1} - \varphi_n + \frac{1}{2\,\gamma_{n+1}} \|x_{n+1} - x_n\|_{\mathcal{X}}^2 \le M \left(\|x_{n+1} - y_n\|_{\mathcal{X}}^2 + \|x_{n+1} - x_n\|_{\mathcal{X}}^2 \right).$$

From the fact that $\lim_{n\to+\infty} \gamma_n = 0$, we immediately derive that for n large enough

$$\varphi_{n+1} - \varphi_n \leq M\, \|x_{n+1} - y_n\|^2_{\mathcal{X}}. \tag{33}$$

Recall that the sequence (ω_n^-) is summable, see the proof of (ii). The summability of $(\|x_{n+1} - y_n\|^2_{\mathcal{X}})$ is then an immediate consequence of inequality (21), with $A = B = \mathcal{I}$ and $\alpha = \nu = 0$. In view of (33), we infer from Lemma 3.2 that the sequence (φ_n) is convergent, hence bounded. Since the function g is bounded from below, the sequence $(f(x_n))$ is majorized. The inf-compactness of f allows to deduce that the sequence (x_n) is relatively compact in $\mathcal{X}$. Hence there exists $\overline{x} \in \mathcal{X}$ along with a subsequence (x_{n_k}) such that $\lim_{k\to+\infty} x_{n_k} = \overline{x}$ strongly in $\mathcal{X}$. From Proposition 5.1, we also have $\lim_{k\to+\infty} y_{n_k} = \overline{x}$ strongly in $\mathcal{X}$. In view of (i), it is clear that $\overline{x} \in \mathrm{Argmin}(f + g)$. Taking $y = \overline{x}$ in assertion (31), we deduce that $\lim_{n\to+\infty} \|y_n - \overline{x}\|_{\mathcal{X}} = 0$. Owing to Proposition 5.1, we conclude that $\lim_{n\to+\infty} x_n = \lim_{n\to+\infty} y_n = \overline{x}$ strongly in $\mathcal{X}$.

Remark 5.1 *Observe that if Argmin$(f + g) = \{\overline{\xi}\}$, Theorem 5.1 (i) shows that any sequence generated by $(\mathcal{A}_0)$ converges strongly to $(\overline{\xi}, \overline{\xi})$, even if $(\gamma_n) \notin l^2$.*

Remark 5.2 *No qualification condition is required in the proof of Theorem 5.1 (i), which is a distinctive mark with respect to the proof of Theorem 4.1 (see specially Claim 4.5).*

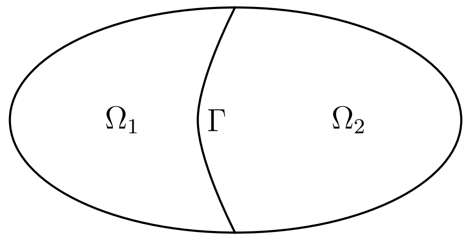
Remark 5.3 *Recall from Remark 2.1 that the iterates of algorithm $(\mathcal{A}_0)$ satisfy the following equalities*

$$\begin{aligned} x_{n+1} &= (I + \gamma_{n+1}\, \partial f)^{-1}(I + \gamma_n\, \partial g)^{-1}\, x_n, \\ y_{n+1} &= (I + \gamma_{n+1}\, \partial g)^{-1}(I + \gamma_{n+1}\, \partial f)^{-1}\, y_n. \end{aligned}$$

This corresponds to a double resolvent scheme studied by Passty in [31]. In this reference, weak ergodic convergence of such sequences is established for general maximal monotone operators such that the sum is itself maximal, provided that $(\gamma_n) \in l^2 \setminus l^1$. Under some inf-compactness assumption, Theorem 5.1 (ii) (resp. (iii)) shows that weak ergodic convergence is replaced by weak (resp. strong) convergence in the subdifferential framework. Hence our result is an improvement of Passty theorem when applied to subdifferential operators.

6 Application to domain decomposition for PDE's

Let us consider a bounded domain $\Omega \subset \mathbb{R}^N$ with $\mathcal{C}^2$ boundary. Assume that the set Ω is decomposed in two nonoverlapping Lipschitz subdomains Ω_1 and Ω_2 with a common interface Γ. This situation is illustrated in the next figure.



6.1 Neumann problem

Given a function $h \in L^2(\Omega)$, let us consider the following Neumann boundary value problem on Ω

$$\begin{cases} -\Delta w = h \text{ on } \Omega \\ \frac{\partial w}{\partial n} = 0 \text{ on } \partial\Omega, \end{cases}$$

where $\frac{\partial w}{\partial n} = \nabla w.\mathbf{n}$ and $\mathbf{n}$ is the unit outward normal to $\partial\Omega$. We assume that $\int_\Omega h = 0$, which is a necessary and sufficient condition for the existence of a solution. The weak solutions of the above Neumann problem satisfy the following minimization problem

$$\min \left\{ \frac{1}{2} \int_\Omega |\nabla w|^2 - \int_\Omega hw; \quad w \in H^1(\Omega) \right\}, \tag{34}$$

see for example [9, 21, 27, 33]. Moreover, denoting by $\widehat{w}$ a particular solution, the solution set of (34) is of the form $\{\widehat{w} + C,\ C \in \mathbb{R}\}$. Assuming that Ω is of class $\mathcal{C}^2$, we know from the regularity theory of weak solutions that $\widehat{w} \in H^2(\Omega)$, see for instance [3, 4, 26]. Notice that, if $w \in H^1(\Omega)$ then the restrictions $u = w_{|\Omega_1}$ and $v = w_{|\Omega_2}$ belong respectively to $H^1(\Omega_1)$ and $H^1(\Omega_2)$ and moreover $u_{|\Gamma} = v_{|\Gamma}$. Conversely, if $u \in H^1(\Omega_1)$, $v \in H^1(\Omega_2)$ and if $u_{|\Gamma} = v_{|\Gamma}$, then the function w defined by $w = \begin{cases} u \text{ on } \Omega_1 \\ v \text{ on } \Omega_2 \end{cases}$ belongs to $H^1(\Omega)$. As a consequence, problem (34) can be reformulated as

$$(\mathcal{P}) \qquad \min \left\{ f(u) + g(v); \quad (u,v) \in H^1(\Omega_1) \times H^1(\Omega_2) \text{ and } u_{|\Gamma} = v_{|\Gamma} \right\},$$

where

$$f(u) = \frac{1}{2} \int_{\Omega_1} |\nabla u|^2 - \int_{\Omega_1} hu \quad \text{and} \quad g(v) = \frac{1}{2} \int_{\Omega_2} |\nabla v|^2 - \int_{\Omega_2} hv. \tag{35}$$

Let us show how the algorithm $(\mathcal{A})$ can be applied so as to solve problem $(\mathcal{P})$. The set $\mathcal{X} = H^1(\Omega_1)$ is equipped with the scalar product $\langle u_1, u_2 \rangle_{\mathcal{X}} = \int_{\Omega_1} (\nabla u_1.\nabla u_2 +$

$u_1 u_2$) and the corresponding norm. The same holds for $\mathcal{Y} = H^1(\Omega_2)$ by replacing Ω_1 with Ω_2. The set $\mathcal{Z} = L^2(\Gamma)$ is equipped with the scalar product $\langle z_1, z_2 \rangle_{\mathcal{Z}} = \int_\Gamma z_1 z_2$ and the corresponding norm. The operators $A : \mathcal{X} \to \mathcal{Z}$ and $B : \mathcal{Y} \to \mathcal{Z}$ are respectively the trace operators on Γ, which are well-defined by the Lipschitz character of the boundaries of Ω_1 and Ω_2 (see [17, Theorem II.46] or [28, Theorem 2]). Algorithm $(\mathcal{A})$ runs as follows

$$\begin{cases} u_{n+1} = \text{Argmin}\{\gamma_{n+1} f(u) + \dfrac{1}{2}\|Au - Bv_n\|^2_{\mathcal{Z}} + \dfrac{\alpha}{2}\|u - u_n\|^2_{\mathcal{X}}; \quad u \in \mathcal{X}\} \\ v_{n+1} = \text{Argmin}\{\gamma_{n+1} g(v) + \dfrac{1}{2}\|Au_{n+1} - Bv\|^2_{\mathcal{Z}} + \dfrac{\nu}{2}\|v - v_n\|^2_{\mathcal{Y}}; \quad v \in \mathcal{Y}\}, \end{cases}$$

where α and ν are fixed positive parameters. An elementary directional derivative computation shows that the weak variational formulation of algorithm $(\mathcal{A})$ is given by

$$\begin{aligned} &\gamma_{n+1} \int_{\Omega_1} \nabla u_{n+1}.\nabla u + \alpha \int_{\Omega_1} (\nabla u_{n+1} - \nabla u_n).\nabla u \\ &\quad + \alpha \int_{\Omega_1} (u_{n+1} - u_n) u + \int_\Gamma (Au_{n+1} - Bv_n) Au = \gamma_{n+1} \int_{\Omega_1} hu \end{aligned}$$

and

$$\begin{aligned} &\gamma_{n+1} \int_{\Omega_2} \nabla v_{n+1}.\nabla v + \nu \int_{\Omega_2} (\nabla v_{n+1} - \nabla v_n).\nabla v \\ &\quad + \nu \int_{\Omega_2} (v_{n+1} - v_n) v + \int_\Gamma (Bv_{n+1} - Au_{n+1}) Bv = \gamma_{n+1} \int_{\Omega_2} hv \end{aligned}$$

for all $u \in \mathcal{X}$ and $v \in \mathcal{Y}$. These are the variational weak formulations of the following mixed Dirichlet-Neumann boundary value problems respectively on Ω_1

$$\begin{cases} -(\gamma_{n+1} + \alpha)\Delta u_{n+1} + \alpha u_{n+1} = \gamma_{n+1} h - \alpha \Delta u_n + \alpha u_n & \text{on } \Omega_1 \\ (\gamma_{n+1} + \alpha)\frac{\partial u_{n+1}}{\partial n} = \alpha \frac{\partial u_n}{\partial n} & \text{on } \partial\Omega_1 \cap \partial\Omega \\ (\gamma_{n+1} + \alpha)\frac{\partial u_{n+1}}{\partial n} + u_{n+1} = \alpha \frac{\partial u_n}{\partial n} + v_n & \text{on } \Gamma, \end{cases}$$

and Ω_2

$$\begin{cases} -(\gamma_{n+1} + \nu)\Delta v_{n+1} + \nu v_{n+1} = \gamma_{n+1} h - \nu \Delta v_n + \nu v_n & \text{on } \Omega_2 \\ (\gamma_{n+1} + \nu)\frac{\partial v_{n+1}}{\partial n} = \nu \frac{\partial v_n}{\partial n} & \text{on } \partial\Omega_2 \cap \partial\Omega \\ (\gamma_{n+1} + \nu)\frac{\partial v_{n+1}}{\partial n} + v_{n+1} = \nu \frac{\partial v_n}{\partial n} + u_{n+1} & \text{on } \Gamma. \end{cases}$$

Let us now check the validity of the assumptions of Theorem 4.1. The qualification condition (QC) is automatically satisfied since $\text{dom} f = \mathcal{X}$ and $\text{dom} g = \mathcal{Y}$.

In view of Proposition 4.1, assumption $(\omega_n^-) \in l^1$ is verified[14] if $(\gamma_n) \in l^2$ and if there exist $(\widehat{u}, \widehat{v}) \in \mathcal{X} \times \mathcal{Y}$ such that $\widehat{u}_{|\Gamma} = \widehat{v}_{|\Gamma}$ along with $z \in \mathcal{Z}$ satisfying

$$-A^* z \in \partial f(\widehat{u}) \quad \text{and} \quad B^* z \in \partial g(\widehat{v}). \tag{36}$$

Take $\widehat{u} = \widehat{w}_{|\Omega_1}$ and $\widehat{v} = \widehat{w}_{|\Omega_2}$ the restrictions of $\widehat{w}$ respectively to Ω_1 and Ω_2. Let us multiply the equality $-\Delta \widehat{u} = h$ by $u \in H^1(\Omega_1)$ and integrate on Ω_1. Using Green's formula and the fact that $\frac{\partial \widehat{u}}{\partial n} = 0$ on $\partial\Omega \cap \partial\Omega_1$, we obtain

$$\forall u \in H^1(\Omega_1), \quad \int_{\Omega_1} \nabla \widehat{u}. \nabla u - \int_\Gamma \frac{\partial \widehat{u}}{\partial n} u = \int_{\Omega_1} hu.$$

Hence we deduce that for every $u \in H^1(\Omega_1)$,

$$f(u) = \frac{1}{2} \int_{\Omega_1} |\nabla u|^2 - \int_{\Omega_1} \nabla \widehat{u}. \nabla u + \int_\Gamma \frac{\partial \widehat{u}}{\partial n}\, u$$

and therefore

$$\begin{aligned} f(u) - f(\widehat{u}) &= \frac{1}{2} \int_{\Omega_1} |\nabla u - \nabla \widehat{u}|^2 + \int_\Gamma \frac{\partial \widehat{u}}{\partial n}(u - \widehat{u}) \\ &\geq \int_\Gamma \frac{\partial \widehat{u}}{\partial n}(u - \widehat{u}) = \left\langle A^* \frac{\partial \widehat{u}}{\partial n}, u - \widehat{u} \right\rangle_{\mathcal{X}}. \end{aligned}$$

This shows that $A^* \frac{\partial \widehat{u}}{\partial n} \in \partial f(\widehat{u})$ and we find in the same way $B^* \frac{\partial \widehat{v}}{\partial n} \in \partial g(\widehat{v})$. Since $\frac{\partial \widehat{u}}{\partial n}_{|\Gamma} = -\frac{\partial \widehat{v}}{\partial n}_{|\Gamma}$, condition (36) is proved with $z = \frac{\partial \widehat{v}}{\partial n}$, which belongs to $L^2(\Gamma)$ since $\widehat{v} \in H^2(\Omega_2)$.

We conclude from Theorem 4.1 (i) and the preceding argument that if $\left(\frac{1}{\gamma_{n+1}} - \frac{1}{\gamma_n}\right)$ is bounded from above and if $(\gamma_n) \in l^2$, then any sequence (u_n, v_n) generated by $(\mathcal{A})$ weakly converges in $H^1(\Omega_1) \times H^1(\Omega_2)$ to a minimum point $(\widehat{u} + C, \widehat{v} + C)$, $(C \in \mathbb{R})$ of problem $(\mathcal{P})$. Without loss of generality, we can assume that $C = 0$. Since Ω_1 and Ω_2 are Lipschitz domains, the injections $H^1(\Omega_1) \hookrightarrow L^2(\Omega_1)$ and $H^1(\Omega_2) \hookrightarrow L^2(\Omega_2)$ are compact by the Rellich-Kondrachov Theorem (see [2, Theorem 6.2] or [17, Theorem II.55]). It ensues that the sequence (u_n, v_n) converges to $(\widehat{u}, \widehat{v})$ strongly in $L^2(\Omega_1) \times L^2(\Omega_2)$. Moreover, from Theorem 4.1 (ii), we have $\lim_{n\to+\infty} f(u_n) = f(\widehat{u})$ and $\lim_{n\to+\infty} g(v_n) = g(\widehat{v})$, hence $\lim_{n\to+\infty} \int_{\Omega_1} |\nabla u_n|^2 = \int_{\Omega_1} |\nabla \widehat{u}|^2$ and $\lim_{n\to+\infty} \int_{\Omega_2} |\nabla v_n|^2 = \int_{\Omega_2} |\nabla \widehat{v}|^2$. As a consequence, we have

[14] Observe that we have $R(A) = R(B) = H^{1/2}(\Gamma)$. Hence the set $R(A) + R(B) = H^{1/2}(\Gamma)$ is dense in $\mathcal{Z} = L^2(\Gamma)$ and condition $(\omega_n^-) \in l^1$ cannot be verified by using assertion (i) of proposition 4.1. This remark may suggest to take $\mathcal{Z} = H^{1/2}(\Gamma)$ endowed with the corresponding norm. In this case, the closedness of the set $R(A) + R(B)$ is automatically ensured. However the practical implementation of algorithm $(\mathcal{A})$ will be more complicated due to the use of the $H^{1/2}(\Gamma)$ norm. The details are out of the scope of the paper.

$$\lim_{n\to+\infty} \|(u_n, v_n)\|_{H^1(\Omega_1)\times H^1(\Omega_2)} = \|(\widehat{u}, \widehat{v})\|_{H^1(\Omega_1)\times H^1(\Omega_2)}.$$

Since (u_n, v_n) weakly converges in $H^1(\Omega_1)\times H^1(\Omega_2)$ toward $(\widehat{u}, \widehat{v})$, the convergence is strong in $H^1(\Omega_1) \times H^1(\Omega_2)$. We have proved the following:

Theorem 6.1 *Let $\Omega \subset \mathbb{R}^N$ be a bounded domain which can be decomposed in two nonoverlapping Lipschitz subdomains Ω_1 and Ω_2 with a common interface Γ. We assume that the set Ω is of class $\mathcal{C}^2$. Let $h \in L^2(\Omega)$ be such that $\int_\Omega h = 0$ and define the functions $f : H^1(\Omega_1) \to \mathbb{R}$ and $g : H^1(\Omega_2) \to \mathbb{R}$ by formulas (35). Assume that (γ_n) is a positive sequence such that $(\gamma_n) \in l^2$ and the sequence $\left(\frac{1}{\gamma_{n+1}} - \frac{1}{\gamma_n}\right)$ is bounded from above. Then any sequence (u_n, v_n) generated by algorithm $(\mathcal{A})$ strongly converges in $H^1(\Omega_1) \times H^1(\Omega_2)$ and the limit $(\widehat{u}, \widehat{v})$ is such that the map $\widehat{w} = \begin{cases} \widehat{u} \text{ on } \Omega_1 \\ \widehat{v} \text{ on } \Omega_2 \end{cases}$ is a solution of the Neumann problem (34).*

Algorithm $(\mathcal{A})$ allows to solve the initial Neumann problem on Ω by solving separately mixed Dirichlet-Neumann problems on Ω_1 and Ω_2. A similar method is developed in [14], where the authors consider alternating minimization algorithms based on augmented Lagrangian approach.

6.2 Problem with an obstacle

As a model situation, let us consider the variational problem with an obstacle constraint

$$\min \left\{ \frac{1}{2}\int_\Omega |\nabla w|^2 - \int_\Omega hw; \quad w \in H^1(\Omega),\ w \geq 0 \ \text{ on } \ \Omega \right\}. \tag{37}$$

It can be cast into our framework by taking

$$f(u) = \frac{1}{2}\int_{\Omega_1} |\nabla u|^2 - \int_{\Omega_1} hu + \delta_{C_1}(u) \quad \text{and} \quad g(v) = \frac{1}{2}\int_{\Omega_2} |\nabla v|^2 - \int_{\Omega_2} hv + \delta_{C_2}(v),$$

where δ_{C_1} is the indicator function of the convex set $C_1 = \{u \geq 0; \quad u \in H^1(\Omega_1)\}$ and δ_{C_2} is the indicator function of the convex set $C_2 = \{v \geq 0; \quad v \in H^1(\Omega_2)\}$. Problem (37) can be reformulated as

$$(\mathcal{P}) \qquad \min \left\{ f(u) + g(v); \quad (u, v) \in H^1(\Omega_1) \times H^1(\Omega_2) \text{ and } u_{|\Gamma} = v_{|\Gamma} \right\}.$$

Let us show that Attouch-Brézis qualification condition (QC) is satisfied in this situation (by contrast with Moreau-Rockafellar condition which fails to be satisfied for $N \geq 2$). Indeed, we are going to verify that

$$\text{dom} f \times \text{dom} g - \boldsymbol{\mathcal{V}} = H^1(\Omega_1) \times H^1(\Omega_2).$$

To that end we introduce two trace lifting operators

$$r_1 : H^{1/2}(\Gamma) \to H^1(\Omega_1)$$
$$r_2 : H^{1/2}(\Gamma) \to H^1(\Omega_2)$$

such that for every $z \in H^{1/2}(\Gamma), \quad z \geq 0 \Rightarrow r_i(z) \geq 0, \ i = 1, 2$. Such operators can be easily obtained by taking any lifting operator and then taking its positive part. Precisely, we use that for any $u \in H^1(\Omega_1)$, $u^+ = \max\{u, 0\} \in H^1(\Omega_1)$, $u^- = \max\{0, -u\} \in H^1(\Omega_1)$ and $u = u^+ - u^-$. Similarly, for any $v \in H^1(\Omega_2)$, $v^+ \in H^1(\Omega_2)$, $v^- \in H^1(\Omega_2)$ and $v = v^+ - v^-$. For any $u \in H^1(\Omega_1)$ and $v \in H^1(\Omega_2)$, we denote respectively by $u_{|\Gamma}$ and $v_{|\Gamma}$ their Sobolev traces on Γ. Let us now perform the following decomposition: for any $(u, v) \in H^1(\Omega_1) \times H^1(\Omega_2)$

$$\begin{aligned}(u, v) &= (u^+ - u^-, v)\\ &= \big(u^+, v + r_2\big((u^-)_{|\Gamma}\big)\big) - \big(u^-, r_2\big((u^-)_{|\Gamma}\big)\big). \end{aligned} \tag{38}$$

Let us notice that $\big(u^-, r_2\big((u^-)_{|\Gamma}\big)\big)$ belongs to $\boldsymbol{\mathcal{V}}$ because u^- and $r_2\big((u^-)_{|\Gamma}\big)$ have the same trace on Γ. Let us perform once more this operation: set $\mathrm{v} = v + r_2\big((u^-)_{|\Gamma}\big)$ which belongs to $H^1(\Omega_2)$.

$$\begin{aligned}(u^+, \mathrm{v}) &= (u^+, \mathrm{v}^+ - \mathrm{v}^-)\\ &= \big(u^+ + r_1\big((\mathrm{v}^-)_{|\Gamma}\big), \mathrm{v}^+\big) - \big(r_1\big((\mathrm{v}^-)_{|\Gamma}\big), \mathrm{v}^-\big). \end{aligned} \tag{39}$$

Combining (38) and (39) we finally obtain

$$(u, v) = \big(u^+ + r_1\big((\mathrm{v}^-)_{|\Gamma}\big), \mathrm{v}^+\big) - \Big[\big(r_1\big((\mathrm{v}^-)_{|\Gamma}\big), \mathrm{v}^-\big) + \big(u^-, r_2\big((u^-)_{|\Gamma}\big)\big)\Big].$$

By construction of the trace lifting operator, and by $\mathrm{v}^- \geq 0$ we have $r_1\big((\mathrm{v}^-)_{|\Gamma}\big) \geq 0$. Thus, we have obtained a decomposition of (u, v) as a difference of an element of $H^1(\Omega_1)^+ \times H^1(\Omega_2)^+$ and an element of $H^1(\Omega)$. The decomposition algorithm can now be developed in a very similar way as in the unconstrained case.

References

[1] F. Acker, M.A. Prestel, Convergence d'un schéma de minimisation alternée, *Annales de la Faculté des Sciences de Toulouse V, série Mathématiques*, 2, (1980), pp. 1-9.

[2] R. Adams, Sobolev spaces, Academic Press, New-York, 1975.

[3] S. Agmon, A. Douglis, L. Nirenberg, Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions I., *Comm. Pure Appl. Math.*, 12 (1959), pp. 623-727.

[4] S. Agmon, A. Douglis, L. Nirenberg, Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions II., *Comm. Pure Appl. Math.*, 17 (1964), pp. 35-92.

[5] F. Álvarez, R. Cominetti, Primal and dual convergence of a proximal point exponential penalty method for linear programming, *Math. Programming*, 93 (2002), pp. 87-96.

[6] F. Álvarez, J. Peypouquet, Asymptotic almost-equivalence of Lipschitz evolution systems in Banach spaces. *Nonlinear Anal.*, 73 (2010), no. 9, 3018-3033.

[7] H. Attouch, J. Bolte, P. Redont, A. Soubeyran, Alternating proximal algorithms for weakly coupled convex minimization problems. Applications to dynamical games and PDE's, *Journal of Convex Analysis*, 15 (2008), No 3, pp. 485-506.

[8] H. Attouch, H. Brézis, Duality for the sum of convex functions in general Banach spaces, *Aspects of Mathematics and Its Applications*, J. Barroso Ed., North-Holland, Amsterdam (1986), pp. 125-133.

[9] H. Attouch, G. Buttazzo, G. Michaille, Variational analysis in Sobolev and BV spaces. Applications to PDE's and optimization, *MPS/SIAM Series on Optimization*, 6. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2006.

[10] H. Attouch, M.-O. Czarnecki, Asymptotic control and stabilization of nonlinear oscillators with non isolated equilibria, *Journal of Differential Equations*, 179 (2002), pp. 278-310.

[11] H. Attouch, M.-O. Czarnecki, Asymptotic behavior of coupled dynamical systems with multiscale aspects, *Journal of Differential Equations*, 248 (2010), pp. 1315-1344.

[12] H. Attouch, M.-O. Czarnecki, J. Peypouquet, Prox-penalization and splitting methods for constrained variational problems, *SIAM Journal on Optimization*, 21 (2011), pp. 149-173.

[13] H. Attouch, P. Redont, A. Soubeyran, A new class of alternating proximal minimization algorithms with costs-to-move, *SIAM Journal on Optimization*, 18 (2007), pp. 1061-1081.

[14] H. Attouch, M. Soueycatt, Augmented Lagrangian and proximal alternating direction methods of multipliers in Hilbert spaces. Applications to games , PDE's and control, *Pacific Journal of Optimization*, 5 (2009), No 1, pp. 17-37.

[15] A. Auslender, Numerical methods for nondifferentiable convex optimization, *Math. Programming Stud.*, 30 (1987), pp. 102-126.

[16] A. Auslender, J.-P. Crouzeix, P. Fedit, Penalty-proximal methods in convex programming, *J. Optim. Theory Appl.*, 55 (1987), pp. 1-21.

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

[18] J.-B. Baillon, R. Cominetti, A convergence result for non-autonomous subgradient evolution equations and its application to the steepest descent exponential penalty trajectory in linear programming, *J. Funct. Anal.*, 187 (2001), pp. 263-273.

[19] H.H. Bauschke, P.L. Combettes, S. Reich, The asymptotic behavior of the composition of two resolvents, *Nonlinear Anal.*, 60 (2005), pp. 283-301.

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

[21] H. Brézis, Analyse fonctionnelle: théorie et applications, Dunod, Paris, 1999.

[22] A. Cabot, Proximal point algorithm controlled by a slowly vanishing term: Applications to hierarchical minimization, *SIAM Journal on Optimization*, 15 (2005), No 2, pp. 555-572.

[23] A. Cabot, P. Frankel, Alternating proximal algorithms with costs-to-move and asymptotically vanishing coupling, *paper under review.*

[24] R. Cominetti, M. Courdourier, Coupling general penalty schemes for convex programming with the steepest descent method and the proximal point algorithm, *SIAM J. Optim*, **13** (2002), 745-765.

[25] I. Ekeland, R. Temam, Convex Analysis and Variational Problems, SIAM Classics in Applied Mathematics, 28, 1999.

[26] D. Gilbarg, N. Trudinger, Elliptic partial differential equations of second order, Springer-Verlag, Berlin, 1977.

[27] J.-L. Lions, E. Magenes, Non-homogeneous boundary value problems and applications, Die Grundlehren der mathematischen Wissenschaften, Springer-Verlag, New York-Heidelberg, 1972.

[28] J. Marschall, The trace of Sobolev-Slobodeckij spaces on Lipschitz domains, *Manuscripta Math.*, **58** (1987), pp. 47–65.

[29] J.J. Moreau, Fonctionnelles convexes, Cours Collège de France 1967, new edition CNR, Facoltà di Ingegneria di Roma, 2003.

[30] Z. Opial, Weak convergence of the sequence of successive approximations for nonexpansive mappings, *Bull. Amer. Math. Soc.*, **73** (1967), pp. 591-597.

[31] G. Passty , Ergodic convergence to a zero of the sum of monotone operators in Hilbert space, *J. Math. Anal. Appl.*, **72** (1979), no. 2, pp. 383-390.

[32] B.T. Polyak, Introduction to optimization, Opimization Software, Inc. Publications Division, New york, 1987.

[33] P.-A. Raviart, J.-M. Thomas, Introduction à l'analyse numérique des équations aux dérivées partielles, Masson, Paris, 1993.

[34] R.T. Rockafellar, Convex Analysis, Princeton University Press, Princeton, 1970.

[35] R.T. Rockafellar, Monotone operators and the proximal point algorithm, *SIAM Journal on Control and Optimization*, 14 (1976), pp. 877-898.

[36] R.T. Rockafellar, R. Wets, Variational analysis, Springer, Berlin, 1998.

[37] M.-V. Solodov, B.-F. Svaiter, An Inexact Hybrid Extragradient-Proximal Point Algorithm using the Enlargement of a Maximal Monotone Operator. *Set-Valued Analysis*, 7 (1999) 323–345.

V. ALGORITHMES PROXIMAUX ALTERNES AVEC COUPLAGE ASYMPTOTIQUEMENT EVANESCENT. APPLICATION A LA DECOMPOSITION DE DOMAINE POUR LES EDP

Alternating proximal algorithms with asymptotically vanishing coupling. Application to domain decomposition for PDE's

A. Cabot[1] and P. Frankel[2]

[1] Université Montpellier II, Place Eugène Bataillon, 34095 Montpellier Cedex 05, France
acabot@math.univ-montp2.fr
[2] Université Montpellier II, Place Eugène Bataillon, 34095 Montpellier Cedex 05, France
p.frankel30@orange.fr

Accepted in *Optimization*

Summary. Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ be real Hilbert spaces, let $f : \mathcal{X} \longrightarrow \mathbb{R} \cup \{+\infty\}$, $g : \mathcal{Y} \longrightarrow \mathbb{R} \cup \{+\infty\}$ be closed convex functions and let $A : \mathcal{X} \longrightarrow \mathcal{Z}$, $B : \mathcal{Y} \longrightarrow \mathcal{Z}$ be linear continuous operators. Given a sequence (γ_n) which increases toward infinity as $n \to +\infty$, we study the following alternating proximal algorithm

$$(\mathcal{A}) \qquad \begin{cases} x_{n+1} = \text{Argmin}\{f(\zeta) + \dfrac{1}{2\gamma_{n+1}}\|A\zeta - By_n\|^2_{\mathcal{Z}} + \dfrac{\alpha}{2}\|\zeta - x_n\|^2_{\mathcal{X}}; \quad \zeta \in \mathcal{X}\} \\ y_{n+1} = \text{Argmin}\{g(\eta) + \dfrac{1}{2\gamma_{n+1}}\|Ax_{n+1} - B\eta\|^2_{\mathcal{Z}} + \dfrac{\nu}{2}\|\eta - y_n\|^2_{\mathcal{Y}}; \quad \eta \in \mathcal{Y}\}, \end{cases}$$

where α and ν are positive parameters. If the sequence (γ_n) increases *moderately slowly* toward infinity, the algorithm $(\mathcal{A})$ tends to minimize the function $(x,y) \mapsto \|Ax - By\|^2_{\mathcal{Z}}$ over the set $C = \text{Argmin} f \times \text{Argmin} g$ (assumed to be nonempty). An illustration of this result is given in the area of domain decomposition for PDE's.

Key words: Convex minimization, alternating minimization, proximal algorithm, hierarchical minimization, domain decomposition for PDE's.
Subject classification: 65K05, 65K10, 49J40, 90C25.

1 Introduction

Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ be real Hilbert spaces respectively endowed with the scalar products $\langle .,. \rangle_{\mathcal{X}}$, $\langle .,. \rangle_{\mathcal{Y}}$ and $\langle .,. \rangle_{\mathcal{Z}}$ and the corresponding norms. Let $f : \mathcal{X} \longrightarrow \mathbb{R} \cup \{+\infty\}$, $g : \mathcal{Y} \longrightarrow \mathbb{R} \cup \{+\infty\}$ be closed convex proper functions and let $A : \mathcal{X} \longrightarrow \mathcal{Z}$, $B : \mathcal{Y} \longrightarrow \mathcal{Z}$ be linear continuous operators. We consider the convex function $\Phi_\gamma : \mathcal{X} \times \mathcal{Y} \longrightarrow \mathbb{R} \cup \{+\infty\}$ defined by

$$\Phi_\gamma(x,y) = f(x) + g(y) + \frac{1}{2\gamma}\|Ax - By\|^2_{\mathcal{Z}},$$

where γ is a positive real parameter. In order to minimize the function Φ_γ, Attouch, Redont and Soubeyran [7] introduced the alternating algorithm with costs-to-move

$$\begin{cases} x_{n+1} = \text{Argmin}\{f(\zeta) + \dfrac{1}{2\gamma}\|A\zeta - By_n\|^2_{\mathcal{Z}} + \dfrac{\alpha}{2}\|\zeta - x_n\|^2_{\mathcal{X}}; \quad \zeta \in \mathcal{X}\} \\ y_{n+1} = \text{Argmin}\{g(\eta) + \dfrac{1}{2\gamma}\|Ax_{n+1} - B\eta\|^2_{\mathcal{Z}} + \dfrac{\nu}{2}\|\eta - y_n\|^2_{\mathcal{Y}}; \quad \eta \in \mathcal{Y}\}, \end{cases}$$

where α and ν are positive real numbers. This algorithm generates a sequence (x_n, y_n) whose convergence is studied in [4]. It is proved that, if $\text{Argmin}\Phi_\gamma \neq \varnothing$, the sequence (x_n, y_n) weakly converges toward a minimum of Φ_γ.

The framework of [4, 7] extends the one of [1, 10] from the strong coupled problem to the weak coupled problem with costs-to-change. More precisely, $Q(x,y) = \|x-y\|^2_{\mathcal{Z}}$ is a strong coupling function with $\mathcal{X} = \mathcal{Y} = \mathcal{Z}$ and $A = B = \mathcal{I}$ while $Q(x,y) = \|Ax - By\|^2_{\mathcal{Z}}$ is now a weak coupling function which allows for asymmetric and partial relations between the variables x and y. Furthermore authors of [1, 10] do not use costs-to-changes $(\alpha/2)\|\zeta - x_n\|^2_{\mathcal{X}}$, $\zeta \in \mathcal{X}$ and $(\nu/2)\|\eta - y_n\|^2_{\mathcal{Y}}$, $\eta \in \mathcal{Y}$, taking $\alpha = \nu = 0$. The interest of the weak coupling term is to cover many situations, ranging from decomposition methods for PDE's to applications in game theory. In decision sciences, the term $Q(x,y) = \|Ax - By\|^2_{\mathcal{Z}}$ allows to consider agents who interplay, only via some components of their decision variables. For further details, the interested reader is referred to [4].

In this study, the constant parameter γ of the above mentioned algorithm is replaced by a sequence (γ_n) which increases toward infinity as $n \to +\infty$. The corresponding algorithm is denoted by $(\mathcal{A})$

$$(\mathcal{A}) \quad \begin{cases} x_{n+1} = \text{Argmin}\{f(\zeta) + \dfrac{1}{2\gamma_{n+1}}\|A\zeta - By_n\|^2_{\mathcal{Z}} + \dfrac{\alpha}{2}\|\zeta - x_n\|^2_{\mathcal{X}}; \quad \zeta \in \mathcal{X}\} \\ y_{n+1} = \text{Argmin}\{g(\eta) + \dfrac{1}{2\gamma_{n+1}}\|Ax_{n+1} - B\eta\|^2_{\mathcal{Z}} + \dfrac{\nu}{2}\|\eta - y_n\|^2_{\mathcal{Y}}; \quad \eta \in \mathcal{Y}\}. \end{cases}$$

The coupling term asymptotically vanishes as $n \to +\infty$. Assuming that the sequence $\left(\frac{1}{\gamma_n}\right)$ is summable, we show that any sequence (x_n, y_n) generated by $(\mathcal{A})$ weakly converges toward a point of $C = \text{Argmin}f \times \text{Argmin}g$ (assumed to be nonempty). The limit does not depend explicitly on the operators A and B because the sequence (γ_n) tends too fast toward infinity. Now consider the case corresponding to $\left(\frac{1}{\gamma_n}\right) \notin l^1$. We prove that, if the sequence (γ_n) increases *moderately slowly*[3], algorithm $(\mathcal{A})$ tends to minimize the function $(x,y) \mapsto \|Ax - By\|^2_{\mathcal{Z}}$ over the set C.

[3] For example, if the functions f and g behave as the square of the distance to their respective argmin sets, the condition on (γ_n) becomes $\left(\frac{1}{\gamma_n}\right) \in l^2 \setminus l^1$.

We apply our abstract results to the framework of splitting methods for PDE's. For that purpose, we consider a domain $\Omega \subset \mathbb{R}^N$ that can be decomposed into two non overlapping subdomains Ω_1, Ω_2 with a common interface Γ. The functional spaces are $\mathcal{X} = H^1(\Omega_1)$, $\mathcal{Y} = H^1(\Omega_2)$ and $\mathcal{Z} = L^2(\Gamma)$, the operators $A : \mathcal{X} \to \mathcal{Z}$ and $B : \mathcal{Y} \to \mathcal{Z}$ being respectively the trace operators on Γ. The term $Au - Bv$ corresponds to the jump of the map $w = \begin{cases} u \text{ on } \Omega_1 \\ v \text{ on } \Omega_2 \end{cases}$ through the interface Γ. Let us consider the set of couples $(u, v) \in H^1(\Omega_1) \times H^1(\Omega_2)$ of solutions that satisfy some boundary value problems, respectively on Ω_1 and Ω_2. Under suitable assumptions, the iterates (u_n, v_n) of algorithm $(\mathcal{A})$ tend toward a couple solution that minimizes the $L^2(\Gamma)$-norm of the jump through the interface Γ.

The paper is organized as follows. General results for algorithm $(\mathcal{A})$ are given in section 2, including the case $(\frac{1}{\gamma_n}) \in l^1$. The case $(\frac{1}{\gamma_n}) \notin l^1$ is analyzed in section 3 and an application to decomposition domain for PDE's is illustrated in section 4. Further convergence results in the finite dimensional setting are given in section 5.

2 General results

Let $f : \mathcal{X} \longrightarrow \mathbb{R} \cup \{+\infty\}$, $g : \mathcal{Y} \longrightarrow \mathbb{R} \cup \{+\infty\}$ be closed convex proper functions and let $A : \mathcal{X} \longrightarrow \mathcal{Z}$, $B : \mathcal{Y} \longrightarrow \mathcal{Z}$ be linear continuous operators. Let (γ_n) be a nondecreasing sequence of positive reals such that $\lim_{n\to+\infty} \gamma_n = +\infty$. Given positive coefficients $\alpha, \nu > 0$ and initial data $(x_0, y_0) \in \mathcal{X} \times \mathcal{Y}$, let us consider the following alternating proximal algorithm

$$(\mathcal{A}) \quad \begin{cases} x_{n+1} = \text{Argmin}\{f(\zeta) + \dfrac{1}{2\gamma_{n+1}}\|A\zeta - By_n\|^2_{\mathcal{Z}} + \dfrac{\alpha}{2}\|\zeta - x_n\|^2_{\mathcal{X}}; \quad \zeta \in \mathcal{X}\} \\ y_{n+1} = \text{Argmin}\{g(\eta) + \dfrac{1}{2\gamma_{n+1}}\|Ax_{n+1} - B\eta\|^2_{\mathcal{Z}} + \dfrac{\nu}{2}\|\eta - y_n\|^2_{\mathcal{Y}}; \quad \eta \in \mathcal{Y}\}. \end{cases}$$

The coupling term asymptotically vanishes as $n \to +\infty$. It is clear from the definition of the sequence (x_n, y_n) that nothing is changed if some constant is added to the function f (resp. g). We will assume in the sequel that $\inf f = \inf g = 0$.
By writing down the optimality conditions, it is immediate to check that points x_{n+1} and y_{n+1} are characterized by

$$\begin{cases} -\dfrac{1}{\gamma_{n+1}} A^*(Ax_{n+1} - By_n) - \alpha(x_{n+1} - x_n) \in \partial f(x_{n+1}) \\ \dfrac{1}{\gamma_{n+1}} B^*(Ax_{n+1} - By_{n+1}) - \nu(y_{n+1} - y_n) \in \partial g(y_{n+1}), \end{cases}$$

where $A^* \in \mathcal{L}(\mathcal{Z}, \mathcal{X})$ and $B^* \in \mathcal{L}(\mathcal{Z}, \mathcal{Y})$ denote the respective adjoint operators of A and B. It ensues that we have, for all $x \in \mathcal{X}$ and $y \in \mathcal{Y}$,

$$\begin{cases} f(x) - f(x_{n+1}) - \dfrac{1}{\gamma_{n+1}}\langle Ax_{n+1} - By_n, Ax_{n+1} - Ax\rangle_{\mathcal{Z}} - \alpha\langle x_{n+1} - x_n, x_{n+1} - x\rangle_{\mathcal{X}} \geq 0 \\ g(y) - g(y_{n+1}) + \dfrac{1}{\gamma_{n+1}}\langle Ax_{n+1} - By_{n+1}, By_{n+1} - By\rangle_{\mathcal{Z}} - \nu\langle y_{n+1} - y_n, y_{n+1} - y\rangle_{\mathcal{Y}} \geq 0. \end{cases} \tag{1}$$

These inequalities will be useful in the sequel. We first give general results which do not depend on the growth speed of the sequence (γ_n) as $n \to +\infty$.

Proposition 2.1 *Let α, $\nu > 0$ and let $A \in \mathcal{L}(\mathcal{X}, \mathcal{Z})$, $B \in \mathcal{L}(\mathcal{Y}, \mathcal{Z})$ be linear continuous operators. Let $f : \mathcal{X} \longrightarrow \mathbb{R}\cup\{+\infty\}$ and $g : \mathcal{Y} \longrightarrow \mathbb{R}\cup\{+\infty\}$ be closed convex functions such that* $\inf f = \inf g = 0$*. Let (γ_n) be a positive nondecreasing sequence such that* $\lim_{n\to+\infty} \gamma_n = +\infty$*. Then, for any sequence (x_n, y_n) generated by algorithm $(\mathcal{A})$ we have*

(i) the sequence $\left(f(x_n) + g(y_n) + \frac{1}{2\gamma_n}\|Ax_n - By_n\|^2_{\mathcal{Z}}\right)$ is nonincreasing and tends toward 0 *as $n \to +\infty$. As a consequence,* $\lim_{n\to+\infty} f(x_n) = \lim_{n\to+\infty} g(y_n) = 0$ *and every weak limit point of the sequence (x_n, y_n) belongs to Argminf $\times$ Argming;*

(ii) the sequences $(\|x_{n+1} - x_n\|^2_{\mathcal{X}})$ and $(\|y_{n+1} - y_n\|^2_{\mathcal{Y}})$ are summable.

Proof. The arguments are similar to those of [10, Theorem 4.6].
(i) Let us set $\theta_n = f(x_n) + g(y_n) + \frac{1}{2\gamma_n}\|Ax_n - By_n\|^2_{\mathcal{Z}}$. From the definition of algorithm $(\mathcal{A})$, we have

$$f(x_{n+1}) + \frac{1}{2\gamma_{n+1}}\|Ax_{n+1} - By_n\|^2_{\mathcal{Z}} + \frac{\alpha}{2}\|x_{n+1} - x_n\|^2_{\mathcal{X}} \leq f(x_n) + \frac{1}{2\gamma_{n+1}}\|Ax_n - By_n\|^2_{\mathcal{Z}},$$
$$g(y_{n+1}) + \frac{1}{2\gamma_{n+1}}\|Ax_{n+1} - By_{n+1}\|^2_{\mathcal{Z}} + \frac{\nu}{2}\|y_{n+1} - y_n\|^2_{\mathcal{Y}} \leq g(y_n) + \frac{1}{2\gamma_{n+1}}\|Ax_{n+1} - By_n\|^2_{\mathcal{Z}}.$$

By using these inequalities, we deduce successively that, for all $n \in \mathbb{N}$,

$$\begin{aligned} \theta_{n+1} &\leq f(x_{n+1}) + g(y_n) + \frac{1}{2\gamma_{n+1}}\|Ax_{n+1} - By_n\|^2_{\mathcal{Z}} \\ &\leq f(x_n) + g(y_n) + \frac{1}{2\gamma_{n+1}}\|Ax_n - By_n\|^2_{\mathcal{Z}}. \end{aligned}$$

Since the sequence (γ_n) is nondecreasing, we finally find $\theta_{n+1} \leq \theta_n$. Let us now use the following lemma borrowed from [10].

Lemma 2.1 *Let $(s, t, u, v, w) \in \mathcal{Z}^5$, then*

$$\begin{aligned} \|s - u\|^2_{\mathcal{Z}} = \|s - w\|^2_{\mathcal{Z}} + \|w - v\|^2_{\mathcal{Z}} - \|s - t\|^2_{\mathcal{Z}} + \|(s - t) - (u - v)\|^2_{\mathcal{Z}} \\ +2\langle s - w, w - v\rangle_{\mathcal{Z}} + 2\langle u - v, v - t\rangle_{\mathcal{Z}}. \end{aligned}$$

For $x \in \mathcal{X}, y \in \mathcal{Y}$, we take $s = By,\ t = Ax,\ u = By_n,\ v = Ax_{n+1},\ w = By_{n+1}$. We obtain

$$\begin{aligned}\|By-By_n\|^2_{\mathcal Z}-\|By-By_{n+1}\|^2_{\mathcal Z} = {} & \|By_{n+1}-Ax_{n+1}\|^2_{\mathcal Z}-\|By-Ax\|^2_{\mathcal Z}\\ & +\|By-Ax-(By_n-Ax_{n+1})\|^2_{\mathcal Z}\\ & +2\langle By-By_{n+1},By_{n+1}-Ax_{n+1}\rangle_{\mathcal Z}\\ & +2\langle By_n-Ax_{n+1},Ax_{n+1}-Ax\rangle_{\mathcal Z}.\end{aligned}$$

Using inequalities (1), we deduce that

$$\begin{aligned}&\|By-By_n\|^2_{\mathcal Z}-\|By-By_{n+1}\|^2_{\mathcal Z}\\ &\geq \|Ax_{n+1}-By_{n+1}\|^2_{\mathcal Z}-\|Ax-By\|^2_{\mathcal Z}\\ &\quad +2\gamma_{n+1}[f(x_{n+1})-f(x)+\alpha\langle x_{n+1}-x_n,x_{n+1}-x\rangle_{\mathcal X}]\\ &\quad +2\gamma_{n+1}[g(y_{n+1})-g(y)+\nu\langle y_{n+1}-y_n,y_{n+1}-y\rangle_{\mathcal Y}]\\ &=2\gamma_{n+1}\left[f(x_{n+1})+g(y_{n+1})-f(x)-g(y)\right]+\|Ax_{n+1}-By_{n+1}\|^2_{\mathcal Z}-\|Ax-By\|^2_{\mathcal Z}\\ &\quad +\gamma_{n+1}\alpha(\|x_{n+1}-x_n\|^2_{\mathcal X}+\|x_{n+1}-x\|^2_{\mathcal X}-\|x_n-x\|^2_{\mathcal X})\\ &\quad +\gamma_{n+1}\nu(\|y_{n+1}-y_n\|^2_{\mathcal Y}+\|y_{n+1}-y\|^2_{\mathcal Y}-\|y_n-y\|^2_{\mathcal Y}).\end{aligned}$$

Finally, we obtain

$$\begin{aligned}&\|By-By_n\|^2_{\mathcal Z}+\gamma_{n+1}\alpha\|x_n-x\|^2_{\mathcal X}+\gamma_{n+1}\nu\|y_n-y\|^2_{\mathcal Y}\\ &-\|By-By_{n+1}\|^2_{\mathcal Z}-\gamma_{n+1}\alpha\|x_{n+1}-x\|^2_{\mathcal X}-\gamma_{n+1}\nu\|y_{n+1}-y\|^2_{\mathcal Y}\\ &\quad\geq 2\gamma_{n+1}\left[f(x_{n+1})+g(y_{n+1})-f(x)-g(y)\right] \qquad (2)\\ &+\|Ax_{n+1}-By_{n+1}\|^2_{\mathcal Z}-\|Ax-By\|^2_{\mathcal Z}+\gamma_{n+1}\alpha\|x_{n+1}-x_n\|^2_{\mathcal X}+\gamma_{n+1}\nu\|y_{n+1}-y_n\|^2_{\mathcal Y}.\end{aligned}$$

Dividing by γ_{n+1} and using $\frac{1}{\gamma_{n+1}}\leq\frac{1}{\gamma_n}$, we deduce that

$$\begin{aligned}&\frac{1}{\gamma_n}\|By-By_n\|^2_{\mathcal Z}+\alpha\|x_n-x\|^2_{\mathcal X}+\nu\|y_n-y\|^2_{\mathcal Y}\\ &-\frac{1}{\gamma_{n+1}}\|By-By_{n+1}\|^2_{\mathcal Z}-\alpha\|x_{n+1}-x\|^2_{\mathcal X}-\nu\|y_{n+1}-y\|^2_{\mathcal Y}\\ &\quad\geq 2\left[f(x_{n+1})+g(y_{n+1})-f(x)-g(y)\right] \qquad (3)\\ &+\frac{1}{\gamma_{n+1}}\|Ax_{n+1}-By_{n+1}\|^2_{\mathcal Z}-\frac{1}{\gamma_{n+1}}\|Ax-By\|^2_{\mathcal Z}+\alpha\|x_{n+1}-x_n\|^2_{\mathcal X}+\nu\|y_{n+1}-y_n\|^2_{\mathcal Y}.\end{aligned}$$

Let us set $h_n=\frac{1}{\gamma_n}\|By-By_n\|^2_{\mathcal Z}+\alpha\|x_n-x\|^2_{\mathcal X}+\nu\|y_n-y\|^2_{\mathcal Y}$. The previous inequality implies that

$$\begin{aligned}&h_{n+1}-h_n+2\left[f(x_{n+1})+g(y_{n+1})-f(x)-g(y)\right]\\ &\qquad+\frac{1}{\gamma_{n+1}}\|Ax_{n+1}-By_{n+1}\|^2_{\mathcal Z}-\frac{1}{\gamma_{n+1}}\|Ax-By\|^2_{\mathcal Z}\leq 0, \qquad (4)\end{aligned}$$

or equivalently

$$2\theta_{n+1}\leq 2(f(x)+g(y))+\frac{1}{\gamma_{n+1}}\|Ax-By\|^2_{\mathcal Z}+h_n-h_{n+1}.$$

Let us sum from $n = 1$ to N. Since $h_{N+1} \geq 0$, we infer that

$$2\sum_{n=1}^{N} \theta_{n+1} \leq 2\,N\,(f(x) + g(y)) + \left(\sum_{n=1}^{N} \frac{1}{\gamma_{n+1}}\right) \|Ax - By\|^2_{\mathcal{Z}} + h_1.$$

Using that the sequence (θ_n) is nonincreasing, we deduce that

$$\theta_{N+1} \leq f(x) + g(y) + \frac{1}{2N}\left(\sum_{n=1}^{N} \frac{1}{\gamma_{n+1}}\right) \|Ax - By\|^2_{\mathcal{Z}} + \frac{h_1}{2N}.$$

Take now the upper limit as $N \to +\infty$ in the above inequality. By using the fact that $\lim_{N\to+\infty} \frac{1}{\gamma_N} = 0$, we derive that $\limsup_{N\to+\infty} \theta_{N+1} \leq f(x) + g(y)$. Since this is true for every $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ and since $\inf f = \inf g = 0$, we infer that $\limsup_{N\to+\infty} \theta_{N+1} \leq 0$. Recalling that the sequence (θ_N) is nonnegative, we conclude that $\lim_{N\to+\infty} \theta_N = 0$, which in turn implies that $\lim_{N\to+\infty} f(x_N) = \lim_{N\to+\infty} g(y_N) = 0$. Finally, let (x_{n_k}, y_{n_k}) be a subsequence which weakly converges toward $(\overline{x}, \overline{y})$. Using the closedness of f and g, we find

$$f(\overline{x}) + g(\overline{y}) \leq \liminf_{k\to+\infty} f(x_{n_k}) + \liminf_{k\to+\infty} g(y_{n_k}) = \lim_{n\to+\infty} f(x_n) + \lim_{n\to+\infty} g(y_n) = 0,$$

hence $(\overline{x}, \overline{y}) \in \text{Argmin} f \times \text{Argmin} g$.
(ii) By applying inequality (3) with $x = x_n$, $y = y_n$ and recalling that $\gamma_n \leq \gamma_{n+1}$, we deduce that

$$\theta_{n+1} - \theta_n \leq -\alpha\|x_{n+1} - x_n\|^2_{\mathcal{X}} - \nu\|y_{n+1} - y_n\|^2_{\mathcal{Y}}.$$

Let us sum from $n = 1$ to N to derive that

$$\alpha \sum_{n=1}^{N} \|x_{n+1} - x_n\|^2_{\mathcal{X}} + \nu \sum_{n=1}^{N} \|y_{n+1} - y_n\|^2_{\mathcal{Y}} \leq \theta_1 - \theta_{N+1} \leq \theta_1.$$

It suffices then to let N tend to infinity.

When the sequence $(\frac{1}{\gamma_n})$ is summable, we can easily establish the weak convergence of the algorithm toward a point of $C = \text{Argmin} f \times \text{Argmin} g$. Notice that the assumption $(\frac{1}{\gamma_n}) \in l^1$ means that (γ_n) increases fast toward infinity.

Proposition 2.2 *Under the hypotheses of Proposition 2.1, assume moreover that $C = Argminf \times Argming \neq \varnothing$ and that $(\frac{1}{\gamma_n}) \in l^1$. If (x_n, y_n) is a sequence generated by $(\mathcal{A})$, then*

(i) for every $(x, y) \in C$, the sequence $(\alpha\,\|x_n - x\|^2_{\mathcal{X}} + \nu\,\|y_n - y\|^2_{\mathcal{Y}})$ is convergent;
(ii) (x_n, y_n) weakly converges toward a point of C.

Proof. (i) Fix $(x,y) \in C$ and set $h_n = \frac{1}{\gamma_n}\|By - By_n\|^2_{\mathcal{Z}} + \alpha\|x_n - x\|^2_{\mathcal{X}} + \nu\|y_n - y\|^2_{\mathcal{Y}}$ as in the proof of Proposition 2.1. From inequality (4) we deduce that

$$h_{n+1} \le h_n + \frac{1}{\gamma_{n+1}}\|Ax - By\|^2_{\mathcal{Z}}.$$

Since the sequence $(\frac{1}{\gamma_{n+1}})$ is summable, we can apply Lemma 2.2 hereafter to prove the convergence of the sequence (h_n). We deduce that the sequence (y_n) is bounded, which in turn implies that $\frac{1}{\gamma_n}\|By - By_n\|^2_{\mathcal{Z}} \to 0$ when $n \to +\infty$. This achieves the proof of (i).

Lemma 2.2 *Let (a_n) and (ε_n) be two real sequences. Assume that (a_n) is minorized, that (ε_n) is summable and that $a_{n+1} \le a_n + \varepsilon_n$ for every $n \in \mathbb{N}$. Then $\lim_{n\to+\infty} a_n$ exists.*

Proof of Lemma 2.2. Define the sequence (w_n) by $w_n = a_n - \sum_{k=0}^{n-1} \varepsilon_k$. The sequence (w_n) is bounded from below and nonincreasing, hence convergent. It follows that $\lim_{n\to+\infty} a_n = \sum_{k=0}^{+\infty} \varepsilon_k + \lim_{n\to+\infty} w_n$.
(ii) The proof of the weak convergence relies on the Opial's lemma [16], that we recall below for the sake of completeness.

Lemma 2.3 (Opial) *Let H be a Hilbert space endowed with the norm N. Let (ξ_n) be a sequence of H such that there exists a nonempty set $\mathcal{S} \subset H$ which verifies*

(a)For all $\xi \in \mathcal{S}$, $\lim_{n\to+\infty} N(\xi_n - \xi)$ exists.
(b)If $(\xi_{n_k}) \rightharpoonup \overline{\xi}$ weakly in H as $k \to +\infty$, we have $\overline{\xi} \in \mathcal{S}$.

Then the sequence (ξ_n) weakly converges in H as $n \to +\infty$ toward a point of $\mathcal{S}$.

Let us define the norm $N(u,v) = (\alpha\|u\|^2_{\mathcal{X}} + \nu\|v\|^2_{\mathcal{Y}})^{1/2}$ on the space $H = \mathcal{X} \times \mathcal{Y}$. Norm N is clearly equivalent to the canonical norm on $\mathcal{X} \times \mathcal{Y}$. In view of (i), $N((x_n,y_n) - (x,y))$ does have a limit for every $(x,y) \in C$, which shows point (a). On the other hand, point (b) is a consequence of Proposition 2.1 (i). Hence we conclude from Opial's lemma that the sequence (x_n,y_n) weakly converges in $\mathcal{X} \times \mathcal{Y}$ toward a point of C.

3 Case of a slowly increasing parameter

The purpose of this section is to study the case $(\frac{1}{\gamma_n}) \notin l^1$ and to bring to light a phenomenon of *selection* with respect to the viscosity function $(x,y) \mapsto \|Ax - By\|^2_{\mathcal{Z}}$. Assuming that the sets Argmin$f$ and Argmin$g$ are nonempty and that $\min f = \min g = 0$, let us consider the following hypotheses introduced by Attouch-Czarnecki [6]

$$(H_f) \qquad \forall p \in R(N_{\text{Argmin}f}), \qquad \left(f^*\left(\frac{p}{\gamma_n}\right) - \sigma_{\text{Argmin}f}\left(\frac{p}{\gamma_n}\right)\right) \in l^1$$

$$(H_g) \qquad \forall q \in R(N_{\text{Argmin}_g}), \qquad \left(g^*\left(\frac{q}{\gamma_n}\right) - \sigma_{\text{Argmin}_g}\left(\frac{q}{\gamma_n}\right)\right) \in l^1.$$

Let us explicit notations: f^* is the Fenchel conjugate of f defined by $f^*(x) = \sup_{\zeta\in\mathcal{X}}\{\langle x,\zeta\rangle_{\mathcal{X}} - f(\zeta)\}$ for every $x \in \mathcal{X}$. Given a subset $D \subset \mathcal{X}$, σ_D is the support function of D: $\sigma_D(x) = \sup_{\zeta\in D}\langle x,\zeta\rangle_{\mathcal{X}}$ for every $x \in \mathcal{X}$. Notice that σ_D coincides with the Fenchel conjugate $(\delta_D)^*$ of the indicator function δ_D. On the other hand, $N_D(x)$ is the normal cone to D at x,

$$N_D(x) = \{p \in \mathcal{X} : \langle p, \zeta - x\rangle_{\mathcal{X}} \leq 0 \quad \forall \zeta \in D\}.$$

$R(N_D)$ is the range of N_D, *i.e.* $p \in R(N_D)$ if and only if $p \in N_D(x)$ for some $x \in D$. Remark that, from the inequality $f \leq \delta_{\text{Argmin}f}$, we get $f^* \geq (\delta_{\text{Argmin}f})^* = \sigma_{\text{Argmin}f}$. In a similar way, we have $g^* \geq \sigma_{\text{Argmin}_g}$, hence the sequences arising in (H_f)-(H_g) are nonnegative.

Example 3.1 Let us illustrate assumptions (H_f) and (H_g). Since they are symmetric, we focus on (H_f) and we suppose that there exist $a > 0$ and $r \geq 1$ such that

$$f \geq a\, d_{\mathcal{X}}^r(., \text{Argmin} f). \tag{5}$$

The notation $d_{\mathcal{X}}(., \text{Argmin} f)$ stands for the distance function to the set Argmin$f$. We have

$$d_{\mathcal{X}}^r(., \text{Argmin} f) = \|.\|_{\mathcal{X}}^r +_e \delta_{\text{Argmin}f},$$

where $+_e$ denotes the epigraphical sum. It ensues that

$$f^* \leq (a\,\|.\|_{\mathcal{X}}^r)^* + \sigma_{\text{Argmin}f}. \tag{6}$$

First assume that $r = 1$. Since $\|.\|_{\mathcal{X}}^* = \delta_{\mathbb{B}_{\mathcal{X}}}$, where $\mathbb{B}_{\mathcal{X}}$ denotes the closed unit ball of $\mathcal{X}$ centered at 0, we deduce from (6) that

$$f^* - \sigma_{\text{Argmin}f} \leq \delta_{a\mathbb{B}_{\mathcal{X}}}.$$

Since $\lim_{n\to+\infty} 1/\gamma_n = 0$, it is clear that for every $p \in \mathcal{X}$, we have $\delta_{a\mathbb{B}_{\mathcal{X}}}(p/\gamma_n) = 0$ for n large enough. Hence condition (H_f) is automatically satisfied.
Now assume that $r > 1$. Since $(\|.\|_{\mathcal{X}}^r/r)^* = (\|.\|_{\mathcal{X}}^{r^*}/r^*)$, where r^* is the conjugate exponent of r, *i.e.* $r^* = 1/(1 - 1/r)$, we deduce from (6) that

$$f^* - \sigma_{\text{Argmin}f} \leq \frac{(ar)^{1-r^*}}{r^*}\,\|.\|_{\mathcal{X}}^{r^*}.$$

Hence condition (H_f) is satisfied if $\left(\frac{1}{\gamma_n}\right) \in l^{r^*}$. Notice finally that the combination of the conditions $\left(\frac{1}{\gamma_n}\right) \notin l^1$ and $\left(\frac{1}{\gamma_n}\right) \in l^{r^*}$ expresses that (γ_n) tends *moderately slowly* toward infinity.

Let us now state the main result of the paper. Some techniques of the proof are similar to the ones of [6, Theorem 3.1] in a continuous framework.

Theorem 3.1 *Under the hypotheses of Proposition 2.1, let us define the set*

$$S = Argmin\{\|Ax - By\|^2_{\mathcal{Z}} \,:\, (x,y) \in Argminf \times Argming\},$$

assumed to be nonempty. Suppose additionally that the sequence $(\gamma_{n+1} - \gamma_n)$ *is bounded and that assumptions* (H_f)*-*(H_g) *are satisfied. Then any sequence* (x_n, y_n) *generated by* $(\mathcal{A})$ *weakly converges toward a point of* S*.*

Proof. For the sake of readibility, we introduce the maps $\Phi : \mathcal{X}\times\mathcal{Y} \longrightarrow \mathbb{R}\cup\{+\infty\}$ and $\Psi : \mathcal{X}\times\mathcal{Y} \longrightarrow \mathbb{R}_+$ respectively defined by $\Phi(x,y) = f(x)+g(y)$ and $\Psi(x,y) = \frac{1}{2}\|Ax - By\|^2_{\mathcal{Z}}$, for every $(x,y) \in \mathcal{X}\times\mathcal{Y}$. Setting $C = \text{Argmin}f \times \text{Argmin}g$, it is easy to check that

$$\Phi^*\left(\frac{(p,q)}{\gamma_n}\right) - \sigma_C\left(\frac{(p,q)}{\gamma_n}\right) = f^*\left(\frac{p}{\gamma_n}\right) - \sigma_{\text{Argmin}f}\left(\frac{p}{\gamma_n}\right) + g^*\left(\frac{q}{\gamma_n}\right) - \sigma_{\text{Argmin}g}\left(\frac{q}{\gamma_n}\right),$$

and $R(N_C) = R(N_{\text{Argmin}f}) \times R(N_{\text{Argmin}g})$. Hence assumptions (H_f)-(H_g) can be equivalently rewritten as

$$\forall (p,q) \in R(N_C), \quad \left(\Phi^*\left(\frac{(p,q)}{\gamma_n}\right) - \sigma_C\left(\frac{(p,q)}{\gamma_n}\right)\right) \in l^1. \tag{7}$$

The proof of Theorem 3.1 is divided into several claims.

Claim 3.1 *For every* $(x,y) \in S$, $\quad \lim_{n\to+\infty} \alpha\,\|x_n - x\|^2_{\mathcal{X}} + \nu\,\|y_n - y\|^2_{\mathcal{Y}}$ *exists in* $\mathbb{R}$*.*

Proof of Claim 3.1. Fix $(x,y) \in S$ and set $h_n = \frac{1}{\gamma_n}\|By_n - By\|^2_{\mathcal{Z}} + \alpha\|x_n - x\|^2_{\mathcal{X}} + \nu\|y_n - y\|^2_{\mathcal{Y}}$ as in the proof of Proposition 2.1. We can rewrite inequality (4) as follows

$$h_{n+1} - h_n + 2\left\{\Phi(x_{n+1},y_{n+1}) + \frac{1}{\gamma_{n+1}}(\Psi(x_{n+1},y_{n+1}) - \Psi(x,y))\right\} \le 0. \tag{8}$$

Since $(x,y) \in S = \text{Argmin}_C\Psi$, we have $-\nabla\Psi(x,y) \in N_C(x,y)$. Setting $(p,q) = -\nabla\Psi(x,y)$, we deduce that

$$\Psi(x_{n+1},y_{n+1}) - \Psi(x,y) \ge \langle -(p,q), (x_{n+1},y_{n+1}) - (x,y)\rangle_{\mathcal{X}\times\mathcal{Y}}. \tag{9}$$

Moreover the definition of the conjugate Φ^* implies that

$$\Phi(x_{n+1},y_{n+1}) \ge \left\langle \frac{(p,q)}{\gamma_{n+1}}, (x_{n+1},y_{n+1})\right\rangle_{\mathcal{X}\times\mathcal{Y}} - \Phi^*\left(\frac{(p,q)}{\gamma_{n+1}}\right), \tag{10}$$

and, as $(x, y) \in C$ and $(p, q) \in N_C(x, y)$,

$$\sigma_C\left(\frac{(p,q)}{\gamma_{n+1}}\right) = \left\langle \frac{(p,q)}{\gamma_{n+1}}, (x,y) \right\rangle_{\mathcal{X}\times\mathcal{Y}}. \tag{11}$$

By combining (9), (10) and (11), inequality (8) gives

$$h_{n+1} - h_n \leq 2\left\{\Phi^*\left(\frac{(p,q)}{\gamma_{n+1}}\right) - \sigma_C\left(\frac{(p,q)}{\gamma_{n+1}}\right)\right\}.$$

Using formulation (7) of assumptions (H_f)-(H_g), we deduce from Lemma 2.2 that $\lim_{n\to+\infty} h_n$ exists. As a consequence, the sequence (x_n, y_n) is bounded and $\lim_{n\to+\infty} \alpha\|x_n - x\|^2_{\mathcal{X}} + \nu\|y_n - y\|^2_{\mathcal{Y}}$ exists and is equal to $\lim_{n\to+\infty} h_n$. □

Claim 3.2 *The sequence $(\Phi(x_n, y_n))$ is summable.*

Proof of Claim 3.2. Let us set $a_N = 2\sum_{n=0}^{N}\left\{\Phi(x_n, y_n) + \frac{1}{\gamma_n}(\Psi(x_n, y_n) - \Psi(x, y))\right\}$. From inequality (8), we can assert that the sequence $(h_n + a_n)$ is nonincreasing. Moreover, the above calculations and condition (7) allow us to assert that, for all $N \in \mathbb{N}$,

$$a_N \geq -2\sum_{n=0}^{+\infty}\left\{\Phi^*\left(\frac{(p,q)}{\gamma_n}\right) - \sigma_C\left(\frac{(p,q)}{\gamma_n}\right)\right\} > -\infty,$$

thus the sequence $(h_n + a_n)$ is bounded from below, hence convergent. As a consequence, $\lim_{n\to+\infty} a_n$ exists, *i.e.*

$$\lim_{N\to+\infty}\sum_{n=0}^{N}\left\{\Phi(x_n, y_n) + \frac{1}{\gamma_n}(\Psi(x_n, y_n) - \Psi(x, y))\right\} \text{ exists in } \mathbb{R}. \tag{12}$$

Recalling that $\Phi(x_{n+1}, y_{n+1}) \geq 0$ and using the fact that $2\,(p, q) \in N_C(x, y)$, we infer from inequality (8) that

$$h_{n+1} - h_n + \left\{-\Phi^*\left(\frac{2(p,q)}{\gamma_{n+1}}\right) + \sigma_C\left(\frac{2(p,q)}{\gamma_{n+1}}\right)\right\} \leq$$

$$h_{n+1} - h_n + 2\left\{\frac{1}{2}\Phi(x_{n+1}, y_{n+1}) + \frac{1}{\gamma_{n+1}}(\Psi(x_{n+1}, y_{n+1}) - \Psi(x, y))\right\} \leq 0.$$

By arguing as above, we deduce from condition (7) that

$$\lim_{N\to+\infty}\sum_{n=0}^{N}\left\{\frac{1}{2}\Phi(x_n, y_n) + \frac{1}{\gamma_n}(\Psi(x_n, y_n)) - \Psi(x, y))\right\} \text{ exists in } \mathbb{R},$$

which, in view of (12), implies that the sequence $(\Phi(x_n, y_n))$ is summable. □

Claim 3.3 $\lim_{n\to+\infty}\Psi(x_n,y_n)=\min_C\Psi$;

Proof of Claim 3.3. Let us define the sequence (E_n) by

$$E_n=\gamma_n\Phi(x_n,y_n)+\Psi(x_n,y_n).$$

By applying inequality (2) with $x=x_n$ and $y=y_n$, we deduce that

$$\begin{aligned}E_{n+1}-E_n&\le(\gamma_{n+1}-\gamma_n)\Phi(x_n,y_n)-\alpha\gamma_{n+1}\|x_{n+1}-x_n\|^2_{\mathcal{X}}-\nu\gamma_{n+1}\|y_{n+1}-y_n\|^2_{\mathcal{Y}}\\&\le(\gamma_{n+1}-\gamma_n)\Phi(x_n,y_n).\end{aligned}$$

As the sequence $(\gamma_{n+1}-\gamma_n)$ is bounded by assumption and the sequence $(\Phi(x_n,y_n))$ is summable from Claim 3.2, Lemma 2.2 shows that $\lim_{n\to+\infty}E_n$ exists in $\mathbb{R}$. On the other hand, we derive from (12) that

$$\lim_{N\to+\infty}\sum_{n=0}^{N}\frac{1}{\gamma_n}\left(E_n-\Psi(x,y)\right)\ \text{exists in }\mathbb{R}. \tag{13}$$

By assumption, there exists $M\ge 0$ such that $\gamma_{n+1}-\gamma_n\le M$ for all $n\in\mathbb{N}$. Hence we have $\gamma_n\le Mn+\gamma_0$, thus implying that $(\frac{1}{\gamma_n})\notin l^1$. We immediately infer from (13) that $\lim_{n\to+\infty}E_n=\Psi(x,y)$. Since $\Psi(x_n,y_n)\le E_n$, we obtain

$$\limsup_{n\to+\infty}\Psi(x_n,y_n)\le\Psi(x,y). \tag{14}$$

Recall now that the sequence (x_n,y_n) is bounded from Claim 3.1. If a subsequence (x_{n_k},y_{n_k}) weakly converges toward $(\overline{x},\overline{y})$, we can assert by Proposition 2.1 (i) that $(\overline{x},\overline{y})\in C$. We deduce that $\langle-(p,q),(\overline{x},\overline{y})-(x,y)\rangle_{\mathcal{X}\times\mathcal{Y}}\ge 0$, hence every limit point of $(\langle-(p,q),(x_n,y_n)-(x,y)\rangle_{\mathcal{X}\times\mathcal{Y}})$ is nonnegative, that is

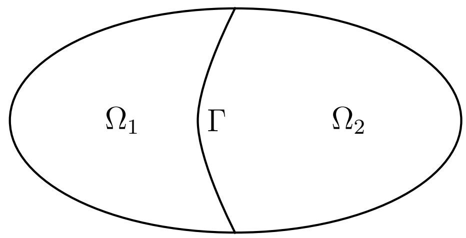
$$\liminf_{n\to+\infty}\langle-(p,q),(x_n,y_n)-(x,y)\rangle_{\mathcal{X}\times\mathcal{Y}}\ge 0.$$

In view of inequality (9), we obtain $\liminf_{n\to+\infty}\Psi(x_n,y_n)\ge\Psi(x,y)$ and we conclude in view of inequality (14). □

To end the proof of Theorem 3.1, we define the norm $N(u,v)=(\alpha\|u\|^2_{\mathcal{X}}+\nu\|v\|^2_{\mathcal{Y}})^{1/2}$, which is equivalent to the canonical norm on $\mathcal{X}\times\mathcal{Y}$. In view of Claim 3.1, $N((x_n,y_n)-(x,y))$ does have a limit for every $(x,y)\in S$, which shows point (a) of Lemma 2.3. Let now (x_{n_k},y_{n_k}) be a subsequence of (x_n,y_n) which weakly converges toward $(\overline{x},\overline{y})$. From Proposition 2.1 (i), we have $(\overline{x},\overline{y})\in C$. Using Claim 3.3 and the closedness of Ψ, we easily infer that $(\overline{x},\overline{y})\in S$, which shows point (b) of Lemma 2.3. Hence we conclude from Opial's lemma that the sequence (x_n,y_n) weakly converges in $\mathcal{X}\times\mathcal{Y}$ toward a point of S.

4 Application to domain decomposition for partial differential equations

Let us consider a bounded domain $\Omega = \Omega_1 \cup \Omega_2 \cup \Gamma$ of $\mathbb{R}^N$ which can be decomposed in two non overlapping Lipschitz subdomains Ω_1 and Ω_2 with a common interface Γ. We assume that $\mathcal{H}^{N-1}(\Gamma) > 0$, where $\mathcal{H}^{N-1}$ is the Hausdorff measure of dimension $N-1$. This situation is illustrated in the next figure.



We consider the following problem

$$(\mathcal{P}) \qquad \min \left\{ \frac{1}{2} \int_\Gamma [w]^2 \right\},$$

where $[w]$ is the jump of w through the interface, $w = \begin{cases} u \text{ on } \Omega_1 \\ v \text{ on } \Omega_2 \end{cases}$ and $u \in H^1(\Omega_1)$, $v \in H^1(\Omega_2)$ are solutions of some boundary value problems, respectively on Ω_1 and Ω_2. This kind of minimization problems often arises in the description of phenomena involving discontinuities on the interfaces between subdomains. To illustrate the results of section 3, we will assume that u, v are respectively weak solutions to the following Neumann boundary value problems

$$\begin{cases} -\Delta u = h \text{ on } \Omega_1 \\ \frac{\partial u}{\partial n} = 0 \text{ on } \partial\Omega_1, \end{cases} \qquad \begin{cases} -\Delta v = h \text{ on } \Omega_2 \\ \frac{\partial v}{\partial n} = 0 \text{ on } \partial\Omega_2, \end{cases}$$

where $h \in L^2(\Omega)$ is a given function. Notice that the corresponding Dirichlet version of these problems was considered in [4] in a slightly different framework, see also [13]. We assume that $\int_{\Omega_1} h = \int_{\Omega_2} h = 0$, which is a necessary and sufficient condition for the existence of a solution. Defining the functions $f : H^1(\Omega_1) \to \mathbb{R}$ and $g : H^1(\Omega_2) \to \mathbb{R}$ by

$$f(u) = \frac{1}{2} \int_{\Omega_1} |\nabla u|^2 - \int_{\Omega_1} hu \quad \text{and} \quad g(v) = \frac{1}{2} \int_{\Omega_2} |\nabla v|^2 - \int_{\Omega_2} hv, \tag{15}$$

it is classical that the solutions u and v of the above Neumann boundary value problems satisfy respectively the following minimization problems

$$(\mathcal{P}_1) \qquad \min\left\{f(u);\quad u\in H^1(\Omega_1)\right\},$$

$$(\mathcal{P}_2) \qquad \min\left\{g(v);\quad v\in H^1(\Omega_2)\right\},$$

see for example [11, 14, 17]. Moreover, denoting by $\widehat{u}$ (resp. $\widehat{v}$) a particular solution of $(\mathcal{P}_1)$ (resp. $(\mathcal{P}_2)$), we have

$$\mathrm{Argmin} f=\{\widehat{u}+c_1;\quad c_1\in\mathbb{R}\},\qquad \mathrm{Argmin} g=\{\widehat{v}+c_2;\quad c_2\in\mathbb{R}\}.$$

Hence in our framework, problem $(\mathcal{P})$ amounts to minimizing over $\mathbb{R}^2$ the map

$$(c_1,c_2)\mapsto \frac{1}{2}\int_\Gamma \left(\widehat{u}_{|\Gamma}-\widehat{v}_{|\Gamma}+c_1-c_2\right)^2.$$

It is immediate to verify that the minimum is reached when

$$c_2-c_1=\frac{1}{\mathcal{H}^{N-1}(\Gamma)}\int_\Gamma \left(\widehat{u}_{|\Gamma}-\widehat{v}_{|\Gamma}\right).$$

Without loss of generality, we can assume that $\int_\Gamma \widehat{u}_{|\Gamma}=\int_\Gamma \widehat{v}_{|\Gamma}=0$. Then the above relation gives $c_1=c_2$, hence the set of solutions of $(\mathcal{P})$ is of the form $\{(\widehat{u}+c,\widehat{v}+c);\quad c\in\mathbb{R}\}$.

Let us now show how algorithm $(\mathcal{A})$ can be applied so as to solve problem $(\mathcal{P})$. The set $\mathcal{X}=H^1(\Omega_1)$ is equipped with the scalar product $\langle u_1,u_2\rangle_{\mathcal{X}}=\int_{\Omega_1}(\nabla u_1.\nabla u_2+u_1u_2)$ and the corresponding norm. The same holds for $\mathcal{Y}=H^1(\Omega_2)$ by replacing Ω_1 with Ω_2. The set $\mathcal{Z}=L^2(\Gamma)$ is equipped with the scalar product $\langle z_1,z_2\rangle_{\mathcal{Z}}=\int_\Gamma z_1z_2$ and the associate norm. The operators $A:\mathcal{X}\longrightarrow\mathcal{Z}$ and $B:\mathcal{Y}\longrightarrow\mathcal{Z}$ are respectively the trace operators on Γ, which are well-defined by the Lipschitz character of the boundaries of Ω_1 and Ω_2 (see [8, Theorem II.46] or [15, Theorem 2]). Algorithm $(\mathcal{A})$ runs as follows

$$\begin{cases} u_{n+1}=\mathrm{Argmin}\left\{f(u)+\dfrac{1}{2\gamma_{n+1}}\|Au-Bv_n\|^2_{\mathcal{Z}}+\dfrac{\alpha}{2}\|u-u_n\|^2_{\mathcal{X}};\quad u\in\mathcal{X}\right\}\\ v_{n+1}=\mathrm{Argmin}\left\{g(v)+\dfrac{1}{2\gamma_{n+1}}\|Au_{n+1}-Bv\|^2_{\mathcal{Z}}+\dfrac{\nu}{2}\|v-v_n\|^2_{\mathcal{Y}};\quad v\in\mathcal{Y}\right\}, \end{cases}$$

where α and ν are fixed positive parameters. An elementary directional derivative computation shows that the weak variational formulation of algorithm $(\mathcal{A})$ is given by

$$\forall u\in\mathcal{X},\qquad \int_{\Omega_1}\nabla u_{n+1}.\nabla u+\frac{1}{\gamma_{n+1}}\int_\Gamma (Au_{n+1}-Bv_n)Au$$
$$+\alpha\int_{\Omega_1}(\nabla u_{n+1}-\nabla u_n).\nabla u+\alpha\int_{\Omega_1}(u_{n+1}-u_n)u=\int_{\Omega_1}hu,$$

$$\forall v \in \mathcal{Y}, \qquad \int_{\Omega_2} \nabla v_{n+1}.\nabla v + \frac{1}{\gamma_{n+1}} \int_\Gamma (Bv_{n+1} - Au_{n+1})Bv$$
$$+\nu \int_{\Omega_2} (\nabla v_{n+1} - \nabla v_n).\nabla v + \nu \int_{\Omega_2} (v_{n+1} - v_n)v = \int_{\Omega_2} hv.$$

These are the variational weak formulations of the following mixed Dirichlet-Neumann boundary value problems respectively on Ω_1

$$\begin{cases} -(1+\alpha)\Delta u_{n+1} + \alpha u_{n+1} = h - \alpha\Delta u_n + \alpha u_n & \text{on } \Omega_1 \\ (1+\alpha)\frac{\partial u_{n+1}}{\partial n} = \alpha\frac{\partial u_n}{\partial n} & \text{on } \partial\Omega_1 \cap \partial\Omega \\ (1+\alpha)\frac{\partial u_{n+1}}{\partial n} + \frac{1}{\gamma_{n+1}} u_{n+1} = \alpha\frac{\partial u_n}{\partial n} + \frac{1}{\gamma_{n+1}} v_n & \text{on } \Gamma, \end{cases}$$

and Ω_2

$$\begin{cases} -(1+\nu)\Delta v_{n+1} + \nu v_{n+1} = h - \nu\Delta v_n + \nu v_n & \text{on } \Omega_2 \\ (1+\nu)\frac{\partial v_{n+1}}{\partial n} = \nu\frac{\partial v_n}{\partial n} & \text{on } \partial\Omega_2 \cap \partial\Omega \\ (1+\nu)\frac{\partial v_{n+1}}{\partial n} + \frac{1}{\gamma_{n+1}} v_{n+1} = \nu\frac{\partial v_n}{\partial n} + \frac{1}{\gamma_{n+1}} u_{n+1} & \text{on } \Gamma. \end{cases}$$

To apply Theorem 3.1, we have to check that assumptions (H_f) and (H_g) are satisfied. In view of the symmetry of f and g, let us focus on (H_f). Since $\text{Argmin} f = \{\widehat{u} + c_1;\quad c_1 \in \mathbb{R}\}$ is an affine space directed by the vector space of constant functions, it is clear that for every $u \in \text{Argmin} f$

$$N_{\text{Argmin}_f}(u) = \{p \in \mathcal{X}, \langle p, 1\rangle_{\mathcal{X}} = 0\} = \left\{p \in \mathcal{X}, \int_{\Omega_1} p = 0\right\}.$$

In the sequel, we denote by V this hyperplane of $\mathcal{X}$. For every $p \in V$, we have

$$\sigma_{\text{Argmin}_f}(p) = \sup_{c_1 \in \mathbb{R}} \langle \widehat{u} + c_1, p\rangle_{\mathcal{X}} = \langle \widehat{u}, p\rangle_{\mathcal{X}}. \tag{16}$$

From the definition of f and since $\widehat{u} \in \text{Argmin} f$, we have for every $u \in \mathcal{X}$

$$f(u) - \min f = \frac{1}{2}\int_{\Omega_1} |\nabla u|^2 - \int_{\Omega_1} hu - \frac{1}{2}\int_{\Omega_1} |\nabla \widehat{u}|^2 + \int_{\Omega_1} h\widehat{u}. \tag{17}$$

Recalling that the weak variational formulation of $(\mathcal{P}_1)$ gives

$$\forall u \in \mathcal{X}, \quad \int_{\Omega_1} \nabla u.\nabla \widehat{u} = \int_{\Omega_1} hu,$$

we derive from (17) that

$$f(u) - \min f = \frac{1}{2}\int_{\Omega_1} |\nabla u - \nabla \widehat{u}|^2. \tag{18}$$

In view of (16) and (18), we find for every $p \in V$

$$(f - \min f)^*(p) - \sigma_{\text{Argmin}_f}(p) = \sup_{u \in \mathcal{X}} \left\{ \langle p, u - \widehat{u} \rangle_{\mathcal{X}} - \frac{1}{2} \int_{\Omega_1} |\nabla u - \nabla \widehat{u}|^2 \right\}$$

$$= \sup_{u \in \mathcal{X}} \left\{ \langle p, u \rangle_{\mathcal{X}} - \frac{1}{2} \int_{\Omega_1} |\nabla u|^2 \right\}. \tag{19}$$

Consider the following minimization problem in V

$$(\mathcal{P}^*) \qquad \inf \left\{ \frac{1}{2} \int_{\Omega_1} |\nabla u|^2 - \langle p, u \rangle_{\mathcal{X}};\ u \in V \right\}.$$

Let us introduce the bilinear form $a : V \times V \to \mathbb{R}$ defined by $a(u,v) = \int_{\Omega_1} \nabla u . \nabla v$. From the Poincaré-Wirtinger inequality, there exists some constant $m > 0$ such that

$$\forall u \in V, \quad \int_{\Omega_1} u^2 \leq m^2 \int_{\Omega_1} |\nabla u|^2.$$

We immediately deduce that

$$\forall u \in V, \quad \int_{\Omega_1} |\nabla u|^2 \geq \frac{1}{1+m^2} \|u\|_{\mathcal{X}}^2. \tag{20}$$

Hence the bilinear form a is coercive on $V \times V$ and we infer that problem $(\mathcal{P}^*)$ has a unique solution, that we denote by u^*. Equivalently u^* is the solution of the variational problem

$$\forall u \in V, \quad \int_{\Omega_1} \nabla u^* . \nabla u = \langle p, u \rangle_{\mathcal{X}} = \int_{\Omega_1} \nabla p . \nabla u + \int_{\Omega_1} pu. \tag{21}$$

It is immediate to check that u^* is a solution of the corresponding problem in $\mathcal{X}$. Hence the supremum in expression (19) is attained at u^* and we derive in view of (21) that

$$(f - \min f)^*(p) - \sigma_{\text{Argmin}_f}(p) = \frac{1}{2} \int_{\Omega_1} |\nabla u^*|^2 = \frac{1}{2} \langle p, u^* \rangle_{\mathcal{X}}. \tag{22}$$

From inequality (20) applied with $u = u^*$, we infer that

$$\frac{1}{1+m^2} \|u^*\|_{\mathcal{X}}^2 \leq \int_{\Omega_1} |\nabla u^*|^2 = \langle p, u^* \rangle_{\mathcal{X}} \leq \|p\|_{\mathcal{X}} \|u^*\|_{\mathcal{X}}.$$

This implies that $\|u^*\|_{\mathcal{X}} \leq (1+m^2)\|p\|_{\mathcal{X}}$. Hence we derive from (22) that for every $p \in V$

$$(f - \min f)^*(p) - \sigma_{\text{Argmin}_f}(p) \leq \frac{1+m^2}{2} \|p\|_{\mathcal{X}}^2.$$

Since the right hand-side of the above inequality is quadratic, assumption (H_f) is satisfied as soon as $\left(\frac{1}{\gamma_n}\right) \in l^2$ and the same holds for (H_g). We conclude from Theorem 3.1 and the above analysis that if $(\gamma_{n+1} - \gamma_n)$ is bounded and if $\left(\frac{1}{\gamma_n}\right) \in l^2$, then any sequence (u_n, v_n) generated by $(\mathcal{A})$ weakly converges in $H^1(\Omega_1) \times H^1(\Omega_2)$ to a minimum point $(\widehat{u} + c, \widehat{v} + c)$, $(c \in \mathbb{R})$ of problem $(\mathcal{P})$. Without loss of generality, we can assume that $c = 0$. Since Ω_1 and Ω_2 are Lipschitz domains, the injections $H^1(\Omega_1) \hookrightarrow L^2(\Omega_1)$ and $H^1(\Omega_2) \hookrightarrow L^2(\Omega_2)$ are compact by the Rellich-Kondrachov Theorem (see [2, Theorem 6.2] or [8, Theorem II.55]). It ensues that the sequence (u_n, v_n) converges to $(\widehat{u}, \widehat{v})$ strongly in $L^2(\Omega_1) \times L^2(\Omega_2)$. Moreover, from Proposition 2.1 (i), we have $\lim_{n\to+\infty} f(u_n) = f(\widehat{u})$ and $\lim_{n\to+\infty} g(v_n) = g(\widehat{v})$, hence $\lim_{n\to+\infty} \int_{\Omega_1} |\nabla u_n|^2 = \int_{\Omega_1} |\nabla \widehat{u}|^2$ and $\lim_{n\to+\infty} \int_{\Omega_2} |\nabla v_n|^2 = \int_{\Omega_2} |\nabla \widehat{v}|^2$. As a consequence, we have $\lim_{n\to+\infty} \|(u_n, v_n)\|_{H^1(\Omega_1)\times H^1(\Omega_2)} = \|(\widehat{u}, \widehat{v})\|_{H^1(\Omega_1)\times H^1(\Omega_2)}$. Since (u_n, v_n) weakly converges in $H^1(\Omega_1) \times H^1(\Omega_2)$ toward $(\widehat{u}, \widehat{v})$, the convergence is strong in $H^1(\Omega_1) \times H^1(\Omega_2)$. We can state the following theorem.

Theorem 4.1 *Let Ω be a bounded domain of $\mathbb{R}^N$ which can be decomposed in two nonoverlapping Lipschitz subdomains Ω_1 and Ω_2 with a common interface Γ. We assume that $\mathcal{H}^{N-1}(\Gamma) > 0$. Let $h \in L^2(\Omega)$ be such that $\int_{\Omega_1} h = \int_{\Omega_2} h = 0$ and define the functions $f : H^1(\Omega_1) \to \mathbb{R}$ and $g : H^1(\Omega_2) \to \mathbb{R}$ by formulas (15). Let (γ_n) be a positive nondecreasing sequence such that $\left(\frac{1}{\gamma_n}\right) \in l^2$ and assume that the sequence $(\gamma_{n+1} - \gamma_n)$ is bounded. Then any sequence (u_n, v_n) generated by algorithm $(\mathcal{A})$ strongly converges in $H^1(\Omega_1) \times H^1(\Omega_2)$ to a minimum point $(\widehat{u}, \widehat{v})$ of problem $(\mathcal{P})$.*

Remark 4.1 *The above analysis gives an exact expression of the conjugate f^*. However, it is possible to verify directly assumption (H_f) without resorting to an exact computation of f^*. The method consists in checking that inequality (5) of Example 3.1 is satisfied with $r = 2$. From the Poincaré-Wirtinger inequality, we have*

$$\int_{\Omega_1} \left| u - \widehat{u} - \frac{1}{|\Omega_1|} \int_{\Omega_1} (u - \widehat{u}) \right|^2 \leq m^2 \int_{\Omega_1} |\nabla u - \nabla \widehat{u}|^2.$$

Since $\widehat{u} + \frac{1}{|\Omega_1|} \int_{\Omega_1} (u - \widehat{u}) \in Argminf$, we deduce that

$$\begin{aligned} \int_{\Omega_1} |\nabla u - \nabla \widehat{u}|^2 &\geq \frac{1}{1+m^2} \left\| u - \widehat{u} - \frac{1}{|\Omega_1|} \int_{\Omega_1} (u - \widehat{u}) \right\|_{\mathcal{X}}^2 \\ &\geq \frac{1}{1+m^2}\, d_{\mathcal{X}}^2(u, Argminf). \end{aligned} \tag{23}$$

In view of (18) and (23), we find, for every $u \in \mathcal{X}$,

$$f(u) - \min f \geq \frac{1}{2(1+m^2)}\, d_{\mathcal{X}}^2(u, Argminf).$$

From Example 3.1, we conclude that the assumption (H_f) is satisfied if $(\frac{1}{\gamma_n}) \in l^2$.

5 Further convergence results in the finite dimensional setting

From now to the end of this section, $\mathcal{X}$ and $\mathcal{Y}$ are finite-dimensional Hilbert spaces. The approach that we now develop relies on topological ingredients that can already be found in [3, 9, 12]. The first result shows that the iterates (x_n, y_n) of algorithm $(\mathcal{A})$ approach the optimal set S as $n \to +\infty$.

Theorem 5.1 *Under the hypotheses of Proposition 2.1, assume that the spaces $\mathcal{X}$ and $\mathcal{Y}$ are finite-dimensional. Let us define the set*

$$S = Argmin\{\|Ax - By\|^2_{\mathcal{Z}} \ : \ (x, y) \in Argminf \times Argming\},$$

assumed to be nonempty and bounded. Suppose that the sequence (γ_n) satisfies $\left(\frac{1}{\gamma_n}\right) \in l^2 \backslash l^1$. Let us consider algorithm $(\mathcal{A})$ with $\alpha = \nu$. Then any sequence (x_n, y_n) generated by $(\mathcal{A})$ satisfies $\lim_{n\to+\infty} d_{\mathcal{X}\times\mathcal{Y}}((x_n, y_n), S) = 0$.

Proof. Let us define the maps $\Phi : \mathcal{X} \times \mathcal{Y} \longrightarrow \mathbb{R} \cup \{+\infty\}$ and $\Psi : \mathcal{X} \times \mathcal{Y} \longrightarrow \mathbb{R}_+$ respectively by $\Phi(x, y) = f(x) + g(y)$ and $\Psi(x, y) = \frac{1}{2}\|Ax - By\|^2_{\mathcal{Z}}$, along with the set $C = \text{Argmin}f \times \text{Argmin}g$. Consider the sequence (h_n) defined by $h_n = \frac{1}{2} d^2_{\mathcal{X}\times\mathcal{Y}}((x_n, y_n), S)$ for every $n \in \mathbb{N}$. Denoting by $P_S(x_n, y_n) = (P^x_n, P^y_n)$ the projection of $(x_n, y_n) \in \mathcal{X} \times \mathcal{Y}$ onto the convex set S, we have

$$\begin{aligned} h_n =& \frac{1}{2}\|(x_n, y_n) - P_S(x_n, y_n)\|^2_{\mathcal{X}\times\mathcal{Y}} \\ =& \frac{1}{2}\|(x_n, y_n) - P_S(x_n, y_n) - ((x_{n+1}, y_{n+1}) - P_S(x_{n+1}, y_{n+1}))\|^2_{\mathcal{X}\times\mathcal{Y}} + h_{n+1} \\ &+ \langle (x_n, y_n) - P_S(x_n, y_n) - ((x_{n+1}, y_{n+1}) - P_S(x_{n+1}, y_{n+1})), (x_{n+1}, y_{n+1}) - P_S(x_{n+1}, y_{n+1})\rangle_{\mathcal{X}\times\mathcal{Y}}. \end{aligned}$$

Since $P_S(x_n, y_n) \in S$, we have

$$\langle P_S(x_n, y_n) - P_S(x_{n+1}, y_{n+1}), (x_{n+1}, y_{n+1}) - P_S(x_{n+1}, y_{n+1})\rangle_{\mathcal{X}\times\mathcal{Y}} \leq \ 0,$$

hence

$$h_{n+1} - h_n \leq \langle x_{n+1} - x_n, x_{n+1} - P^x_{n+1}\rangle_{\mathcal{X}} + \langle y_{n+1} - y_n, y_{n+1} - P^y_{n+1}\rangle_{\mathcal{Y}}. \tag{24}$$

Using subdifferential inequalities (1) with $\alpha = \nu$, we have

$$\begin{cases} \langle x_{n+1} - x_n, x_{n+1} - P^x_{n+1}\rangle_{\mathcal{X}} \leq -\dfrac{1}{\alpha} f(x_{n+1}) - \dfrac{1}{\alpha\gamma_{n+1}}\langle Ax_{n+1} - By_n, Ax_{n+1} - AP^x_{n+1}\rangle_{\mathcal{Z}} \\ \langle y_{n+1} - y_n, y_{n+1} - P^y_{n+1}\rangle_{\mathcal{Y}} \leq -\dfrac{1}{\alpha} g(y_{n+1}) + \dfrac{1}{\alpha\gamma_{n+1}}\langle Ax_{n+1} - By_{n+1}, By_{n+1} - BP^y_{n+1}\rangle_{\mathcal{Z}}. \end{cases} \tag{25}$$

On the other hand, since $\nabla\Psi(x, y) = (A^*(Ax - By), -B^*(Ax - By))$ we have

$$\Psi(P^x_{n+1},P^y_{n+1})-\Psi(x_{n+1},y_{n+1})\geq\langle Ax_{n+1}-By_{n+1},AP^x_{n+1}-Ax_{n+1}\rangle_{\mathcal{Z}}$$
$$-\langle Ax_{n+1}-By_{n+1},BP^y_{n+1}-By_{n+1}\rangle_{\mathcal{Z}}. \quad (26)$$

Then inequalities (24), (25) and (26) give

$$h_{n+1}-h_n+\frac{1}{\alpha}\Phi(x_{n+1},y_{n+1})+\frac{1}{\alpha\gamma_{n+1}}(\Psi(x_{n+1},y_{n+1})-\min_C\Psi)$$
$$+\frac{1}{\alpha\gamma_{n+1}}\langle By_{n+1}-By_n,Ax_{n+1}-AP^x_{n+1}\rangle_{\mathcal{Z}}\leq 0,$$

hence

$$h_{n+1}-h_n+\frac{1}{\alpha\gamma_{n+1}}(\Psi(x_{n+1},y_{n+1})-\min_C\Psi)\leq\frac{1}{2\alpha}\|By_{n+1}-By_n\|^2_{\mathcal{Z}}+\frac{1}{2\alpha\gamma^2_{n+1}}\|Ax_{n+1}-AP^x_{n+1}\|^2_{\mathcal{Z}}.$$

Let us set $\mu_n=\frac{1}{2\alpha}\|By_{n+1}-By_n\|^2_{\mathcal{Z}}$. As the linear operator B is continuous and since $(\|y_{n+1}-y_n\|^2_{\mathcal{Y}})$ is summable, (μ_n) is also summable. In the same way, there exists a constant $M\geq 0$ such that $\frac{1}{2}\|Ax_{n+1}-AP^x_{n+1}\|^2_{\mathcal{Z}}\leq\frac{M}{2}\|x_{n+1}-P^x_{n+1}\|^2_{\mathcal{X}}\leq Mh_{n+1}$. Setting $\rho_n=1-\frac{M}{\alpha\gamma_n^2}$, there exists $n_0\in\mathbb{N}$ such that $\rho_n>0$ for all $n\geq n_0$. Finally we find

$$\rho_{n+1}h_{n+1}-h_n+\frac{1}{\alpha\gamma_{n+1}}(\Psi(x_{n+1},y_{n+1})-\min_C\Psi)\leq\mu_n.$$

Let us note $\rho'_n=\prod_{i=n_0}^n\rho_i$ and $h'_n=\rho'_n(h_n+\sum_{i=n}^{+\infty}\mu_i)$, then, proceeding as in the proof of [3, Theorem 3], we find, for all $n\geq n_0$,

$$h'_{n+1}+\rho'_n\frac{1}{\alpha\gamma_{n+1}}(\Psi(x_{n+1},y_{n+1})-\min_C\Psi)\leq h'_n. \quad (27)$$

We now follow the same arguments as those used by the first author in [12, Theorem 3.1]. We distinguish two cases:

(a)There exists $n_1\geq n_0$ such that for all $n\geq n_1$, $\Psi(x_{n+1},y_{n+1})>\min_C\Psi$.
(b)For all $n_1\geq n_0$, there exists $n\geq n_1$ such that $\Psi(x_{n+1},y_{n+1})\leq\min_C\Psi$.

Case (a). We assume that there exists $n_1\geq n_0$ such that for all $n\geq n_1$, $\Psi(x_{n+1},y_{n+1})>\min_C\Psi$. Then $(h'_n)_{n\geq n_1}$ is nonincreasing, hence convergent. Remark that, since $(\frac{1}{\gamma_n^2})$ is summable, $\lim_{n\to+\infty}\rho'_n=\rho\in]0,1[$, therefore (h_n) is also convergent. We must prove that $\lim_{n\to+\infty}h_n=0$. Using inequality (27), we can assert that $(\frac{1}{\gamma_n}(\Psi(x_n,y_n)-\min_C\Psi))$ is summable. Moreover, since $(\frac{1}{\gamma_n})$ is not summable, we have $\liminf_{n\to+\infty}\Psi(x_n,y_n)=\min_C\Psi$. Consider a subsequence of (x_n,y_n), still denoted by (x_n,y_n), such that $\lim_{n\to+\infty}\Psi(x_n,y_n)=\min_C\Psi$. As the sequence (h_n) converges and since the set S is bounded, we infer that the sequence (x_n,y_n) is bounded. Since (x_n,y_n) lies in the finite-dimensional space $\mathcal{X}\times\mathcal{Y}$, we can extract a subsequence (x_{n_k},y_{n_k}) of (x_n,y_n) which converges toward $(\overline{x},\overline{y})\in$

$\mathcal{X} \times \mathcal{Y}$. In view of Proposition 2.1 (i), we have $(\overline{x}, \overline{y}) \in C$. Moreover the map Ψ is continuous, hence $\lim_{k\to+\infty} \Psi(x_{n_k}, y_{n_k}) = \Psi(\overline{x}, \overline{y}) = \min_C \Psi$. Finally $(\overline{x}, \overline{y}) \in S$, and $\lim_{k\to+\infty} h_{n_k} = \lim_{k\to+\infty} \frac{1}{2} d^2_{\mathcal{X}\times\mathcal{Y}}((x_{n_k}, y_{n_k}), S) = \frac{1}{2} d^2_{\mathcal{X}\times\mathcal{Y}}((\overline{x}, \overline{y}), S) = 0$. Recalling that the sequence (h_n) is convergent, we conclude that $\lim_{n\to+\infty} h_n = 0$. *Case* (b). We assume that, for all $n_1 \geq n_0$, there exists $n \geq n_1$ such that $\Psi(x_{n+1}, y_{n+1}) \leq \min_C \Psi$. Let us define

$$\tau_N = \max\{n \in \mathbb{N}, n \leq N \text{ and } \Psi(x_n, y_n) \leq \min_C \Psi\}.$$

The integer τ_N is well-defined for N large enough and $\lim_{N\to+\infty} \tau_N = +\infty$. If $\tau_N \leq N-1$, we have $h'_{n+1} \leq h'_n$ for all $n \in \{\tau_N, N-1\}$, therefore

$$h'_N \leq h'_{\tau_N}. \tag{28}$$

If $\tau_N = N$, inequality (28) is still true. Because of Proposition 2.1 (i), we have $\lim_{n\to+\infty} \Phi(x_n, y_n) = 0$, hence there exists $M_0 > 0$ such that, for every $n \in \mathbb{N}$, $\Phi(x_n, y_n) \leq M_0$. From the definition of τ_N, we have $\Psi(x_{\tau_N}, y_{\tau_N}) \leq \min_C \Psi$. Hence we have

$$(x_{\tau_N}, y_{\tau_N}) \in [\Phi \leq M_0] \cap [\Psi \leq \min_C \Psi],$$

as soon as (x_{τ_N}, y_{τ_N}) is defined. It is proved in [12, Lemma 3.3] that the boundedness of $S = \operatorname{Argmin}_C \Psi$ implies the boundedness of the set $[\Phi \leq M_0] \cap [\Psi \leq \min_C \Psi]$. It ensues that the sequence (h_{τ_N}) is bounded and the same holds true for the sequence (h'_{τ_N}). Let us show that $\lim_{N\to+\infty} h'_{\tau_N} = 0$. For that purpose, we prove that 0 is the unique limit point of the bounded sequence (h'_{τ_N}). Considering a converging subsequence $(h'_{\tau_{N_k}})$ of (h'_{τ_N}), we can extract a subsequence of $(x_{\tau_{N_k}}, y_{\tau_{N_k}})$, still denoted by $(x_{\tau_{N_k}}, y_{\tau_{N_k}})$, which converges toward $(\overline{x}, \overline{y}) \in \mathcal{X} \times \mathcal{Y}$. Using Proposition 2.1 (i) and since $[\Psi \leq \min_C \Psi]$ is closed as a sublevel set of the continuous function Ψ, we infer that $(\overline{x}, \overline{y}) \in S$. Therefore $\lim_{k\to+\infty} h_{\tau_{N_k}} = \lim_{k\to+\infty} \frac{1}{2} d^2_{\mathcal{X}\times\mathcal{Y}}((x_{\tau_{N_k}}, y_{\tau_{N_k}}), S) = \frac{1}{2} d^2_{\mathcal{X}\times\mathcal{Y}}((\overline{x}, \overline{y}), S) = 0$ and hence $\lim_{k\to+\infty} h'_{\tau_{N_k}} = 0$. We immediately deduce that the whole sequence (h'_{τ_N}) converges toward 0. Then inequality (28) implies that $\lim_{N\to+\infty} h'_N = 0$, which allows to conclude that $\lim_{N\to+\infty} h_N = 0$.

Under the additional assumptions (H_f)-(H_g) introduced in section 3, one can obtain the convergence of the whole sequence (x_n, y_n) toward a point $(\overline{x}, \overline{y}) \in S$.

Proposition 5.1 *Under the hypotheses of Theorem 5.1, assume moreover that conditions (H_f)-(H_g) hold. Then, any sequence (x_n, y_n) generated by algorithm $(\mathcal{A})$ converges to a point $(\overline{x}, \overline{y}) \in S$.*

Proof. Let $(x, y) \in S$ and define the sequence (g_n) by $g_n = \frac{1}{2} \|(x_n, y_n) - (x, y)\|^2_{\mathcal{X}\times\mathcal{Y}}$ for every $n \in \mathbb{N}$. By arguing as in the proof of Theorem 5.1, we obtain

$$g_{n+1} - g_n + \frac{1}{\alpha}\left[\Phi(x_{n+1}, y_{n+1}) + \frac{1}{\gamma_{n+1}}(\Psi(x_{n+1}, y_{n+1}) - \min_C \Psi)\right] \\ + \frac{1}{\alpha\gamma_{n+1}}\langle By_{n+1} - By_n, Ax_{n+1} - Ax\rangle_{\mathcal{Z}} \leq 0.$$

Setting $(p, q) = -\nabla\Psi(x, y)$, we obtain as in the proof of Theorem 3.1 that

$$\Phi(x_{n+1}, y_{n+1}) + \frac{1}{\gamma_{n+1}}(\Psi(x_{n+1}, y_{n+1}) - \min_C \Psi) \geq -\Phi^*\left(\frac{(p, q)}{\gamma_{n+1}}\right) + \sigma_C\left(\frac{(p, q)}{\gamma_{n+1}}\right).$$

We infer from the above inequality that

$$g_{n+1}-g_n \leq \frac{1}{\alpha}\left[\Phi^*\left(\frac{(p, q)}{\gamma_{n+1}}\right) - \sigma_C\left(\frac{(p, q)}{\gamma_{n+1}}\right)\right]+\frac{1}{2\alpha}\|By_{n+1}-By_n\|^2_{\mathcal{Z}}+\frac{1}{2\alpha\gamma^2_{n+1}}\|Ax_{n+1}-Ax\|^2_{\mathcal{Z}}.$$

Let us set $\lambda_n = \frac{1}{\alpha}\left[\Phi^*\left(\frac{(p,q)}{\gamma_{n+1}}\right) - \sigma_C\left(\frac{(p,q)}{\gamma_{n+1}}\right)\right]$ and $\mu_n = \frac{1}{2\alpha}\|By_{n+1} - By_n\|^2_{\mathcal{Z}}$. The sequence (λ_n) is summable from assumptions (H_f)-(H_g), see formula (7). Since the linear operator B is continuous and since $(\|y_{n+1} - y_n\|^2_{\mathcal{Y}})$ is summable, the sequence (μ_n) is also summable. On the other hand, there exists $M \geq 0$ such that $\frac{1}{2}\|Ax_{n+1} - Ax\|^2_{\mathcal{Z}} \leq \frac{M}{2}\|x_{n+1} - x\|^2_{\mathcal{X}} \leq Mg_{n+1}$. Let us note $\rho_n = 1 - \frac{M}{\alpha\gamma_n^2}$. There exists $n_0 \in \mathbb{N}$ such that $\rho_n > 0$ for all $n \geq n_0$ and we find

$$\rho_{n+1}g_{n+1} - g_n \leq \lambda_n + \mu_n.$$

By setting $\rho'_n = \prod_{i=n_0}^n \rho_i$ and $g'_n = \rho'_n\, g_n$, we infer that for all $n \geq n_0$,

$$g'_{n+1} - g'_n \leq \rho'_n(\lambda_n + \mu_n).$$

Recall that, because $(\frac{1}{\gamma_n^2})$ is summable, $\lim_{n\to+\infty} \rho'_n = \rho \in]0, 1[$. Since $(\lambda_n + \mu_n)$ is summable, we deduce that the right member of the previous inequality is also summable. Then we can apply Lemma 2.2 to assert that $\lim_{n\to+\infty} g'_n$ exists, hence $\lim_{n\to+\infty} g_n$ also exists. We have proved that

$$\lim_{n\to+\infty} \|(x_n, y_n) - (x, y)\|_{\mathcal{X}\times\mathcal{Y}} \text{ exists for any } (x, y) \in S. \tag{29}$$

Hence the sequence (x_n, y_n) is bounded, therefore we can extract a subsequence (x_{n_k}, y_{n_k}) which converges toward $(\overline{x}, \overline{y})$. From Theorem 5.1, we have $(\overline{x}, \overline{y}) \in S$. Taking $(x, y) = (\overline{x}, \overline{y})$ in (29), we deduce that $\lim_{n\to+\infty} \|(x_n, y_n) - (\overline{x}, \overline{y})\|_{\mathcal{X}\times\mathcal{Y}}$ exists and finally $\lim_{n\to+\infty} \|(x_n, y_n) - (\overline{x}, \overline{y})\|_{\mathcal{X}\times\mathcal{Y}} = 0$.

Remark 5.1 *Theorem 5.1 and Proposition 5.1 rely on techniques which differ from the ones of section 3. Notice that they do not assume that the sequence* $(\gamma_{n+1} - \gamma_n)$ *is bounded, as it is the case in Theorem 3.1. The main drawback is that the involved spaces* $\mathcal{X}$ *and* $\mathcal{Y}$ *are finite-dimensional, which precludes the potential applications to PDE's.*

Concluding comments. In this paper, we focused our attention on the case $\gamma_n \to +\infty$ as $n \to +\infty$. Another situation of interest corresponds to a decreasing sequence (γ_n) tending to 0 as $n \to +\infty$. Under suitable conditions on the decay rate of (γ_n), the associated algorithm minimizes the function $(x, y) \mapsto f(x)+g(y)$ over the space $\mathcal{V} = \{(x, y) \in \mathcal{X} \times \mathcal{Y},\ Ax = By\}$. This situation is analyzed in the companion paper [5].

Acknowledgments

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. Acker, M.A. Prestel, Convergence d'un schéma de minimisation alternée, *Annales de la Faculté des Sciences de Toulouse V, série Mathématiques*, 2 (1980), pp. 1-9.

[2] R. Adams, Sobolev spaces, Academic Press, New-York, 1975.

[3] F. Alvarez, R. Cominetti, Primal and dual convergence of a proximal point exponential penalty method for linear programming, *Math. Programming*, 93 (2002), pp. 87-96.

[4] H. Attouch, J. Bolte, P. Redont, A. Soubeyran, Alternating proximal algorithms for weakly coupled convex minimization problems. Applications to dynamical games and PDE's, *Journal of Convex Analysis*, 15 (2008), No 3, pp. 485-506.

[5] H. Attouch, A. Cabot, P. Frankel, J. Peypouquet, Alternating proximal algorithms for constrained variational inequalities. Application to domain decomposition for PDE's, submitted to *Nonlinear Analysis: Theory, Methods & Applications*.

[6] H. Attouch, M.-O. Czarnecki, Asymptotic behavior of coupled dynamical systems with multiscale aspects, *Journal of Differential Equations*, 248 (2010), pp. 1315-1344.

[7] H. Attouch, P. Redont, A. Soubeyran, A new class of alternating proximal minimization algorithms with costs-to-move, *SIAM Journal on Optimization*, 18 (2007), pp. 1061-1081.

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

[9] J.-B. Baillon, R. Cominetti, A convergence result for non-autonomous subgradient evolution equations and its application to the steepest descent exponential penalty trajectory in linear programming, *J. Funct. Anal.*, 187 (2001), pp. 263-273.

[10] H.H. Bauschke, P.-L. Combettes, S. Reich, The asymptotic behavior of the composition of two resolvents, *Nonlinear analysis*, 60 (2005), pp. 283-301.

[11] H. Brezis, Analyse fonctionnelle : théorie et applications, Dunod, Paris, 1999.

[12] A. Cabot, Proximal point algorithm controlled by a slowly vanishing term: Applications to hierarchical minimization, *SIAM Journal on Optimization*, 15 (2005), No 2, pp. 555-572.

[13] P. Frankel, Alternating proximal algorithm with costs-to-move, dual description and application to PDE's, accepted in *Discrete and Continuous Dynamical Systems- Series S.*

[14] J.-L. Lions, E. Magenes, Non-homogeneous boundary value problems and applications, Die Grundlehren der mathematischen Wissenschaften, Springer-Verlag, New York-Heidelberg, 1972.

[15] J. Marschall, The trace of Sobolev-Slobodeckij spaces on Lipschitz domains, *Manuscripta Math.*, **58** (1987), pp. 47-65.

[16] Z. Opial, Weak convergence of the sequence of successive approximations for nonexpansive mappings, *Bull. Amer. Math. Soc.*, 73 (1967), pp. 591-597.

[17] P.-A. Raviart, J.-M. Thomas, Introduction à l'analyse numérique des équations aux dérivées partielles, Masson, Paris, 1993.

VI. ALGORITHME LAGRANGIEN DE PENALISATION POUR PROBLEME D'OPTIMISATION AVEC CONTRAINTES ET INEGALITES VARIATIONNELLES

Lagrangian-penalization algorithm for constrained optimization and variational inequalities

P. Frankel[1] and J. Peypouquet[2]

[1] Université Montpellier II, Place Eugène Bataillon, 34095 Montpellier Cedex 05, France
p.frankel30@orange.fr

[2] Departamento de Matemática, Universidad Técnica Federico Santa María, Avenida España 1680, Valparaíso, Chile
juan.peypouquet@usm.cl

Submitted to ***Set-valued and variational analysis***

Summary. Let X, Y be real Hilbert spaces. Consider a bounded linear operator $A : X \to Y$ and a nonempty closed convex set $\mathcal{C} \subset Y$. In this paper we propose an inexact proximal-type algorithm to solve constrained optimization problems

$$(\mathcal{P}) \qquad \inf\{f(x) \ : \ Ax \in \mathcal{C}\},$$

where f is a proper lower-semicontinuous convex function on X; and variational inequalities

$$(\mathcal{VI}) \qquad 0 \in \mathcal{M}x + A^* N_{\mathcal{C}}(Ax),$$

where $\mathcal{M} : X \rightrightarrows X$ is a maximal monotone operator and $N_{\mathcal{C}}$ denotes the normal cone to the set $\mathcal{C}$. Our method combines an exact penalization procedure involving a bounded sequence of parameters, with the predictor corrector proximal multiplier method of [12]. Under suitable assumptions the sequences generated by our algorithm are proved to converge weakly to solutions of $(\mathcal{P})$ and $(\mathcal{VI})$. As applications, we describe how the algorithm can be used to find sparse solutions of linear inequality systems and solve partial differential equations by domain decomposition.

Key words: Convex optimization, proximal methods, Lagrangian, domain decomposition for PDE's.
Subject classification: 65K05, 65K10, 46N10, 49J40, 49M27, 90C25
[3]

Introduction

Let X, Y be real Hilbert spaces. Given a proper lower-semicontinuous function $f : X \to \mathbb{R} \cup \{+\infty\}$, a nonempty closed convex subset $\mathcal{C}$ of Y and a bounded linear operator $A : X \to Y$, consider the following problem

$$(\mathcal{P}) \qquad \min\{f(x) \ : \ Ax \in \mathcal{C}\}.$$

Here f is the *objective function* and $\mathcal{C}$ is a *set of constraints* for the *observations* of x given by Ax. Denote by S the solution set of $(\mathcal{P})$. Let us mention two simple instances of this problem:

1. Inequality constraints in mathematical programming. Let $A = (A_m^n)$ be a $M \times N$ matrix and let $b \in \mathbb{R}^M$. For the problem of minimizing $f : \mathbb{R}^N \to \mathbb{R}$ subject to $Ax \leq b$ the set $\mathcal{C}$ is given by $\mathcal{C} = \{y \in \mathbb{R}^M : y_m \leq b_m,\ m = 1, \ldots, M\}$. More generally, one can require the observations Ax of the vector x to take values under given thresholds $c_1, \ldots, c_J$ for valuation functions $g_1, \ldots, g_J$. In that case, $\mathcal{C} = \{y \in \mathbb{R}^M : g_j(y) \leq c_j,\ j = 1, \ldots, J\}$. □

2. Domain decomposition for partial differential equations. Let us consider a bounded domain $\Omega \subset \mathbb{R}^N$ which is decomposed in two non-overlapping subdomains Ω_1 and Ω_2 with a common interface Γ. Consider the problem of finding a function on Ω satisfying some elliptic differential equations on Ω_1 and Ω_2 and such that the jump when passing from Ω_1 to Ω_2 is nonnegative. For the Poisson equation with right-hand side h and Neumann boundary conditions, the variational formulation is

$$\inf \left\{ \frac{1}{2} \int_{\Omega_1} |\nabla u|^2 - \int_{\Omega_1} hu + \frac{1}{2} \int_{\Omega_2} |\nabla v|^2 - \int_{\Omega_2} hv;\ (u,v) \in H^1(\Omega_1) \times H^1(\Omega_2) \text{ and } u_{|\Gamma} \geq v_{|\Gamma} \right\}.$$

Here $X = H^1(\Omega_1) \times H^1(\Omega_2)$, $Y = L^2(\Gamma)$, $A(u,v) = u_{|\Gamma} - v_{|\Gamma}$, $\mathcal{C} = \{y \in Y : y \geq 0\}$ and $f(u,v) = \frac{1}{2}\int_{\Omega_1} |\nabla u|^2 - \int_{\Omega_1} hu + \frac{1}{2}\int_{\Omega_2} |\nabla v|^2 - \int_{\Omega_2} hv$. □

This paper is concerned with a new algorithm of proximal type that provides a solution for problem $(\mathcal{P})$. It can also be applied to solve constrained variational inequalities of the form

$$(\mathcal{VI}) \qquad 0 \in \mathcal{M}x + A^* N_{\mathcal{C}}(Ax),$$

where $\mathcal{M} : X \rightrightarrows X$ is a maximal monotone operator and $N_{\mathcal{C}}$ denotes the normal cone to the set $\mathcal{C}$.

Notice that x is a solution of problem $(\mathcal{P})$ if and only if $0 \in \partial(f + \delta_{\mathcal{C}} \circ A)(x)$, where $\delta_{\mathcal{C}}$ is the indicator function of the set $\mathcal{C}$. Recalling that $\partial \delta_{\mathcal{C}} = N_{\mathcal{C}}$, we observe that if $\mathcal{M} = \partial f$ then any solution of $(\mathcal{VI})$ is a solution of $(\mathcal{P})$. Equivalence holds under qualification conditions. It occurs, for instance, if $\mathcal{C} - A(\mathrm{dom} f)$ is a neighborhood of the origin (see [9, Theorem 2.168]).

Our method has been inspired by two classical approaches:

1. Penalization. Let us introduce an exact penalization function $P : Y \to [0, +\infty)$ such that $P(y) = 0$ if, and only if, $y \in \mathcal{C}$. Following [7], [14] or [4], one way to approximate points in S is to apply either a diagonal or an alternating proximal point algorithm to the family (f_k) of functions given by

$$f_k(x) = f(x) + \beta_k P(Ax), \tag{1}$$

while letting $\beta_k \to +\infty$. The idea behind is that, since the proximal point algorithm tends to minimize the function f_k, once β_k is large, the cost given by $\beta_k P(Ax)$ will force Ax to be close to $\mathcal{C}$ in some sense. This approach is especially useful when the set $\mathcal{C}$ is expressed as a sublevel set of a convex function or as intersections of such sets. Several theoretical or practical choices for the function P are available. For instance, one can take $P(\cdot) = d(\cdot, \mathcal{C})$, the distance function to $\mathcal{C}$. For the case of linear inequality constraints one can use $P(y) = \sum_{m=1}^{M} \left[y_m - b_m\right]_+$, where $[r]_+$ denotes the positive part of $r \in \mathbb{R}$.

The penalization procedure described above using (1) provides a solution of $(\mathcal{P})$. However, it often involves parameters that tend either to 0 or $+\infty$, which might lead to numerical instabilities or ill-conditioning. □

2. Lagrangian duality. Let $\sigma_{\mathcal{C}}$ denote the support function of the set $\mathcal{C}$ and define the Lagrangian function $\mathcal{L}(x,\mu) = f(x) + \langle \mu, Ax\rangle - \sigma_{\mathcal{C}}(\mu)$, where $\langle\cdot,\cdot\rangle$ denotes the inner product in Y. Observe that problem $(\mathcal{P})$ is

$$(\mathcal{P}) \qquad \inf_{x\in X} \sup_{\mu\in Y} \mathcal{L}(x,\mu) = \inf_{x\in X} \{ f(x) : Ax \in \mathcal{C}\}$$

(see [8, Chapter V]). If (x^*, μ^*) is a saddle point of $\mathcal{L}$ then $Ax^* \in \mathcal{C}$ and x^* is a solution of $(\mathcal{P})$[4]. The operator $T : X \times Y \rightrightarrows X \times Y$ defined by $T(x,\mu) = (\partial f(x) + A^*\mu, \partial\sigma_{\mathcal{C}}(\mu) - Ax)$ is maximal monotone and its zeroes coincide with the saddle points of $\mathcal{L}$ (see [20]). Therefore, one can obtain solutions of $(\mathcal{P})$ by applying the proximal point algorithm to T (see [10], [21] or [19]). One drawback is the implementation complexity due to the presence of the support function $\sigma_{\mathcal{C}}$. □

In order to solve problems $(\mathcal{P})$ and $(\mathcal{VI})$ we propose a Lagrangian-based approach that incorporates a sort of penalization function for the set $\mathcal{C}$. It is worth mentioning that neither divergent penalization parameters nor vanishing step sizes come into play. The method uses the prediction-correction ideas introduced in [12] for minimization problems, but keeping a multiplier for the constraint involving P. This multiplier can also be interpreted as a vector of penalization parameters with an updating rule that prevents them from growing indefinitely. The prediction-correction steps also allow to circumvent the problem of computing resolvents of sums. All the analysis is carried out in a Hilbert space setting.

[4] Also μ^* is a solution of the dual problem

$$(\mathcal{P}^*) \qquad \sup_{\mu\in Y} \inf_{x\in X} \mathcal{L}(x,\mu) = \sup_{\mu\in Y} \{ f^*(-A^*\mu) - \sigma_{\mathcal{C}}(\mu)\}.$$

This paper is organized as follows: In Section 1 we discuss on the problems $(\mathcal{P})$ and $(\mathcal{VI})$, alternative formulations and their sets of solutions. We present our Lagrangian-based algorithm with explicitly evaluated prediction/correction steps for the Lagrange multipliers and describe our main results. The convergence analysis in the context of problem $(\mathcal{VI})$ is presented in Section 2. Section 3 contains additional results for problem $(\mathcal{P})$. The remainder is devoted to applications. In Section 4 we explain how the algorithm can be used to obtain sparse solutions for a system of linear inequalities. Section 5 contains a domain decomposition method for partial differential equations with a unilateral transfer through the boundary.

1 Preliminaries

Since no confusion should arise, all inner products (in X, Y and $\mathbb{R}^M$) will be denoted by $\langle\cdot,\cdot\rangle$ and the corresponding norms by $|\cdot|$.

Let $P=(p_m)_{m=1}^M$ be a l-Lipschitz vector-valued function on Y such that each component p_m is nonnegative and convex. Assume that the set $\mathcal{C}$ is defined by

$$\mathcal{C}=\{\ y\in Y : P(y)=0\ \}.$$

Set $\mathbf{H}=X\times Y\times Y\times\mathbb{R}^M$. In order to simplify the notation, let us write $\partial P=(\partial p_m)_{m=1}^M$. Following [20, 6], given a maximal monotone operator $\mathcal{M}:X\rightrightarrows X$ we define the monotone[5] operator $\mathcal{N}_{\mathcal{M}}:\mathbf{H}\rightrightarrows\mathbf{H}$ by

$$\mathcal{N}_{\mathcal{M}}(x,y,\mu,\nu)=(\mathcal{M}x+A^*\mu,-\mu+\langle\nu,\partial P(y)\rangle,-Ax+y,-P(y)).$$

Since each component p_m is continuous, for each fixed $\nu\in\mathbb{R}^M$ we have $\partial(\langle\nu,P(\cdot)\rangle)(y)=\langle\nu,\partial P(y)\rangle$ for all $y\in Y$. Therefore, the operator $\langle\nu,\partial P\rangle : Y\rightrightarrows Y$ is maximal monotone. Write $\mathbf{S}_{\mathcal{M}}=\mathcal{N}_{\mathcal{M}}^{-1}0$ and observe that a point $(x^*,y^*,\mu^*,\nu^*)\in\mathbf{H}$ belongs to $\mathbf{S}_{\mathcal{M}}$ if, and only if,

$$-A^*\mu^*\in\mathcal{M}x^*,\quad \mu^*\in\langle\nu^*,\partial P(y^*)\rangle,\quad Ax^*=y^*,\quad \text{and}\quad P(y^*)=0.$$

If $(x^*,y^*,\mu^*,\nu^*)\in\mathbf{S}_{\mathcal{M}}$ then x^* satisfies $(\mathcal{VI})$ because $\langle\nu^*,\partial P(y^*)\rangle\subset N_{\mathcal{C}}(y^*)$. The converse depends on the function P. For example, if $P(\cdot)=d(\cdot,\mathcal{C})$ and x^* satisfies $(\mathcal{VI})$, then there exist y^*, μ^* and ν^* such that $(x^*,y^*,\mu^*,\nu^*)\in\mathbf{S}_{\mathcal{M}}$.

On the other hand, by introducing an auxiliary variable $y\in Y$ we can rewrite $(\mathcal{P})$ as

$$\inf\{f(x) : Ax=y \text{ and } P(y)=0\}=\inf\{f(x) : (x,y)\in\mathbf{C}\},$$

where

[5] Maximality is irrelevant for our convergence analysis.

$$\mathbf{C} = \{ (x, y) \in X \times Y \;:\; Ax = y,\, y \in \mathcal{C} \}$$

is the set of *primal feasible points.*

Define the Lagrangian function $L : \mathbf{H} \to \mathbb{R} \cup \{+\infty\}$ by

$$L(x, y, \mu, \nu) = f(x) + \langle \mu, Ax - y \rangle + \langle \nu, P(y) \rangle. \tag{2}$$

A point $w^* = (x^*, y^*, \mu^*, \nu^*) \in \mathbf{H}$ is a *saddle point* of L if

$$L(x^*, y^*, \mu, \nu) \leq L(x^*, y^*, \mu^*, \nu^*) \leq L(x, y, \mu^*, \nu^*) \tag{3}$$

for all $(x, y, \mu, \nu) \in \mathbf{H}$. The set of saddle points of L coincides with $\mathcal{S}_{\partial f}$ (see [20]). Observe that if (x^*, y^*, μ^*, ν^*) is a saddle point of the Lagrangian then $(x^*, y^*) \in \mathbf{C}$ and x^* is a solution of $(\mathcal{P})$.

In order to find points in $\mathbf{S}_{\mathcal{M}}$, we propose the following method. Let us take $w^0 \in \mathbf{H}$ and define the sequence (w^k) inductively as follows: given $w^{k-1} = (x^{k-1}, y^{k-1}, \mu^{k-1}, \nu^{k-1})$ we introduce a prediction $(\tilde{\mu}^k, \tilde{\nu}^k)$ for the multipliers using the proximal point algorithm. This idea is motivated by [12]. By linearity, this accounts to

$$(\mathbf{A1}) \qquad \begin{cases} \tilde{\mu}^k = \mu^{k-1} + \lambda_k (A x^{k-1} - y^{k-1}) \\ \tilde{\nu}^k = \nu^{k-1} + \lambda_k P(y^{k-1}). \end{cases}$$

Proximal steps with respect to the state variables (x, y) read

$$-\frac{x^k - x^{k-1}}{\lambda_k} - A^* \tilde{\mu}^k \in \mathcal{M} x^k \qquad \text{and} \qquad -\frac{y^k - y^{k-1}}{\lambda_k} + \tilde{\mu}^k \in \sum_{m=1}^{M} \tilde{\nu}_m^k \partial p_m(y^k), \tag{4}$$

respectively. If $\mathcal{M} = \partial f$ these correspond to

$$\begin{cases} x^k = \mathrm{Argmin}_{x \in X} \left\{ L(x, y^{k-1}, \tilde{\mu}^k, \tilde{\nu}^k) + \frac{1}{2\lambda_k} |x - x^{k-1}|^2 \right\} \\ y^k = \mathrm{Argmin}_{y \in Y} \left\{ L(x^{k-1}, y, \tilde{\mu}^k, \tilde{\nu}^k) + \frac{1}{2\lambda_k} |y - y^{k-1}|^2 \right\}. \end{cases}$$

Due to the maximal monotonicity of $\mathcal{M}$ and $\langle \nu, \partial P \rangle$, each of the inclusions given by (4) has a unique solution by virtue of Minty's Theorem. However, since they might be difficult to solve it is important to use approximate or relaxed versions. For $\varepsilon \geq 0$ set

$$\mathcal{M}_\varepsilon x = \{ x^* \in X \;:\; \langle x^* - u^*, x - u \rangle \geq -\varepsilon \quad \text{for all } u^* \in \mathcal{M} u \}.$$

We always have $\mathcal{M} \subset \mathcal{M}_\varepsilon$. Moreover, if $\mathcal{M} = \partial f$ then $\partial f \subset \partial_\varepsilon f \subset (\partial f)_\varepsilon$, where ∂_ε denotes the standard ε-approximate subdifferential. We consider the inclusions

(**A2**)

$$-\frac{x^k - x^{k-1}}{\lambda_k} - A^*\tilde{\mu}^k \in \mathcal{M}_{\varepsilon_k} x^k \qquad \text{and} \qquad -\frac{y^k - y^{k-1}}{\lambda_k} + \tilde{\mu}^k \in \sum_{m=1}^{M} \tilde{\nu}_m^k \partial_{\varepsilon_k} p_m(y^k),$$

for $\varepsilon_k \geq 0$. Finally, the multipliers are updated using:

$$(\mathbf{A3}) \qquad \begin{cases} \mu^k = \mu^{k-1} + \lambda_k(Ax^k - y^k) \\ \nu^k = \nu^{k-1} + \lambda_k P(y^k). \end{cases}$$

In the following sections we shall prove the weak convergence of the sequence (w^k) generated by $(\mathbf{A1}) - (\mathbf{A3})$ to a point in $\mathcal{S}_\mathcal{M}$ under a summability assumption on the error sequence (ε_k) and a boundedness assumption on the step sizes (λ_k). For a general maximal monotone operator $\mathcal{M}$ we require Y to be finite-dimensional, an assumption that is already present in [12]. When $\mathcal{M}$ is the subdifferential of some proper lower-semicontinuous function $f : X \to \mathbb{R} \cup \{+\infty\}$, this hypothesis on the dimension of Y can be eliminated. Moreover, we also establish the existence of $\lim_{k\to+\infty} L(x^k, y^k, \mu^k, \nu^k)$ and $\lim_{k\to+\infty} f(x^k)$, which provide a key tool for upgrading convergence from weak to strong in the application described in Section 5.

2 Convergence toward $\mathcal{S}_\mathcal{M}$

The purpose of this section is to prove the following:

Theorem 2.1 *Let X be a real Hilbert space and $Y = \mathbb{R}^p$. Let $\mathbf{S}_\mathcal{M} \neq \emptyset$ and assume $(\varepsilon_k) \in \ell^1$ and $0 < \underline{\lambda} \leq \lambda_k \leq \overline{\lambda} < \max\{\frac{1}{\sqrt{2}\|A\|}, \frac{1}{\sqrt{2+l^2}}\}$. Any sequence (x^k, y^k, μ^k, ν^k) generated by Algorithm $(\mathbf{A1}) - (\mathbf{A3})$ converges weakly as $k \to +\infty$ to some $(x^\infty, y^\infty, \mu^\infty, \nu^\infty) \in \mathbf{S}_\mathcal{M}$.*

We start by deriving the fundamental estimations that will support the convergence analysis. For $w \in \mathbf{H}$, let us write

$$|\!|\!|w|\!|\!|^2 = |x|^2 + |y|^2 + |\mu|^2 + |\nu|^2.$$

Lemma 2.1 *Let $(x^*, y^*, \mu^*, \nu^*) \in \mathbf{S}_\mathcal{M}$. Then for all $k \in \mathbb{N}$ we have*

$$\begin{aligned} &|\!|\!|w^k - w^*|\!|\!|^2 - |\!|\!|w^{k-1} - w^*|\!|\!|^2 + |\tilde{\mu}^k - \mu^{k-1}|^2 + |\tilde{\nu}^k - \nu^{k-1}|^2 \\ &+ \left(1 - 2\lambda_k^2\|A\|^2\right)|x^k - x^{k-1}|^2 + \left(1 - \lambda_k^2(2 + l^2)\right)|y^k - y^{k-1}|^2 \leq 2\lambda_k(M+1)\varepsilon_k \end{aligned} \quad (5)$$

Proof. Let $(x^*, y^*, \mu^*, \nu^*) \in \mathbf{S}_\mathcal{M}$. From the definition of $\mathcal{M}_\varepsilon$ and $(\mathbf{A2})$ we have

$$\left\langle A^*\mu^* - \frac{x^k - x^{k-1}}{\lambda_k} - A^*\tilde{\mu}^k, x^* - x^k \right\rangle \leq \varepsilon_k,$$

and we infer that

$$|x^k - x^*|^2 - |x^{k-1} - x^*|^2 + |x^k - x^{k-1}|^2 + 2\lambda_k\langle\tilde{\mu}^k - \mu^*, A(x^k - x^*)\rangle \le 2\lambda_k\varepsilon_k. \quad (6)$$

On the other hand, the ε_k-approximate subdifferential inequality for each $\tilde{\nu}^k p_m$ gives

$$2\lambda_k\langle\tilde{\nu}^k, P(y^*) - P(y^k)\rangle \ge -2\lambda_k\left\langle \frac{y^k - y^{k-1}}{\lambda_k} - \tilde{\mu}^k, y^* - y^k\right\rangle - 2\lambda_k M\varepsilon_k$$

by summation. Hence

$$|y^k - y^*|^2 - |y^{k-1} - y^*|^2 + |y^k - y^{k-1}|^2 + 2\lambda_k\langle\tilde{\nu}^k, P(y^k) - P(y^*)\rangle + 2\lambda_k\langle\tilde{\mu}^k, y^* - y^k\rangle \le 2\lambda_k M\varepsilon_k. \quad (7)$$

Moreover we have $\mu^* \in \langle\nu^*, \partial P(y^*)\rangle$, and so

$$2\lambda_k\langle -\mu^*, y^* - y^k\rangle - 2\lambda_k\langle\nu^*, P(y^k) - P(y^*)\rangle \le 0. \quad (8)$$

Summing up inequalities (6), (7) and (8), and using that $Ax^* = y^*$, one obtains

$$\begin{aligned} &|x^k - x|^2 - |x^{k-1} - x|^2 + |x^k - x^{k-1}|^2 \\ &+ |y^k - y|^2 - |y^{k-1} - y|^2 + |y^k - y^{k-1}|^2 \\ &+2\lambda_k\left[\langle\tilde{\mu}^k - \mu^*, Ax^k - y^k\rangle + \langle\tilde{\nu}^k - \nu^*, P(y^k)\rangle\right] \le 2\lambda_k(M+1)\varepsilon_k. \end{aligned} \quad (9)$$

We rewrite the term in the bracket as follows

$$\begin{aligned} &\langle\tilde{\mu}^k - \mu^*, Ax^k - y^k\rangle + \langle\tilde{\nu}^k - \nu^*, P(y^k)\rangle \\ &= \langle\tilde{\mu}^k - \mu^k, Ax^k - y^k\rangle + \langle\tilde{\nu}^k - \nu^k, P(y^k)\rangle + \langle\mu^k - \mu^*, Ax^k - y^k\rangle + \langle\nu^k - \nu^*, P(y^k)\rangle \\ &= \frac{1}{\lambda_k}\langle\tilde{\mu}^k - \mu^k, \mu^k - \mu^{k-1}\rangle + \frac{1}{\lambda_k}\langle\tilde{\nu}^k - \nu^k, \nu^k - \nu^{k-1}\rangle \\ &\quad + \frac{1}{\lambda_k}\langle\mu^k - \mu^*, \mu^k - \mu^{k-1}\rangle + \frac{1}{\lambda_k}\langle\nu^k - \nu^*, \nu^k - \nu^{k-1}\rangle \\ &= \frac{1}{2\lambda_k}[|\tilde{\mu}^k - \mu^{k-1}|^2 - |\tilde{\mu}^k - \mu^k|^2 - |\mu^k - \mu^{k-1}|^2] + \frac{1}{2\lambda_k}[|\tilde{\nu}^k - \nu^{k-1}|^2 - |\tilde{\nu}^k - \nu^k|^2 - |\nu^k - \nu^{k-1}|^2] \\ &\quad + \frac{1}{2\lambda_k}[|\mu^k - \mu^*|^2 + |\mu^k - \mu^{k-1}|^2 - |\mu^{k-1} - \mu^*|^2] + \frac{1}{2\lambda_k}[|\nu^k - \nu^*|^2 + |\nu^k - \nu^{k-1}|^2 - |\nu^{k-1} - \nu^*|^2]. \end{aligned} \quad (10)$$

To simplify the notation, define

$$\rho_k = |x^k - x^{k-1}|^2 + |y^k - y^{k-1}|^2 + |\tilde{\mu}^k - \mu^{k-1}|^2 + |\tilde{\nu}^k - \nu^{k-1}|^2.$$

Recall that $\|w\|^2 = |x|^2 + |y|^2 + |\mu|^2 + |\nu|^2$ for $w \in \mathbf{H}$. Replacing equality (10) in (9), we deduce that

$$\|w^k - w^*\|^2 - \|w^{k-1} - w^*\|^2 + \rho_k - |\tilde{\mu}^k - \mu^k|^2 - |\tilde{\nu}^k - \nu^k|^2 \le 2\lambda_k(M+1)\varepsilon_k.$$

To conclude, observe that

$$|\tilde{\mu}^k - \mu^k|^2 = \lambda_k^2 |A(x^{k-1} - x^k) - (y^{k-1} - y^k)|^2 \leq 2\lambda_k^2 \|A\|^2 |x^k - x^{k-1}|^2 + 2\lambda_k^2 |y^k - y^{k-1}|^2,$$

while

$$|\tilde{\nu}^k - \nu^k|^2 = \lambda_k^2 |P(y^{k-1}) - P(y^k)|^2 \leq \lambda_k^2 l^2 |y^k - y^{k-1}|^2.$$

Adding the last three inequalities we obtain (5). ■

In order to prove the convergence of the algorithm first recall the following elementary result for real sequences. A proof can be found, for instance, in [5, Lemma 2].

Lemma 2.2 *Let (a_k), (b_k) and (η_k) be real sequences. Assume that (a_k) is bounded from below, (b_k) is nonnegative and $(\eta_k) \in l^1$. Assume also that $a_{k+1} - a_k + b_k \leq \eta_k$ for every $k \in \mathbb{N}$. Then (a_k) converges and $(b_k) \in l^1$.*

An immediate consequence of Lemmas 2.1 and 2.2 is the following:

Proposition 2.1 *Let $\mathbf{S}_{\mathcal{M}} \neq \emptyset$ and assume $(\varepsilon_k) \in \ell^1$ and $0 < \underline{\lambda} \leq \lambda_k \leq \overline{\lambda} < \max\{\frac{1}{\sqrt{2}\|A\|}, \frac{1}{\sqrt{2+l^2}}\}$. We have the following:*

(i) the sequences $(|x^k - x^{k-1}|^2)$, $(|y^k - y^{k-1}|^2)$, $(|Ax^k - y^k|^2)$, $(|P(y^k)|^2)$ are summable;

(ii) for every $(x^*, y^*, \nu^*, \mu^*) \in \mathbf{S}_{\mathcal{M}}$, $\lim_{k\to+\infty} \|(x^k, y^k, \mu^k, \nu^k) - (x^*, y^*, \mu^*, \nu^*)\|$ exists in $\mathbb{R}$.

In order to prove the main result of this section we shall use Opial's Lemma [18], which we recall for the sake of completeness:

Lemma 2.3 (Opial) *Let H be a Hilbert space endowed with the norm $\|\,.\,\|$. Let (ξ_n) be a sequence of H such that there exists a nonempty set $\Xi \subset H$ which verifies*

(a) for all $\xi \in \Xi$, $\lim_{n\to+\infty} \|\xi_n - \xi\|$ exists,

(b) if $(\xi_{n_k}) \rightharpoonup \overline{\xi}$ weakly in H as $k \to +\infty$, we have $\overline{\xi} \in \Xi$.

Then the sequence (ξ_n) converges weakly in H as $n \to +\infty$ to a point in Ξ.

We are now in position to prove the main result of this section.

Proof of Theorem 2.1. Let (x^k, y^k, μ^k, ν^k) be a sequence generated by Algorithm $(\mathbf{A1}) - (\mathbf{A3})$. In view of item (ii) of Proposition 2.1, the quantity $\|w^k - w\|$ has a limit as $n \to +\infty$ for every $w \in \mathbf{S}_{\mathcal{M}}$. This shows point (a) in Opial's Lemma. To prove point (b), suppose a subsequence of (x^k, y^k, μ^k, ν^k), still denoted (x^k, y^k, μ^k, ν^k), that converges weakly to $(x^\infty, y^\infty, \mu^\infty, \nu^\infty)$, *i.e.* (x^k) weakly converges toward x^∞ in X and (y^k, μ^k, ν^k) strongly converges toward $(y^\infty, \mu^\infty, \nu^\infty)$ in $Y \times Y \times \mathbb{R}^M$. We must show that $(x^\infty, y^\infty, \mu^\infty, \nu^\infty) \in \mathbf{S}_\mathcal{M}$. Using the closedness of the function $(x, y) \in X \times Y \to |Ax - y|^2 \in \mathbb{R}_+$ and the continuity of the function P and item (i) of Proposition 2.1, we have

$$|Ax^\infty - y^\infty|^2 \leq \liminf_{k\to+\infty} |Ax^k - y^k|^2 = 0,$$
$$P(y^\infty) = \lim_{k\to+\infty} P(y^k) = 0,$$

hence $Ax^\infty - y^\infty = 0$ and $P(y^\infty) = 0$, which implies $(x^\infty, y^\infty) \in \mathbf{C}$. Let (x, x^*) be in the graph of $\mathcal{M}$. In view of (**A2**), we have

$$\langle -\frac{x^k - x^{k-1}}{\lambda_k} - A^*\tilde{\mu}^k - x^*, x^k - x\rangle \geq -\varepsilon_k.$$

Notice that, in view of Proposition 2.1(i), $\lim_{k\to+\infty} -\frac{x^k - x^{k-1}}{\lambda_k} = 0$. Moreover $\lim_{k\to+\infty} |Ax^k - y^k| = 0$, hence the sequence $(\tilde{\mu}^k)$ strongly converges in Y toward μ^∞. Using also the continuity of the operator A^*, we can pass to the limit in the above inequality to obtain

$$\langle -A^*\mu^\infty - x^*, x^\infty - x\rangle \geq 0.$$

Using the maximality of the operator $\mathcal{M}$, this implies $-A^*\mu^\infty \in \mathcal{M}x^\infty$. Let now (y, y^*) in the graph of $\langle \nu^\infty, \partial P\rangle$, we have

$$\langle \nu^\infty, P(y^k) - P(y)\rangle \geq \langle y^*, y^k - y\rangle.$$

Moreover in view of (**A2**), we have

$$\langle \tilde{\nu}^k, P(y) - P(y^k)\rangle \geq \left\langle -\frac{y^k - y^{k-1}}{\lambda_k} + \tilde{\mu}^k, y - y^k \right\rangle - M\varepsilon_k.$$

Adding these two last inequalities, we obtain

$$\langle \nu^\infty - \tilde{\nu}^k, P(y^k) - P(y)\rangle \geq \left\langle y^* + \frac{y^k - y^{k-1}}{\lambda_k} - \tilde{\mu}^k, y^k - y \right\rangle - M\varepsilon_k.$$

In view of Proposition 2.1(i), $\lim_{k\to+\infty} \frac{y^k - y^{k-1}}{\lambda_k} = 0$. Moreover $\lim_{k\to+\infty} P(y^k) = 0$, hence the sequence $(\tilde{\nu}^k)$ strongly converges in Y toward ν^∞. We can pass to the limit in the above inequality to obtain

$$\langle \mu^\infty - y^*, y^\infty - y\rangle \geq 0.$$

By maximality of the operator $\langle \nu^\infty, \partial P\rangle$, this implies that $\mu^\infty \in \langle \nu^\infty, \partial P(y^\infty)\rangle$. This achieves to prove that $(x^\infty, y^\infty, \mu^\infty, \nu^\infty) \in \mathbf{S}_\mathcal{M}$. ■

Remark 2.1 *If $\mathcal{M}$ is strongly monotone with parameter $\alpha > 0$, the algorithm can be slightly modified in order to obtain strong convergence in Theorem 2.1. It suffices to redefine the operator $\mathcal{M}_\varepsilon$ for $\varepsilon \geq 0$ as*

$$\widetilde{\mathcal{M}}_\varepsilon x = \{x^* \in X \;:\; \langle x^* - u^*, x - u\rangle \geq \alpha\|x-u\|^2 - \varepsilon \quad \text{for all } u^* \in \mathcal{M}u\ \}.$$

The strong monotonicity of $\mathcal{M}$ implies that one still has $\mathcal{M} \subset \widetilde{\mathcal{M}}_\varepsilon$. Following the argument in Lemma 2.1 we deduce that

$$\|w^k - w^*\|^2 - \|w^{k-1} - w^*\|^2 + 2\alpha\lambda_k\|x^k - x^*\|^2 \leq 2\lambda_k(M+1)\varepsilon_k$$

for all $k \in \mathbb{N}$, where x^* is the unique solution of $(\mathcal{VI})$ and w^* is any element in $\mathbf{S}_{\mathcal{M}}$. The details are left to the reader. This immediately implies that x^k converges strongly to x^* as $k \to +\infty$.

3 Further results for $\mathcal{M} = \partial f$

If $\mathcal{M} = \partial f$ a more detailed analysis can be carried out and some results can be improved. In particular, the assumption on the dimension of Y can be omitted. Moreover, part (ii) in Proposition 3.1 below is used in Section 5 to upgrade convergence from weak to strong in a domain decomposition method for partial differential equations. In this section, we assume that the primal steps are computed using the approximate subdifferentials. Namely,
($\mathbf{A2'}$)

$$-\frac{x^k - x^{k-1}}{\lambda_k} - A^*\tilde{\mu}^k \in \partial_{\varepsilon_k} f(x^k) \qquad \text{and} \qquad -\frac{y^k - y^{k-1}}{\lambda_k} + \tilde{\mu}^k \in \sum_{m=1}^M \tilde{\nu}_m^k \partial_{\varepsilon_k} p_m(y^k),$$

for $\varepsilon_k \geq 0$. We shall prove the following:

Theorem 3.1 *Let X and Y be real Hilbert spaces. Let $\mathcal{S}_{\partial f} \neq \emptyset$ and assume $(\varepsilon_k) \in \ell^1$ and $0 < \underline{\lambda} \leq \lambda_k \leq \overline{\lambda} < \max\{\frac{1}{\sqrt{2}\|A\|}, \frac{1}{\sqrt{2+l^2}}\}$. Any sequence (x^k, y^k, μ^k, ν^k) generated by Algorithm $(\mathbf{A1}) - (\mathbf{A2'}) - (\mathbf{A3})$ converges weakly as $k \to +\infty$ to some $(x^\infty, y^\infty, \mu^\infty, \nu^\infty) \in \mathcal{S}_{\partial f}$.*

We begin with a reinforced version of Lemma 2.1:

Lemma 3.1 *Let (x^*, y^*, μ^*, ν^*) have the saddle-point property. Then for all $k \in \mathbb{N}$ we have*

$$\begin{aligned}\|w^k - w^*\|^2 - \|w^{k-1} - w^*\|^2 + |\tilde{\mu}^k - \mu^{k-1}|^2 + |\tilde{\nu}^k - \nu^{k-1}|^2 \\ + \left(1 - 2\lambda_k^2\|A\|^2\right)|x^k - x^{k-1}|^2 + \left(1 - \lambda_k^2(2+l^2)\right)|y^k - y^{k-1}|^2 \\ +2\lambda_k\left[L(x^k, y^k, \mu^*, \nu^*) - L(x^*, y^*, \mu^*, \nu^*)\right] \leq 2\lambda_k(M+1)\varepsilon_k\end{aligned} \quad (11)$$

Proof. The subdifferential inequality for f gives

$$\begin{aligned}2\lambda_k(f(x)-f(x^k)) &\geq -2\lambda_k\left\langle \frac{x^k-x^{k-1}}{\lambda_k}+A^*\tilde{\mu}^k, x-x^k\right\rangle -2\lambda_k\varepsilon_k\\ &= |x^k-x|^2-|x^{k-1}-x|^2+|x^k-x^{k-1}|^2+2\lambda_k\langle\tilde{\mu}^k, A(x^k-x)\rangle-2\lambda_k\varepsilon_k\end{aligned}$$

for all $x\in X$. On the other hand, the subdifferential inequality for each $\tilde{\nu}^k p_m$ gives

$$\begin{aligned}2\lambda_k\langle\tilde{\nu}^k, P(y)-P(y^k)\rangle &\geq -2\lambda_k\left\langle \frac{y^k-y^{k-1}}{\lambda_k}-\tilde{\mu}^k, y-y^k\right\rangle -2\lambda_k M\varepsilon_k\\ &= |y^k-y|^2-|y^{k-1}-y|^2+|y^k-y^{k-1}|^2+2\lambda_k\langle\tilde{\mu}^k, y-y^k\rangle-2\lambda_k M\varepsilon_k\end{aligned}$$

for all $y\in Y$. Summing up, one obtains

$$\begin{aligned}&|x^k-x|^2-|x^{k-1}-x|^2+|x^k-x^{k-1}|^2\\ &+|y^k-y|^2-|y^{k-1}-y|^2+|y^k-y^{k-1}|^2\\ &+2\lambda_k\Big[L(x^k,y^k,\tilde{\mu}^k,\tilde{\nu}^k)-L(x,y,\tilde{\mu}^k,\tilde{\nu}^k)\Big]\leq 2\lambda_k(M+1)\varepsilon_k.\end{aligned}\tag{12}$$

Let (x^*,y^*,μ^*,ν^*) have the saddle-point property and take $x=x^*$ and $y=y^*$ in (12). Since $L(x^*,y^*,\tilde{\mu}^k,\tilde{\nu}^k)\leq L(x^*,y^*,\mu^*,\nu^*)$, we obtain

$$\begin{aligned}&|x^k-x^*|^2-|x^{k-1}-x^*|^2+|x^k-x^{k-1}|^2\\ &+|y^k-y^*|^2-|y^{k-1}-y^*|^2+|y^k-y^{k-1}|^2\\ &+2\lambda_k\Big[L(x^k,y^k,\tilde{\mu}^k,\tilde{\nu}^k)-L(x^*,y^*,\mu^*,\nu^*)\Big]\leq 2\lambda_k(M+1)\varepsilon_k.\end{aligned}\tag{13}$$

We can write

$$\begin{aligned}L(x^k,y^k,\tilde{\mu}^k,\tilde{\nu}^k)-L(x^*,y^*,\mu^*,\nu^*) &= L(x^k,y^k,\tilde{\mu}^k,\tilde{\nu}^k)-L(x^k,y^k,\mu^*,\nu^*)\\ &\quad +L(x^k,y^k,\mu^*,\nu^*)-L(x^*,y^*,\mu^*,\nu^*)\\ &= \langle\tilde{\mu}^k-\mu^*, Ax^k-y^k\rangle+\langle\tilde{\nu}^k-\nu^*, P(y^k)\rangle\\ &\quad +L(x^k,y^k,\mu^*,\nu^*)-L(x^*,y^*,\mu^*,\nu^*).\end{aligned}$$

Using equality (10), complete the proof of (11) as in Lemma 2.1. ■

The following complements Proposition 2.1.

Proposition 3.1 *Let $\mathcal{S}_{\partial f}\neq\emptyset$ and assume $(\varepsilon_k)\in\ell^1$ and $0<\underline{\lambda}\leq\lambda_k\leq\overline{\lambda}<\max\{\frac{1}{\sqrt{2}\|A\|},\frac{1}{\sqrt{2+l^2}}\}$. We have the following:*

(i) for each $(x^*,y^*,\nu^*,\mu^*)\in\mathcal{S}_{\partial f}$, the sequence $(L(x^k,y^k,\mu^*,\nu^*)-L(x^*,y^*,\mu^*,\nu^*))$ is in ℓ^1;

(ii) $\lim_{k\to+\infty}L(x^k,y^k,\mu^k,\nu^k)=L(x^*,y^*,\mu^*,\nu^*)$ and $\lim_{k\to+\infty}f(x^k)=f(x^*)$.

Proof. Item (i) is an immediate consequence of Lemmas 3.1 and 2.2 because each of the terms $L(x^k,y^k,\mu^*,\nu^*)-L(x^*,y^*,\mu^*,\nu^*)$ is nonnegative in view of the saddle-point property. We deduce that $\lim_{k\to+\infty}(L(x^k,y^k,\mu^*,\nu^*)-L(x^*,y^*,\mu^*,\nu^*))=0$. By Proposition 2.1(i), $\lim_{k\to+\infty} Ax^k-y^k=0$, $\lim_{k\to+\infty} P(y^k)=0$ and the sequences (μ^k) and (ν^k) are bounded. This easily implies (ii). ■

We can now prove the main result of this section.

Proof of Theorem 3.1. Let (x^k,y^k,μ^k,ν^k) be any sequence generated by Algorithm $(\mathbf{A1})-(\mathbf{A2'})-(\mathbf{A3})$. In view of item (ii) of Proposition 2.1, the quantity $\|w^k-w\|$ has a limit as $n\to+\infty$ for every $w\in\mathcal{S}_{\partial f}$. This shows point (a) in Opial's Lemma. To prove point (b), suppose a subsequence of (x^k,y^k,μ^k,ν^k), still denoted (x^k,y^k,μ^k,ν^k), converges weakly to $(x^\infty,y^\infty,\mu^\infty,\nu^\infty)$. We must show that $(x^\infty,y^\infty,\mu^\infty,\nu^\infty)$ is a saddle-point for the Lagrangian function L. Using the closedness of the functions $(x,y)\in X\times Y\to|Ax-y|^2\in\mathbb{R}_+$ and $|P|$ and item (i) of Proposition 2.1, we have

$$|Ax^\infty-y^\infty|^2\le\liminf_{k\to+\infty}|Ax^k-y^k|^2=0,$$
$$|P(y^\infty)|\le\liminf_{k\to+\infty}|P(y^k)|=0,$$

hence $Ax^\infty-y^\infty=0$ and $P(y^\infty)=0$, which implies $(x^\infty,y^\infty)\in\mathbf{C}$. Let us fix $(x,y)\in X\times Y$. For all $N\in\mathbb{N}$ we have

$$2\sum_{k=1}^{N}\lambda_k(L(x^k,y^k,\tilde\mu^k,\tilde\nu^k)-L(x,y,\tilde\mu^k,\tilde\nu^k))\le|x^0-x|^2+|y^0-y|^2+2\overline{\lambda}(M+1)\sum_{k=1}^{\infty}\varepsilon_k$$

in view of inequality (12). Therefore, $\liminf_{k\to+\infty}(L(x^k,y^k,\tilde\mu^k,\tilde\nu^k)-L(x,y,\tilde\mu^k,\tilde\nu^k))\le 0$. Notice that, since $\lim\limits_{k\to+\infty}|Ax^k-y^k|=\lim\limits_{k\to+\infty}|P(y^k)|=0$, the sequence $(\tilde\mu^k,\tilde\nu^k)$ converges weakly to $(\mu^\infty,\nu^\infty)\in Y\times\mathbb{R}$. We deduce that

$$\begin{aligned}\lim_{k\to+\infty}L(x,y,\tilde\mu^k,\tilde\nu^k)&=\lim_{k\to+\infty}\big(f(x)+\langle\tilde\mu^k,Ax-y\rangle+\langle\tilde\nu^k,P(y)\rangle\big)\\&=f(x)+\langle\mu^\infty,Ax-y\rangle+\langle\nu^\infty,P(y)\rangle\\&=L(x,y,\mu^\infty,\nu^\infty).\end{aligned}$$

Moreover

$$L(x^k,y^k,\tilde\mu^k,\tilde\nu^k)=f(x^k)+\langle\tilde\mu^k,Ax^k-y^k\rangle+\langle\tilde\nu^k,P(y^k)\rangle \tag{14}$$

and the last two terms tend to 0 as $k\to+\infty$. Whence

$$\begin{aligned}\liminf_{k\to+\infty}f(x^k)=\liminf_{k\to+\infty}L(x^k,y^k,\tilde\mu^k,\tilde\nu^k)\le\liminf_{k\to+\infty}L(x,y,\tilde\mu^k,\tilde\nu^k)=\lim_{k\to+\infty}L(x,y,\tilde\mu^k,\tilde\nu^k)\\=L(x,y,\mu^\infty,\nu^\infty).\end{aligned}$$

Finally, using the fact that every limit point of (x^k, y^k) is feasible along with closedness of the function f, we infer that

$$L(x^\infty, y^\infty, \mu^\infty, \nu^\infty) = f(x^\infty) \leq \liminf_{k\to+\infty} f(x^k) \leq L(x, y, \mu^\infty, \nu^\infty).$$

We now must prove that, for every $(\mu, \nu) \in Y \times \mathbb{R}$, we have

$$L(x^\infty, y^\infty, \mu, \nu) \leq L(x^\infty, y^\infty, \mu^\infty, \nu^\infty).$$

This is clear since $Ax^\infty - y^\infty = 0$ and $P(y^\infty) = 0$. We have proved that every weak cluster point of the sequence (x^k, y^k, μ^k, ν^k) is a saddle-point for the Lagrangian function L and the result follows from Opial's Lemma. ■

Remark 3.1 *Our penalization scheme is exact in the following sense: Let* $(x^*, y^*, \mu^*, \nu^*) \in \mathcal{S}_{\partial f}$ *and let*

$$\widehat{x} \in Argmin\{f(x) + \langle \nu, P(Ax)\rangle,$$

with $\nu_m > \nu_m^*$ *for* $m = 1, \ldots, M$*. Then* $\widehat{x}$ *is a solution of* $(\mathcal{P})$*. Indeed, from the definition of* $\widehat{x}$ *and the saddle-point property* (3)*, we have*

$$f(\widehat{x}) + \langle \nu, P(A\widehat{x})\rangle \leq f(x^*) \leq f(\widehat{x}) + \langle \nu^*, P(A\widehat{x})\rangle.$$

Since $\nu_m > \nu_m^*$ *for each* $m = 1, \ldots, M$ *one must have* $P(A\widehat{x}) = 0$ *and* $f(\widehat{x}) \leq f(x^*)$*, which implies* $\widehat{x}$ *is a solution of* $(\mathcal{P})$*.*

4 Sparse solutions for linear inequality systems

Let $A = (A_m^n)$ be a $M \times N$ matrix and let $b \in \mathbb{R}^M$ and consider the problem

$$\min\{ \|x\|_1 \ : \ Ax \leq b \}. \tag{15}$$

This is the convex relaxation of the nonconvex problem (see [17]) of finding the sparsest solutions to the system of inequalities $Ax \leq b$, which is stated as

$$\min\{ \|x\|_0 \ : \ Ax \leq b \},$$

where $\| \cdot \|_0$ denotes the *counting norm* (number of nonzero entries). The interested reader may consult [11], [13], [16].

The problem defined in (15) can be restated as

$$\min\{ \|x\|_1 \ : \ Ax = y, \ y \leq b \}.$$

For $m = 1, \ldots, M$ take

$$p_m(y) = \Big[y_m - b_m\Big]_+ .$$

Begin with $(x^{k-1}, y^{k-1}, \mu^{k-1}, \nu^{k-1})$ and apply the multiplier prediction steps following (**A1**):

$$\tilde{\mu}^k = \mu^{k-1} + \lambda_k(Ax^{k-1} - y^{k-1})$$

and for $m = 1, \ldots, M$

$$\tilde{\nu}_m^k = \begin{cases} \nu_m^{k-1} & \text{if} \quad y_m^{k-1} \le b_m \\ \nu_m^{k-1} + \lambda_k(y_m^{k-1} - b_m) & \text{otherwise.} \end{cases}$$

Next, the exact primal step with respect to the x-variable

$$-\frac{x^k - x^{k-1}}{\lambda_k} - A^*\tilde{\mu}^k \in \partial f(x^k)$$

reduces to

$$x_n^k = \begin{cases} x_n^{k-1} - \lambda_k(A^*\tilde{\mu}^k)_n - \lambda_k & \text{if } x_n^{k-1} - \lambda_k(A^*\tilde{\mu}^k)_n > \lambda_k \\ x_n^{k-1} - \lambda_k(A^*\tilde{\mu}^k)_n + \lambda_k & \text{if } x_n^{k-1} - \lambda_k(A^*\tilde{\mu}^k)_n < -\lambda_k \\ 0 & \text{if } x_n^{k-1} - \lambda_k(A^*\tilde{\mu}^k)_n \in [-\lambda_k, \lambda_k] \end{cases}$$

for $n = 1, \ldots, N$. On the other hand, for the y-variable we have

$$-\frac{y^k - y^{k-1}}{\lambda_k} + \tilde{\mu}^k \in \sum_{m=1}^{M} \tilde{\nu}_m^k \partial p_m(y^k),$$

which we rewrite as

$$y_m^k = \begin{cases} y_m^{k-1} + \lambda_k\tilde{\mu}_m^k - \lambda_k\tilde{\nu}_m^k & \text{if } y_m^{k-1} + \lambda_k\tilde{\mu}_m^k - b_m > \lambda_k\tilde{\nu}_m^k \\ y_m^{k-1} + \lambda_k\tilde{\mu}_m^k & \text{if } y_m^{k-1} + \lambda_k\tilde{\mu}_m^k - b_m < 0 \\ b_m & \text{if } y_m^{k-1} + \lambda_k\tilde{\mu}_m^k - b_m \in [0, \lambda_k\tilde{\nu}_m^k] \end{cases}$$

for $m = 1, \ldots, M$.

Finally we update the multipliers

$$\mu^k = \mu^{k-1} + \lambda_k(Ax^k - y^k)$$

and for $m = 1, \ldots, M$

$$\nu_m^k = \begin{cases} \nu_m^{k-1} & \text{if} \quad y_m^k \le b_m \\ \nu_m^{k-1} + \lambda_k(y_m^k - b_m) & \text{otherwise.} \end{cases}$$

A simple illustration. With no intention to test the numerical performance of the method we present the following academic example to illustrate the implementation. Let

$$A = \begin{pmatrix} -1 & 0 & -1 & 1 & 0 & 1 \\ 0 & -1 & 0 & -1 & 0 & 1 \\ 0 & 1 & -1 & 0 & 1 & 0 \\ -1 & 1 & 0 & 1 & -1 & 0 \\ 1 & 0 & -1 & -1 & -1 & 1 \\ -1 & -1 & 0 & 0 & 0 & 1 \\ 0 & -1 & 1 & -1 & -1 & 1 \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} -2 \\ -1 \\ -1 \\ 0 \\ -2 \\ -1 \\ 0 \end{pmatrix}.$$

The sparsest solution of the system of inequalitites given by $Ax \leq b$ is

$$\hat{x} = (0\ 0\ 1\ 0\ 0\ -1)'.$$

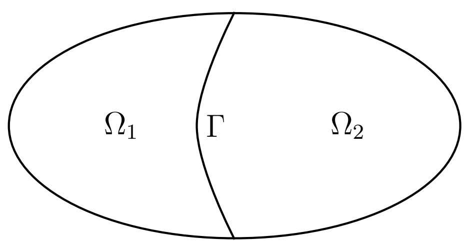
We implement our algorithm in SCILAB with $\lambda_k \equiv 0.4$, starting from 10 randomly generated initial points in $[-2, 2]^6$. The average outcome after 20 iterations was

$$\tilde{x} = (0\ 0\ 1.0052\ 0\ 0\ -0.9913)'$$

and the average processing time was 0.1 seconds in a laptop computer with a U9300 Intel(R) Core(TM)2 CPU and 3 GB of RAM.

5 Domain decomposition for partial differential equations

Let us consider a bounded domain $\Omega = \Omega_1 \cup \Omega_2 \cup \Gamma$ of $\mathbb{R}^N$ which can be decomposed in two non overlapping Lipschitz subdomains Ω_1 and Ω_2 with a common interface Γ. We assume that $\mathcal{H}^{N-1}(\Gamma) > 0$, where $\mathcal{H}^{N-1}$ is the Hausdorff measure of dimension $N - 1$. This situation is illustrated in the next figure.



Let $h \in L^2(\Omega)$. We consider the following problem

$$\begin{aligned} &\text{Minimize } \left\{ \tfrac{1}{2} \int_{\Omega_1} |\nabla u|^2 - \int_{\Omega_1} hu + \tfrac{1}{2} \int_{\Omega_2} |\nabla v|^2 - \int_{\Omega_2} hv \right\}; \\ &\text{subject to } (u, v) \in H^1(\Omega_1) \times H^1(\Omega_2) \text{ and } u_{|\Gamma} \geq v_{|\Gamma}. \end{aligned} \tag{16}$$

This kind of minimization problems often arises in the description of phenomena where the boundary is free, *i.e.* no external action is exerted on $\partial\Omega$, and involving discontinuities through the interface Γ. Here we consider the problem where the jump when passing from Ω_1 to Ω_2 is nonnegative. The case with

no condition on the jump through the interface is treated in [1] with Dirichlet conditions on the boundary of Ω and in [15] with Neumann conditions. In [6] (respectively [3]) the authors consider the problem with a no-jump condition through the interface and Dirichlet conditions on the boundary of Ω (respectively Neumann conditions), which amounts to solving a Dirichlet problem (respectively Neumann problem) on the whole set Ω by decomposition.
Notice that Problem (16) is not coercive. Under the assumptions $\int_\Omega h = 0$ and $\int_{\Omega_1} h < 0$, we can use [2, Theorem 15.1.2] to prove the existence of solutions.
Let us now show how the algorithm described by $(\mathbf{A1})-(\mathbf{A2'})-(\mathbf{A3})$ can be applied to solve problem (16). The space $X = H^1(\Omega_1)\times H^1(\Omega_2)$ is equipped with the scalar product $\langle (u_1,v_1),(u_2,v_2)\rangle_X = \int_{\Omega_1}(\nabla u_1.\nabla u_2 + u_1u_2) + \int_{\Omega_2}(\nabla v_1.\nabla v_2 + v_1v_2)$ and the corresponding norm. The space $Y = L^2(\Gamma)$ is equipped with the scalar product $\langle y_1, y_2\rangle_Y = \int_\Gamma y_1y_2$ and the associated norm. We denote respectively $A_1 : H^1(\Omega_1) \to Y$ and $A_2 : H^1(\Omega_2) \to Y$ the trace operators on Γ. Problem (16) can be reformulated as problem $(\mathcal{P})$ with the following notations

$$(\mathcal{P}) \qquad \min\{f(u)+g(v); \quad (u,v)\in X, \quad A(u,v)-y=0, \quad y\in\mathcal{C}\},$$

where

$$f(u) = \frac{1}{2}\int_{\Omega_1}|\nabla u|^2 - \int_{\Omega_1} hu \quad \text{and} \quad g(v) = \frac{1}{2}\int_{\Omega_2}|\nabla v|^2 - \int_{\Omega_2} hv, \tag{17}$$

the operator $A : X \longrightarrow Y$ is defined by $A(u,v) = A_1u - A_2v$ and the set $\mathcal{C}$ is the closed convex cone of the space Y defined by $\mathcal{C} = \{y \in Y; \quad y \geq 0\}$. We now describe the computation of the primal steps. The auxiliary varibles y and ν are used in the computation of the Lagrange multiplier approximations $\tilde{\mu}^k$ and μ^k. Their definition depends on the particular choice of the function P. One can take $P(y) = d(y,\mathcal{C})$, which in this case is the L^2-norm of the negative part of y.

Description of the primal steps. A derivative computation allows to express the exact primal steps

$$\begin{cases} u^k = \text{Argmin}\left\{f(u) + \langle\tilde{\mu}^k, A_1u\rangle + \frac{1}{2\lambda_k}|u-u^{k-1}|^2; \quad u\in H^1(\Omega_1)\right\} \\ v^k = \text{Argmin}\left\{g(v) - \langle\tilde{\mu}^k, A_2v\rangle + \frac{1}{2\lambda_k}|v-v^{k-1}|^2; \quad v\in H^1(\Omega_2)\right\}, \end{cases} \tag{18}$$

as

$$\begin{cases} \int_{\Omega_1}\nabla u^k\cdot\nabla u + \frac{1}{\lambda_k}\int_{\Omega_1}\nabla(u^k-u^{k-1})\cdot\nabla u + \frac{1}{\lambda_k}\int_{\Omega_1}(u^k-u^{k-1})u = \int_{\Omega_1} hu - \int_\Gamma \tilde{\mu}^k A_1u \\ \int_{\Omega_2}\nabla v^k\cdot\nabla v + \frac{1}{\lambda_k}\int_{\Omega_2}\nabla(v^k-v^{k-1})\cdot\nabla v + \frac{1}{\lambda_k}\int_{\Omega_2}(v^k-v^{k-1})v = \int_{\Omega_2} hv + \int_\Gamma \tilde{\mu}^k A_2v, \end{cases}$$

for all $u \in H^1(\Omega_1)$ and $v \in H^1(\Omega_2)$. These are the variational weak formulations of the following mixed Dirichlet-Neumann boundary value problems respectively on Ω_1

$$\begin{cases} -(1+\frac{1}{\lambda_k})\Delta u^k + \frac{1}{\lambda_k}u^k = h - \frac{1}{\lambda_k}\Delta u^{k-1} + \frac{1}{\lambda_k}u^{k-1} & \text{on } \Omega_1 \\ (1+\frac{1}{\lambda_k})\frac{\partial u_k}{\partial \nu} = \frac{1}{\lambda_k}\frac{\partial u^{k-1}}{\partial \nu} & \text{on } \partial\Omega_1 \cap \partial\Omega \\ (1+\frac{1}{\lambda_k})\frac{\partial u_k}{\partial \nu} = \frac{1}{\lambda_k}\frac{\partial u^{k-1}}{\partial \nu} - \tilde{\mu}^k & \text{on } \Gamma, \end{cases}$$

and Ω_2

$$\begin{cases} -(1+\frac{1}{\lambda_k})\Delta v^k + \frac{1}{\lambda_k}v^k = h - \frac{1}{\lambda_k}\Delta v^{k-1} + \frac{1}{\lambda_k}v^{k-1} & \text{on } \Omega_2 \\ (1+\frac{1}{\lambda_k})\frac{\partial v_k}{\partial \nu} = \frac{1}{\lambda_k}\frac{\partial v^{k-1}}{\partial \nu} & \text{on } \partial\Omega_2 \cap \partial\Omega \\ (1+\frac{1}{\lambda_k})\frac{\partial v_k}{\partial \nu} = \frac{1}{\lambda_k}\frac{\partial v^{k-1}}{\partial \nu} + \tilde{\mu}^k & \text{on } \Gamma. \end{cases}$$

Convergence. Since this matter is out of the scope of this paper, we shall not enter into the details concerning the existence of saddle points here. Instead we shall assume that there are such points. Under these conditions, any sequence (u^k, v^k) generated by (18) converges *strongly* in $H^1(\Omega_1) \times H^1(\Omega_2)$ to a solution $(\overline{u}, \overline{v})$ of problem (16). Indeed, let $((u^k, v^k), y^k, \mu^k, \nu^k)$ be a sequence generated by $(\mathbf{A1})-(\mathbf{A2'})-(\mathbf{A3})$ so that (u^k, v^k) satisfies (18). In view of Theorem 3.1, (u^k, v^k) converges weakly in $H^1(\Omega_1) \times H^1(\Omega_2)$ to a minimum point $(\overline{u}, \overline{v})$ of problem $(\mathcal{P})$. For the strong convergence, observe that, by the Rellich-Kondrachov Theorem, the sequence (u^k, v^k) converges to $(\overline{u}, \overline{v})$ strongly in $L^2(\Omega_1) \times L^2(\Omega_2)$. Moreover, from Proposition 3.1 (ii), we have $\lim_{k\to+\infty} f(u^k) + g(v^k) = f(\overline{u}) + g(\overline{v})$, which in turn implies that

$$\lim_{k\to+\infty} \int_{\Omega_1} |\nabla u^k|^2 + \int_{\Omega_2} |\nabla v^k|^2 = \int_{\Omega_1} |\nabla \overline{u}|^2 + \int_{\Omega_2} |\nabla \overline{v}|^2.$$

As a consequence, we have $\lim_{k\to+\infty} |(u^k, v^k)|_{H^1(\Omega_1)\times H^1(\Omega_2)} = |(\overline{u}, \overline{v})|_{H^1(\Omega_1)\times H^1(\Omega_2)}$ and we conclude that the convergence is strong.

Observe that the algorithm allows to solve the initial problem on Ω by solving separately problems on Ω_1 and Ω_2.

Acknowledgements

The authors are deeply indebted to Professor H. Attouch for his valuable comments and advice.

References

[1] H. Attouch, J. Bolte, P. Redont, A. Soubeyran, Alternating proximal algorithms for weakly coupled convex minimization problems. Applications to dynamical games and PDE's, *J. Convex Anal.*, 15 (2008), No 3, pp. 485-506.

[2] H. Attouch, G. Buttazzo, G. Michaille, Variational analysis in Sobolev and BV spaces. Applications to PDE's and optimization, *MPS/SIAM Series on Optimization*, 6. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2006.

[3] H. Attouch, A. Cabot, P. Frankel, J. Peypouquet, Alternating proximal algorithms for constrained variational inequalities. Applications to domain decomposition for PDE's, *Paper under review.*

[4] H. Attouch, M.O. Czarnecki, J. Peypouquet, Prox-penalization and splitting methods for constrained variational problems, *SIAM J. Opt.*, **21** (2011), pp. 149–173.

[5] H. Attouch, M.O. Czarnecki, J. Peypouquet, Coupling forward-backward with penalty schemes and parallel splitting for constrained variational inequalities. To appear in *SIAM J. Opt.*

[6] H. Attouch, M. Soueycatt, Augmented Lagrangian and proximal alternating direction methods of multipliers in Hilbert spaces. Applications to games , PDE's and control, *Pacific J. Opt.*, 5 (2009), No 1, pp. 17-37.

[7] A. Auslender, J.-P. Crouzeix, P. Fedit, Penalty-proximal methods in convex programming, *J. Optim. Theory Appl.* **55** (1987), no. 1, 1-21.

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

[9] F. Bonnans, A. Shapiro, Perturbation Analysis of Optimization Problems, *Springer Series in Operations Research*, Springer-Verlag New York Inc., 2000.

[10] H. Brézis, P.-L. Lions, Produits infinis de résolvantes, *Israel J. Math.*, **29** (1978), 329-345.

[11] Candès E.-J., Romberg J.-K., Tao T., Stable signal recovery from incomplete and inaccurate measurements, *Comm. Pure Appl. Math.*, **59** (2006), no. 8, 1207-1223.

[12] G. Chen, M. Teboulle, A proximal-based decomposition method for convex minimization problems, *Math. Programming* **64** (1994), no. 1, Ser. A, 81-101.

[13] S.-S. Chen, D.-L. Donoho, M.-A. Saunders, Atomic decomposition by basis pursuit, *SIAM J. Sci. Comput.*, **20** (1998), no. 1, 33-61.

[14] R. Cominetti, M. Courdurier, Coupling general penalty schemes for convex programming with the steepest descent method and the proximal point algorithm, *SIAM J. Optim*, **13** (2002), 745-765.

[15] P. Frankel, Alternating proximal algorithm with costs-to-move, dual description and application to PDE's, *to appear in Discrete Contin. Dynam. Systems - Series S.*

[16] M.-P. Friedlander, P. Tseng, Exact regularization of convex programs, *SIAM J. Optim.*, **18** (2007), no. 4, 1326-1350.

[17] J.-B. Hiriart-Urruty, H.-Y. Le, Convexifying the set of matrices of bounded rank. Applications to the quasiconvexification and convexification of the rank function, *preprint.*

[18] Z. Opial, Weak convergence of the sequence of successive approximations for nonexpansive mappings, *Bull. Amer. Math. Soc.*, **73** (1967), pp. 591-597.

[19] J. Peypouquet, S. Sorin, Evolution equations for maximal monotone operators: asymptotic analysis in continuous and discrete time. *J. Convex Anal.* **17** (2010), no. 3-4, 1113–1163.

[20] R.T. Rockafellar, Monotone operators associated with saddle-functions and minimax problems. 1970 Nonlinear Functional Analysis (Proc. Sympos. Pure Math., Vol. XVIII, Part 1, Chicago, Ill., 1968) pp. 241–250

[21] R.T. Rockafellar, Monotone Operators and the Proximal Point Algorithm, *SIAM J. Control Optim.* **14** (1976), 877-898.