Variational problems with weak coupling. Alternating minimization algorithms with friction. Links with dynamical games and P.D.E.

H. Attouch

Institut de Mathématiques et de Modélisation de Montpellier

UMR CNRS 5149

programme ANR *décisionprox* (UMII-Paris 6)

Rate-Independance, Homogenization and Multiscaling Workshop

Centro De Giorgi-Pisa

Organizing Committee: **A. Mielke and A. Visintin**

November 15-17, 2007

Pisa, Italy.

Contents:

1 Variational problems with weak coupling.

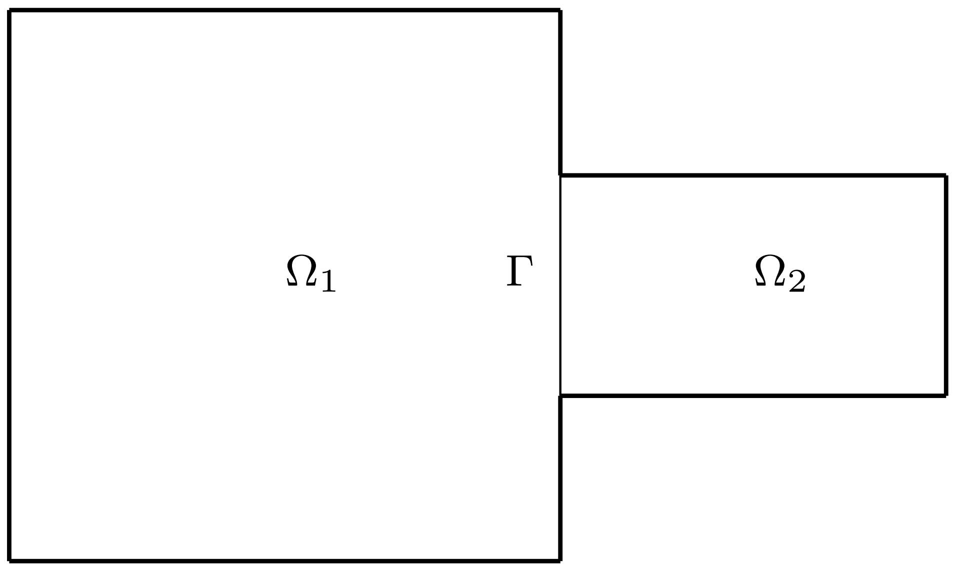
Abstract setting: $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ real Hilbert spaces.
$A : \mathcal{X} \longmapsto \mathcal{Z}$ and $B : \mathcal{Y} \longmapsto \mathcal{Z}$ linear continuous operators.

$$\boxed{\min\{f(x) + g(y) + \tfrac{\mu}{2} \parallel Ax - By \parallel_{\mathcal{Z}}^2 : \; x \in \mathcal{X}, \; y \in \mathcal{Y}\}.}$$

1. Transmission problem: $h \in L^2(\Omega)$, $\mu > 0$ given. Solve

$$\min\{\frac{1}{2}\int_{\Omega_1} |\nabla v_1|^2 + \frac{1}{2}\int_{\Omega_2} |\nabla v_2|^2 + \frac{\mu}{2}\int_{\Gamma} [v]^2 - \int_{\Omega} hv \; : \; v_i \in H^1(\Omega_i)\}$$

$v = v_1$ on Ω_1, $v = v_2$ on Ω_2 and $[v] =$ jump of v through the interface Γ.



$$\begin{cases} \mathcal{X} = H^1(\Omega_1), \; \mathcal{Y} = H^1(\Omega_2), \; \mathcal{Z} = L^2(\Gamma) \\ A : \mathcal{X} = H^1(\Omega_1) \to \mathcal{Z} = L^2(\Gamma) \text{ trace operator} \\ B : \mathcal{Y} = H^1(\Omega_2) \to \mathcal{Z} = L^2(\Gamma) \text{ trace operator} \end{cases}$$

2. Optimal control theory

State equation: $Ax = By$, $x =$ state variable, $y =$ control variable.

$$\boxed{\min\{f(x) + g(y) : Ax - By = 0,\ x \in \mathcal{X},\ y \in \mathcal{Y}\}.}$$

Penalized approximation (J.L Lions, 1983).
Augmented Lagrangian methods (Glowinski - Le Tallec, 1989).

$$\min\{f(x) + g(y) + \frac{\mu}{2} \parallel Ax - By \parallel_{\mathcal{Z}}^2 :\ x \in \mathcal{X},\ y \in \mathcal{Y}\}. \tag{1}$$

3. Potential Games

Two interrelated players 1 and 2:

- $f : \xi \in \mathcal{X} \mapsto f(\xi) \in \mathbb{R}$ **individual** payoff of player 1
- $g : \eta \in \mathcal{Y} \mapsto g(\eta) \in \mathbb{R}$ **individual** payoff of player 2.
- c: $\mathcal{X} \times \mathcal{Y} \longrightarrow \mathbb{R}$ **joint payoff** of the two players, **coupling** term. attractive case (team games), repulsive case (congestion games).

<u>Static normal form of the potential game</u> (Monderer-Shapley, 1996)

$$\boxed{\begin{cases} F(\xi,\eta) = f(\xi) \ + \ \beta c(\xi,\eta) \\ G(\xi,\eta) = g(\xi) \ + \ \mu c(\xi,\eta) \end{cases}}$$

Convex setting: Nash equilibria = solutions of

$$\min\{\frac{1}{\beta} f(x) + \frac{1}{\mu} g(y) + c(\xi,\eta) :\ x \in \mathcal{X},\ y \in \mathcal{Y}\}. \tag{2}$$

Compromise solutions between the individual and collective aspects.

4. Questions

$$\boxed{\min\{f(x)+g(y)+\frac{\mu}{2}\parallel Ax-By\parallel^2_{\mathcal{Z}}:\ x\in\mathcal{X},\ y\in\mathcal{Y}\}.}$$

- **Numerical optimization**: splitting, decomposition methods for complex variational systems involving n coupled variables (n large), parallel computing.
- **Modelling in decision sciences**: Incremental processes leading to equilibria and selection of equilibria in **the real world**: Best response dynamic, Nash equilibration processes. Friction, inertia, routines, learning, memory, path dependence aspects.

Dynamical approach to equilibria, discrete (algorithms) and continuous (differential) ***dissipative*** *systems.*

Results:

- Inertia, anchoring, frictions (viscous, dry) and memory aspects (short, long) play an important role.
- New alternating algorithms (with inertia), proximal algorithms with costs to move. New parallel algorithms for large systems (co-ordinated games).
- Convex and Nonconvex (analyticity, subanalyticity) results. Nonlinear coupling, unilateral transmission.
- Coupled inclusions associated to maximal monotone operators, coupled dissipative dynamical systems.

2 Alternating minimization algorithms. Some classical results.

2.1 Von Neumann theorem

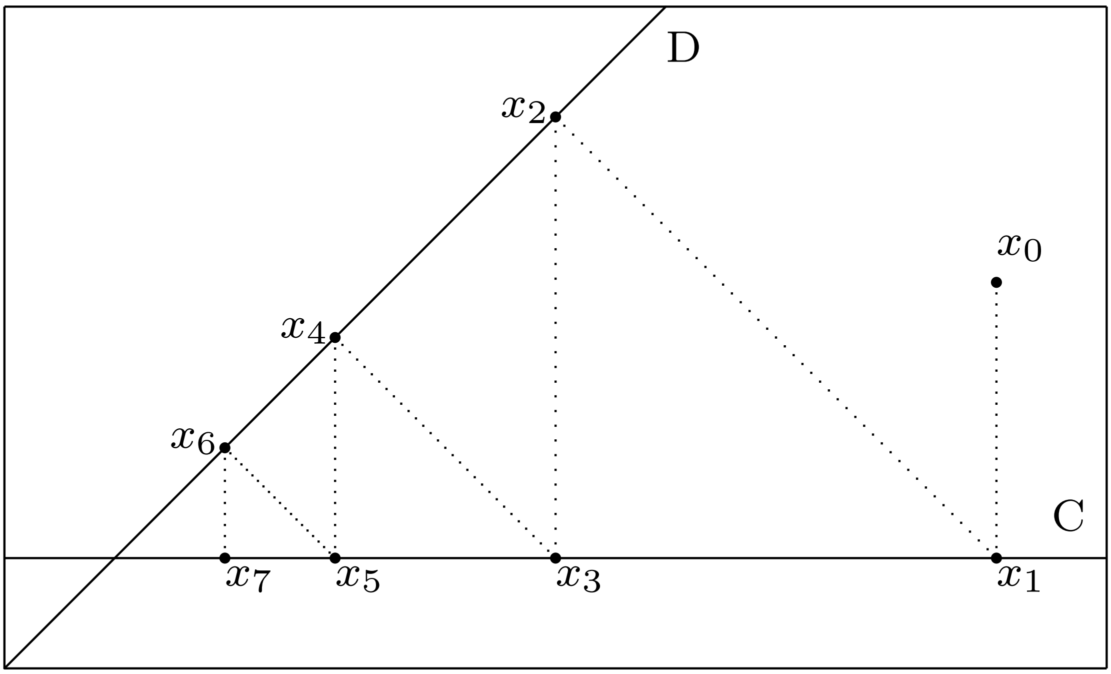
(1933), *Annals of Math.* (1950)

$\mathcal{H}$ Hilbert

C, D closed affine subspaces of $\mathcal{H}$, $C \cap D \neq \emptyset$

$x_0 \in \mathcal{H}$, $x_1 = \mathrm{proj}_C x_0$, $x_2 = \mathrm{proj}_D x_1, \ldots$

Then $x_k \overset{s-\mathcal{H}}{\longrightarrow} \mathrm{proj}_{C \cap D} x_0$ as $k \to +\infty$.



$$\lim x_{2k} = \lim x_{2k+1} = \mathrm{proj}_{C \cap D} x_0$$

Extension: Halperin (1962), N subspaces: $C_1, \ldots, C_N$

$x_0 \in \mathcal{H}$, $x_1 = \mathrm{proj}_{C_1} x_0$, $x_2 = \mathrm{proj}_{C_2} x_1, \ldots, x_N = \mathrm{proj}_{C_N} x_{N-1}$, $x_{N+1} = \mathrm{proj}_{C_1} x_N$

Solving numerically linear systems (Kaczmarz, 1937).

2.2 From *linear* → *convex* analysis

Theorem (Acker and Prestel, *Ann. Toulouse*, 1980) $\mathcal{H}$ Hilbert
$f, g : \mathcal{H} \mapsto \mathbb{R} \cup \{+\infty\}$ convex, lower semicontinuous, $\not\equiv +\infty$

$$(x_k, y_k) \to (x_{k+1}, y_k) \to (x_{k+1}, y_{k+1}):$$

$$\begin{cases} x_{k+1} = \operatorname{argmin}\{f(\xi) + \frac{\mu}{2} \| \xi - y_k \|^2 : \ \xi \in \mathcal{H}\} \\ \text{then} \\ y_{k+1} = \operatorname{argmin}\{g(\eta) + \frac{\mu}{2} \| x_{k+1} - \eta \|^2 : \ \eta \in \mathcal{H}\} \end{cases}$$

Then $(x_k, y_k) \overset{w-\mathcal{H}\times\mathcal{H}}{\longrightarrow} (x_\infty, y_\infty)$ as $k \to +\infty$, with

$$(x_\infty, y_\infty) \in \operatorname{argmin}\{f(\xi) + g(\eta) + \frac{\mu}{2} \| \xi - \eta \|^2 : \ \xi \in \mathcal{H}, \eta \in \mathcal{H}\}$$

Algorithm = alternating minimization of the bivariate function

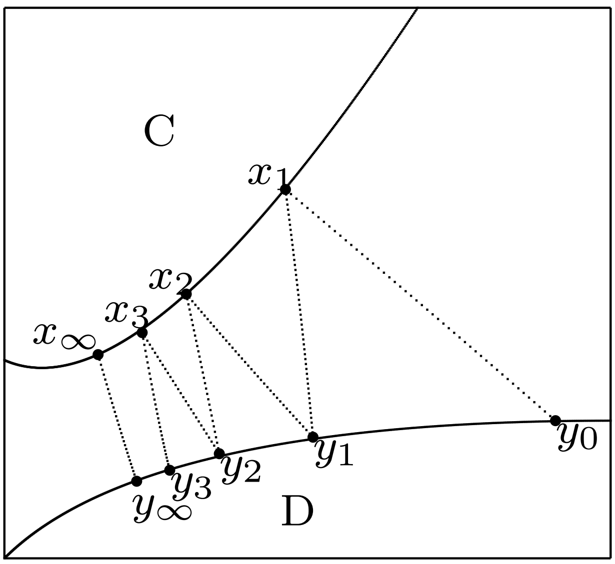
$$L_\mu(\xi, \eta) = f(\xi) + g(\eta) + \frac{\mu}{2} \| \xi - \eta \|^2$$

Application to inverse problems

(Cheney and Goldstein 59; Gubin, Polyak and Raik 67)

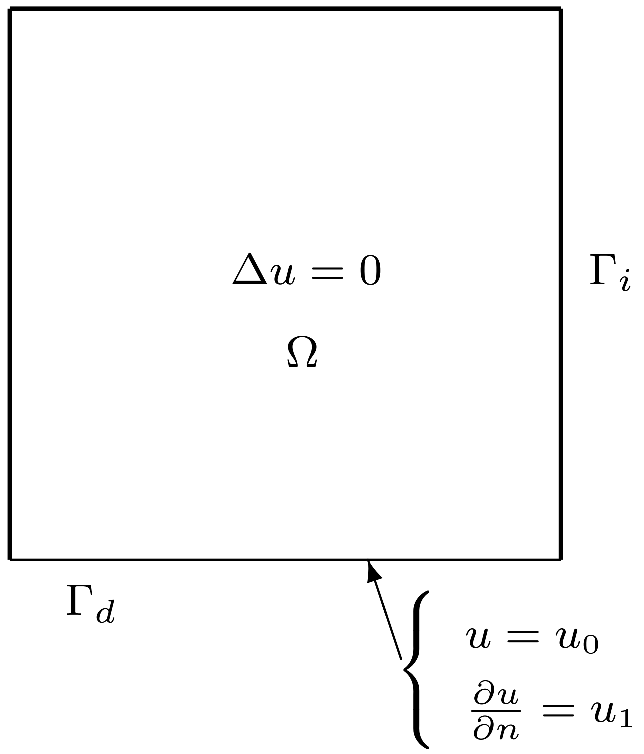
$f = \delta_C$, $g = \delta_D$, C, D closed convex sets in $\mathcal{H}$

$$\| x_\infty - y_\infty \| \ = \ \min\{\| x - y \| : \ x \in C, y \in D\}$$



2.3 Application to the inverse Cauchy problem

Cimetière, Delvare, Jaoua, Pons, *Inverse problems* (2001)



Lacking data on $\Gamma_i \neq$ superabundant measured data on Γ_d.

$$\begin{cases} \Delta u = 0 \ \text{ on } \ \Omega \\ u = u_0, \ \frac{\partial u}{\partial n} = u_1 \ \text{ on } \Gamma_d \\ u \ \text{ unknown on } \Gamma_i \end{cases}$$

Ill-posed inverse problem:

- data may be incompatible (measurements, noise).
- Hadamard example of instability: $u_n(x, y) = \frac{1}{n} \sin(nx) \mathrm{ch}(ny)$.

Alternating projection method in $\mathcal{H} = H^1(\Omega)$

$$D = \{v \in H^1(\Omega) : \ \Delta v = 0 \text{ on } \Omega, \ v = u_0 \text{ on } \ \Gamma_d\}$$

$$N = \{v \in H^1(\Omega) : \ \Delta v = 0 \text{ on } \Omega, \ \frac{\partial v}{\partial n} = u_1 \ \text{ on } \Gamma_d\}$$

2.4 Passty theorem

Theorem (Passty, *J. Math Analysis Appl.*, 1979) $\mathcal{H}$ Hilbert
$f, g : \mathcal{H} \mapsto \mathbb{R} \cup \{+\infty\}$ <u>convex</u>, lower semicontinuous, $\not\equiv +\infty$
$\partial(f+g) = \partial f + \partial g$

. $(x_0, y_0) \in \mathcal{H} \times \mathcal{H}$ given
. $(x_k, y_k) \to (x_{k+1}, y_k) \to (x_{k+1}, y_{k+1})$:

$$\begin{cases} x_{k+1} = \operatorname{argmin}\{f(\xi) + \frac{\mu_k}{2} \parallel \xi - y_k \parallel^2 : \xi \in \mathcal{H}\} \\ \text{then} \\ y_{k+1} = \operatorname{argmin}\{g(\eta) + \frac{\mu_k}{2} \parallel x_{k+1} - \eta \parallel^2 : \eta \in \mathcal{H}\} \end{cases}$$

Suppose $\mu_k \to +\infty$ *not to fast and not to slow*:
$\frac{1}{\mu_k} = \lambda_k \to 0$ with $\sum \lambda_k^2 \prec \infty$ and $\sum \lambda_k = \infty$.

Define $s_k = \sum_1^k \lambda_p$, $X_k = \frac{1}{s_k} \sum_1^k \lambda_p x_p$, $Y_k = \frac{1}{s_k} \sum_1^k \lambda_p y_p$
Then, $X_k \overset{w-\mathcal{H}}{\longrightarrow} z_\infty$, $Y_k \overset{w-\mathcal{H}}{\longrightarrow} z_\infty$ as $k \to +\infty$, with

$$z_\infty \in \operatorname{argmin}\{f(\xi) + g(\xi) : \ \xi \in \mathcal{H}\}$$

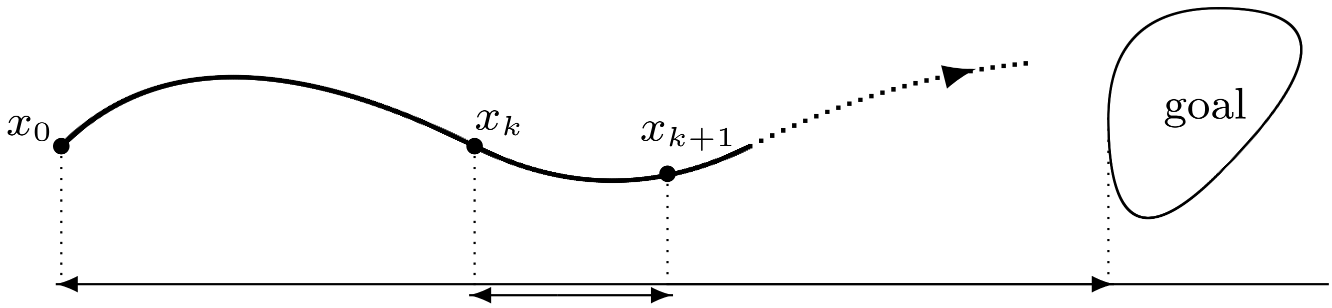
<u>Remarks on Passty theorem:</u>
1. Valid for A, B maximal monotone operators, A+B max. monotone. Extension of Brezis-Lions thm. (*Israel J. Math.*, 1978).
2. Time scaling problem with different internal and external (coupling) time scales.
3. Weak ergodic $\longrightarrow$ weak convergence? (Bruck, *JFA.*, 1975).
4. $\sum \lambda_k = \infty$ is enough? (Güler, *SIAM J. Control Opt.*, 1991).

3 Link with decision sciences and dynamical games

3.1 Dynamical decision with costs to change

- **Goal oriented human cognition (motivated agent):** Vroom (1964), Locke and Lathan (1990). Agent = problem solver.

  Solving problem = reducing (gradually) unsatisfied needs.



- **Satisficing, bounded rationality:** (Carnegie school), H. Simon, Nobel prize of economics (1978), artificial intelligence.
- **Landscape theory:** The environment (landscape) is largely unknown, exploration aspects: Levinthal and Warglien (1999), learning aspects: Sobel (2000), Berthoz (2003).
- **Worthwhile to move and inertia:** Temporary satisficing with not too much sacrificing. Attouch and Soubeyran (*J. Math. Psy.*)

$$g(x_{k+1}) - g(x_k) \;\geq\; \theta_k \; c(x_{k+1}, x_k)$$
$$\Updownarrow$$
Marginal gain $\geq$ Cost to change

$$\begin{cases} \mathcal{X} \;\; \text{state, performance, strategy space} \\ g : \mathcal{X} \longrightarrow \mathbb{R} \;\; \text{gain function} \\ c : \mathcal{X} \times \mathcal{X} \longrightarrow \mathbb{R}^+ \; \text{cost to change} \end{cases}$$

3.2 Proximal algorithms in decision sciences.

Worthwhile to move principle + Optimization $\Rightarrow$ *Proximal Dynamics.*

The vision the agent has of his environment and of his gain function depends on his current position (local aspects).

Difficulty to change, inertia, frictions, anchoring effect $\Rightarrow$ at stage k, his gain function is given by $\xi \longmapsto g(\xi) - c(\xi, x_k)$.

Proximal dynamics: $x_k \to x_{k+1}$

$$\boxed{x_{k+1} \ \in \ \operatorname{argmax}\{g(\xi) - c(\xi, x_k) \ : \ \xi \in \mathcal{X}\}}$$

Take $\xi = x_k$, sum w.r. to k, " *Finite ressource assumption* " $\sup_{\mathcal{X}} g \prec +\infty$

$$g(x_{k+1}) - g(x_k) \geq c(x_{k+1}, x_k)$$

$$\Downarrow$$

$$\sum_k c(x_{k+1}, x_k) \leq \sup_{\mathcal{X}} g - g(x_0) \prec +\infty$$

$$\Downarrow$$

$$c(x_{k+1}, x_k) \longrightarrow 0 \text{ as } \ k \to +\infty$$

Classical prox. dynamics: $\mathcal{H}$ Hilbert, $f = -g$, $c(\xi, x) = \| \xi - x \|^2$, $\lambda_k \succ 0$

$$\boxed{x_{k+1} \ \in \ \operatorname{argmin}\{f(\xi) + \frac{1}{2\lambda_k} \| \xi - x_k \|^2 \ : \ \xi \in \mathcal{H}\}}$$

Local search proximal algorithms:

$$\boxed{x_{k+1} \ \in \ \epsilon_k - \operatorname{argmax}\{g(\xi) - \theta_k c(\xi, x_k) \ : \ \xi \in E(x_k, r_k)\}}$$

- ϵ_k : psychological features (motivation, degree of resolution)
- θ_k : cognitive features (speed, learning, reactivity).
- $E(x_k, r_k)$: exploration set.

3.3 Alternating algorithms with friction and Nash potential games

2 agents (players)

- f: $\mathcal{X} \longrightarrow \mathbb{R}$ payoff of player 1, **individual** behaviour
- g: $\mathcal{Y} \longrightarrow \mathbb{R}$ payoff of player 2, **individual** behaviour
- c: $\mathcal{X} \times \mathcal{Y} \longrightarrow \mathbb{R}$ **joint payoff** of the two players, **coupling** term

Static normal form of the potential game (Monderer-Shapley, 1996)

$$\begin{cases} F(\xi, \eta) = f(\xi) \ + \ \beta c(\xi, \eta) \\ G(\xi, \eta) = g(\xi) \ + \ \mu c(\xi, \eta) \end{cases}$$

Dynamical game: Inertial non autonomous Nash equilibration process

- h: $\mathcal{X} \times \mathcal{X} \longrightarrow \mathbb{R}^+$ cost to move in $\mathcal{X}$ for player 1
- g: $\mathcal{Y} \times \mathcal{Y} \longrightarrow \mathbb{R}^+$ cost to move in $\mathcal{Y}$ for player 2
- non autonomous aspects: $\alpha_k \succ 0, \ \nu_k \succ 0$

$$\boxed{\begin{cases} x_{k+1} \in \operatorname{argmin}\{f(\xi) + \beta c(\xi, y_k) + \alpha_k h(x_k, \xi) : \xi \in \mathcal{X}\} \\ \text{then} \\ y_{k+1} \in \operatorname{argmin}\{g(\eta) + \mu c(x_{k+1}, \eta) + \nu_k k(y_k, \eta) : \eta \in \mathcal{Y}\} \end{cases}}$$

Nash equilibrium

$(\overline{x}, \overline{y})$ is a Nash equilibrium of the normal form game if

$$\begin{cases} \overline{x} \in \operatorname{argmin}\{f(\xi) + \beta c(\xi, \overline{y}) : \xi \in \mathcal{X}\} \\ \overline{y} \in \operatorname{argmin}\{g(\eta) + \mu c(\overline{x}, \eta) : \eta \in \mathcal{Y}\} \end{cases}$$

4 Convergence of proximal dynamics

4.1 The convex case

Theorem (Martinet 70, Rockafellar 76, Güler 91)

- $\mathcal{H}$ Hilbert space
- $f : \mathcal{H} \mapsto \mathbb{R} \cup \{+\infty\}$ <u>convex</u>, lower semicontinuous, $\inf_X f \succ -\infty$
- $\lambda_k \succ 0, \ \sum \lambda_k = +\infty$
- $x_{k+1} \ = \ \text{argmin}\{f(\xi) + \frac{1}{2\lambda_k} \parallel \xi - x_k \parallel^2 \ : \xi \in \mathcal{H}\}$

Then,

- a) $(x_k)_{k \in N}$ is a minimizing sequence for f.
- b) If $\text{argmin} f \neq \emptyset$
  $x_k \overset{w-\mathcal{H}}{\longrightarrow} x_\infty$ as $k \to +\infty$, with $x_\infty \in \text{argmin} f$.

1. Proximal algorithm = implicit discretization of the generalized steepest descent differential inclusion ($x(t_k) = x_k$)

$$0 \in \dot{x}(t) + \partial f(x(t))$$

$$0 \in \frac{1}{\lambda_k} \left(x(t_{k+1}) - x(t_k)\right) + \partial f \left(x(t_{k+1})\right)$$

$$\lambda_k = t_{k+1} - t_k, \ \sum_k \lambda_k = +\infty \Longleftrightarrow t_k \to +\infty$$

   Hence, asymptotic behaviour of proximal algorithms $\approx$ asymptotic behaviour of the steepest descent (Bruck theorem, Opial lemma).

2. Weak convergence of $(x_k)_{k \in N}$ and not strong convergence: see Hundal counterexample (*Nonlinear Analysis*, 2004).

4.2 The analytic and subanalytic case

- $f : \mathbb{R}^n \mapsto \mathbb{R} \cup \{+\infty\}$ lower semicontinuous, $\inf_{\mathcal{X}} f \succ -\infty$.
- The restriction of f to its domain is a continuous function.
- f satisfies the Lojasiewicz property : (L) For any limiting-critical point $\hat{x}$, that is $\partial f(\hat{x}) \ni 0$, there exist $C, \epsilon > 0$ and $\theta \in [0, 1)$ s.t.

$$|f(x) - f(\hat{x})|^{\theta} \leq C|x^*|, \forall x \in B(\hat{x}, \epsilon), \forall x^* \in \partial f(x).$$

Theorem (Attouch and Bolte, *Math. Programming*, 2007)
Let $(x_k)_{k \in N}$ be a bounded sequence generated by the proximal algorithm, then

$$\sum_{k=0}^{+\infty} |x_{k+1} - x_k| \ < +\infty,$$

and the whole sequence (x_k) converges to some critical point of f.

The rate of convergence is intimately related to the Lojasiewicz exponent which measures some flatness (curvature) property of f.

Theorem (rate of convergence) Let $x_k \longrightarrow x_\infty$ be a convergent sequence generated by the proximal algorithm. Let us denote by θ a Lojasiewicz exponent of $x_\infty \in$ crit f. Then,
(i) If $\theta = 0$, the sequence (x_k) converges in a finite number of steps,
(ii) If $\theta \in (0, \frac{1}{2}]$ then there exist $c > 0$ and $Q \in [0, 1)$ such that

$$|x_k - x_\infty| \leq c\, Q^k,$$

(iii) If $\theta \in (\frac{1}{2}, 1)$ then there exists $c > 0$ such that

$$|x_k - x_\infty| \leq c\, k^{-\frac{1-\theta}{2\theta-1}}.$$

4.3 Subdifferential calculus for nonsmooth subanalytic functions

Definition: Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a proper lower semicontinuous function.

For each $x \in \mathrm{dom} f$, the **Fréchet subdifferential** of f at x, written $\widehat{\partial}$ f(x), is the set of vectors $x^* \in \mathbb{R}^n$ which satisfy

$$\liminf_{\substack{y \neq x \\ y \to x}} \frac{1}{|x-y|}[f(y) - f(x) - \langle x^*, y - x\rangle] \geq 0.$$

If $x \notin \mathrm{dom} f$, then $\widehat{\partial}$ f(x)) $= \emptyset$.

The **limiting-subdifferential** of f at $x \in \mathbb{R}^n$, written ∂f, is defined as follows

$$\partial f(x) := \{x^* \in \mathbb{R}^n : \ \exists x_n \to x, \ f(x_n) \to f(x), \ \widehat{\partial} f(x_n) \to x^*\}.$$

Remarks: The above definition implies that $\widehat{\partial}$ f(x) $\subset \partial$ f(x), where the first set is convex while the second one is closed.

Clearly a necessary condition for $x \in \mathbb{R}^n$ to be a minimizer of f is

$$\partial f(x) \ni 0.$$

Unless f is convex the above is not a sufficient condition. A point $x \in \mathbb{R}^n$ that satisfies $\partial f(x) \ni 0$ is called *limiting-critical* or *critical*. The set of critical points of f is denoted by $\mathrm{crit} f$.

On the class of nonsmooth real valued functions involving analytic features

1. **Real-analytic functions** have the Lojasiewicz property, see Lojasiewicz (*Editions du CNRS*, Paris, 1963).

2. If f is subanalytic and is continuous on its domain with dom f closed in $\mathbb{R}^n$, in particular if f is continuous and subanalytic, it has the Lojasiewicz property, see Bolte-Daniilidis-Lewis (*JMAA.*, 2005), Kurdyka-Parusinski (*CRAS.*, 1994).

3. An interesting class of functions satisfying the Lojasiewicz property is given by **semialgebraic functions**. These are functions whose graphs can be expressed as

$$\cup_{i=1}^{p} \cap_{j=1}^{q} \{x \in \mathbb{R}^n : \ P_{ij}(x) = 0, Q_{ij}(x) \succ 0\}$$

   where for all $1 \leq i \leq p, \ 1 \leq j \leq q$ the $P_{ij}, Q_{ij} : \mathbb{R}^n \mapsto \mathbb{R}$ are polynomial functions. Due to the Tarski-Seidenberg principle, which asserts that the linear projection of a set of the above type remains of this type, semialgebraic objects enjoy remarkable stability properties.

4. Convex functions satisfying the following growth condition near the infimal set do satisfy the Lojasiewicz property:
   $\forall \widehat{x} \in \text{argmin} f, \exists C \succ 0, \exists r \geq 1, \epsilon \succ 0, \forall x \in B(\widehat{x}, \epsilon),$

$$f(x) \geq f(\widehat{x}) + Cd(x, \text{argmin} f)^r$$

5. Infinite-dimensional versions of the Lojasiewicz property have been developed in view of the asymptotic analysis of dissipative evolution equations, see Simon (*Ann. Math.*, 1983) and Haraux.

6. Kurdyka has recently established a Lojasiewicz-like inequality for functions definable in an arbitrary o-minimal structure (*Ann. Inst. Fourier*, 1998).

5 Alternating proximal minimization: the convex case

5.1 Strong coupling

Setting:

- $\mathcal{X} = \mathcal{Y} = \mathcal{H}$ Hilbert spaces.
- f, g : $\mathcal{H} \mapsto \mathbb{R} \cup \{+\infty\}$ convex, lower semicontinuous, $\not\equiv +\infty$
- $\alpha_k, \nu_k \geq 0$, $\sum_{k=0}^{+\infty} |\alpha_{k+1} - \alpha_k| \prec \infty$, $\sum_{k=0}^{+\infty} |\nu_{k+1} - \nu_k| \prec \infty$

Algorithm:

$$\boxed{\left\{\begin{array}{l} x_{k+1} = \text{argmin}_\xi \{f(\xi) + \frac{\beta}{2} \parallel \xi - y_k \parallel^2 + \frac{\alpha_k}{2} \parallel \xi - x_k \parallel^2\} \\ \text{then} \\ y_{k+1} = \text{argmin}_\eta \{g(\eta) + \frac{\mu}{2} \parallel \eta - x_{k+1} \parallel^2 + \frac{\nu_k}{2} \parallel \eta - y_k \parallel^2\} \end{array}\right.}$$

- Alternating minimization with anchoring (inertia) of

$$L_{\beta,\mu}(\xi,\eta) = \mu f(\xi) + \beta g(\eta) + \frac{\beta\mu}{2} \parallel \xi - \eta \parallel^2$$

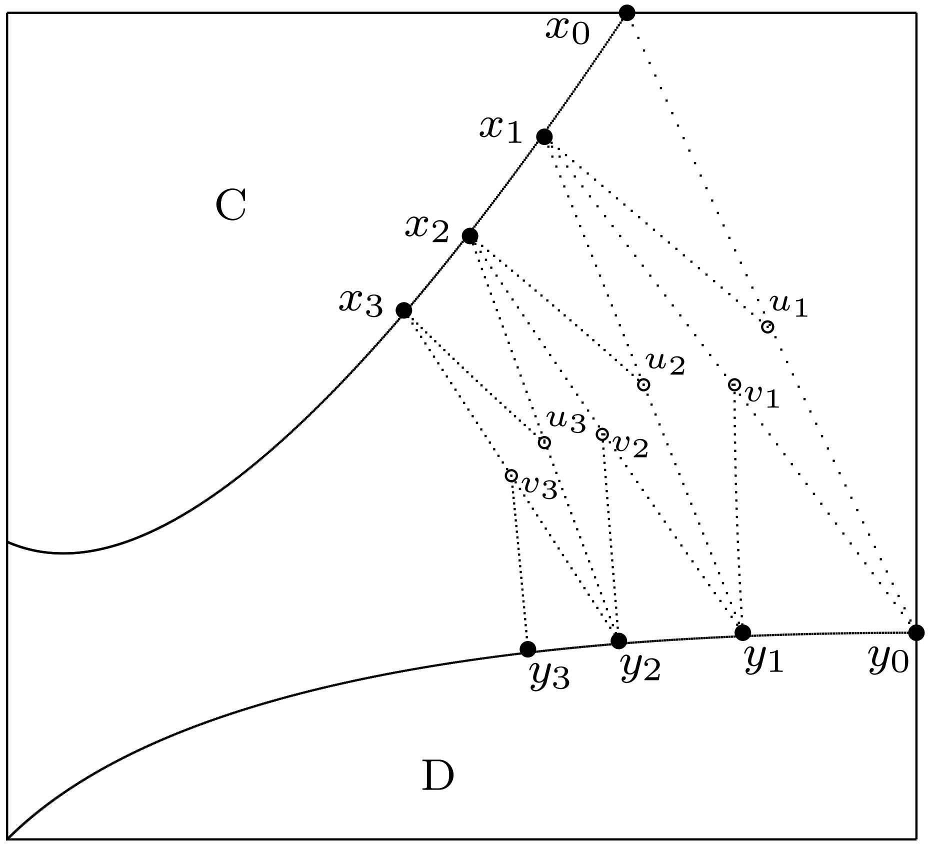
- Equilibrium: solutions of the minimization problem

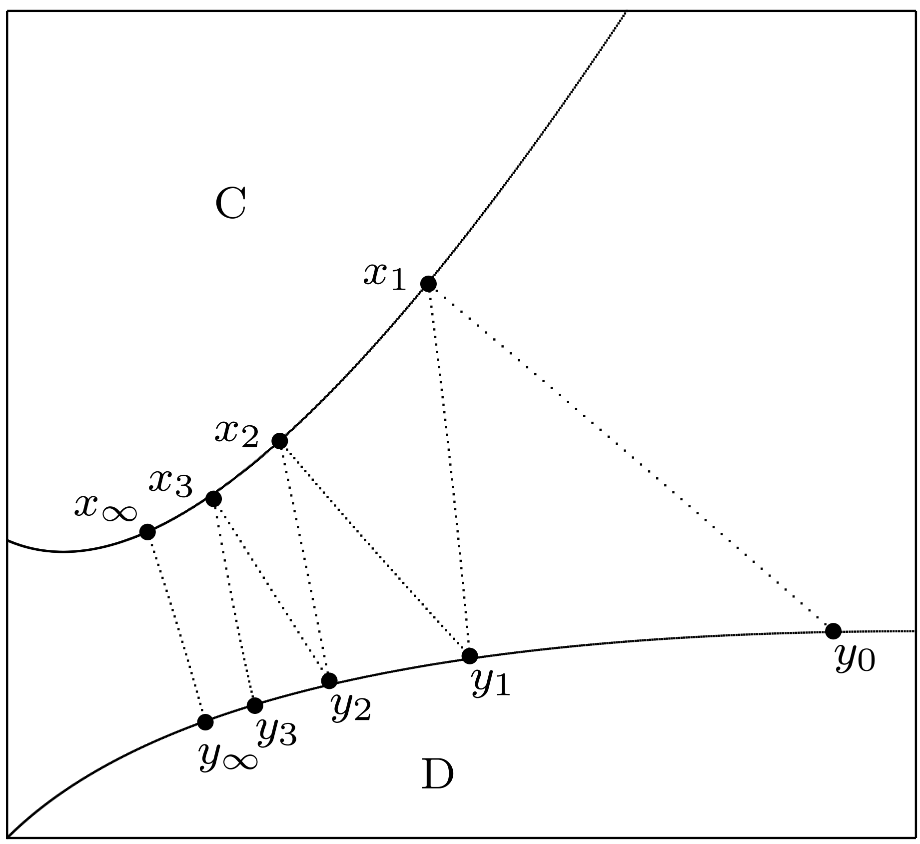
$$\min\{\mu f(\xi) + \beta g(\eta) + \frac{\beta\mu}{2} \parallel \xi - \eta \parallel^2 \colon\ \xi \in \mathcal{H}, \eta \in \mathcal{H}\}$$

Theorem (Attouch-Redont-Soubeyran, SIAM J. Optimization, 2007)
Assume S = argmin $L_{\beta,\mu} \neq \emptyset$. Then,

- a) $(x_k, y_k)_{k\in\mathbb{N}}$ is a minimizing sequence for $L_{\beta,\mu}$.
- b) $x_k \xrightarrow{w-\mathcal{H}} x_\infty$, $y_k \xrightarrow{w-\mathcal{H}} y_\infty$ with $(x_\infty, y_\infty) \in S = \text{argmin} L_{\beta,\mu}$
- c) $x_{k+1} - x_k \xrightarrow{s-\mathcal{H}} 0$, $y_{k+1} - y_k \xrightarrow{s-\mathcal{H}} 0$
- $x_k - y_k \xrightarrow{s-\mathcal{H}} x_\infty - y_\infty$ as $k \to +\infty$

Alternating projection with anchoring versus classical version





5.2 Weak coupling

- $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ Hilbert spaces.
- f : $\mathcal{X} \mapsto \mathbb{R} \cup \{+\infty\}$ convex, lower semicontinuous, $\not\equiv +\infty$
- g : $\mathcal{Y} \mapsto \mathbb{R} \cup \{+\infty\}$ convex, lower semicontinuous, $\not\equiv +\infty$
- $\mathcal{Q} : \mathcal{X} \times \mathcal{Y} \mapsto \mathbb{R}^+$ nonnegative quadratic form (convex, but possibly nondefinite), for example $\mathcal{Q}(x,y) = \| Ax - By \|^2_{\mathcal{Z}}$

Algorithm: Alternating minimization with anchoring (inertia) of

$$L_\mu(\xi,\eta) = f(\xi) + g(\eta) + \tfrac{\mu}{2}\mathcal{Q}(\xi,\eta).$$

$$\begin{cases} x_{k+1} = \text{argmin}_\xi \{f(\xi) + \frac{\mu}{2}\mathcal{Q}(\xi, y_k) + \frac{\alpha}{2} \| \xi - x_k \|^2\} \\ \text{then} \\ y_{k+1} = \text{argmin}_\eta \{g(\eta) + \frac{\mu}{2}\mathcal{Q}(x_{k+1}, \eta) + \frac{\nu}{2} \| \eta - y_k \|^2\} \end{cases}$$

Equilibrium: solutions of the minimization problem

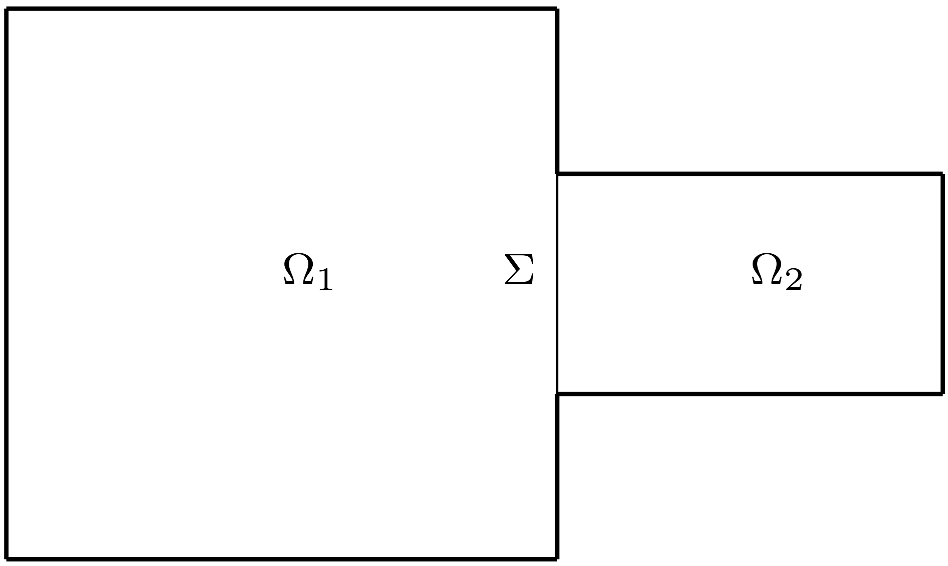
$$\min\{f(\xi) + g(\eta) + \tfrac{\mu}{2}Q(\xi,\eta) : \xi \in \mathcal{X}, \eta \in \mathcal{Y}\}$$

Theorem (Attouch-Bolte-Redont-Soubeyran, JCA, 2008)
Assume S = argmin $L_\mu \neq \emptyset$. Then,

- a) $(x_k, y_k)_{k\in\mathbb{N}}$ is a minimizing sequence for L_μ.
- b) $x_k \overset{w-X}{\longrightarrow} x_\infty$, $y_k \overset{w-Y}{\longrightarrow} y_\infty$ with $(x_\infty, y_\infty) \in S = \text{argmin} L_\mu$
- c) $x_{k+1} - x_k \overset{s-X}{\longrightarrow} 0$, $y_{k+1} - y_k \overset{s-Y}{\longrightarrow} 0$
- $f(x_k) \to f(x_\infty), g(y_k) \to g(y_\infty), \mathcal{Q}(x_k, y_k) \to \mathcal{Q}(x_\infty, y_\infty)$ as $k \to +\infty$.

5.3 Decomposition and splitting for P.D.E

Transmission through a thin weakly conducting layer, contact problems.



$$\min\{\frac{1}{2}\int_{\Omega_1}|\nabla v_1|^2+\frac{1}{2}\int_{\Omega_2}|\nabla v_2|^2+\frac{1}{2\lambda}\int_{\Sigma}[v]^2-\int_{\Omega}hv \ : v_1\in\mathcal{X}, v_2\in\mathcal{Y}\}$$

$\mathcal{X}=\{v\in\ H^1(\Omega_1),\ v=0 \text{ on } \partial\Omega\cap\partial\Omega_1\}$, $A: H^1(\Omega_1)\to\mathcal{Z}=L^2(\Gamma)$ trace
$\mathcal{Y}=\{v\in\ H^1(\Omega_2),\ v=0 \text{ on } \partial\Omega\cap\partial\Omega_2\}$, $B: H^1(\Omega_2)\to\mathcal{Z}=L^2(\Gamma)$ trace

$$f(v_1)=\frac{1}{2}\int_{\Omega_1}|\nabla v_1|^2-\int_{\Omega_1}hv_1$$
$$g(v_2)=\frac{1}{2}\int_{\Omega_2}|\nabla v_2|^2-\int_{\Omega_2}hv_2.$$

$$\begin{cases} v=v_1 \text{ on } \Omega_1\\ v=v_2 \text{ on } \Omega_2\end{cases}$$

$$A(v_1)-B(v_2)=[v]= \text{ jump of v through } \Sigma.$$

Algorithm: $u_k=(u_{1,k},u_{2,k})\in\mathcal{X}\times\mathcal{Y}=$ current point at stage k.

$$
\left\{
\begin{array}{l}
u_{1,k+1} = \text{argmin}_{v_1}\{f_1(v_1) + \frac{\mu}{2} \parallel Av_1 - Bu_{2,k} \parallel^2_{\mathcal{Z}} + \frac{\alpha}{2} \parallel v_1 - u_{1,k} \parallel^2_{\mathcal{X}}\} \\
\\
u_{2,k+1} = \text{argmin}_{v_2}\{f_2(v_2) + \frac{\mu}{2} \parallel Au_{1,k+1} - Bv_2 \parallel^2_{\mathcal{Z}} + \frac{\nu}{2} \parallel v_2 - u_{2,k} \parallel^2_{\mathcal{Y}}\}
\end{array}
\right.
\tag{3}
$$

where α and ν are given fixed positive parameters.

Optimality conditions lead to the following Dirichlet-Neumann boundary value problems respectively on Ω_1

$$
\left\{
\begin{array}{l}
-(1-\alpha)\Delta u_{1,k+1} = h + \alpha \Delta u_{1,k} \quad \text{on } \Omega_1 \\
\\
(1+\alpha)\frac{\partial u_{1,k+1}}{\partial \nu_1} + \mu u_{1,k+1} = \mu u_{2,k} + \alpha \frac{\partial u_{1,k}}{\partial \nu_1} \quad \text{on } \Gamma \\
\\
u_{1,k+1} = 0 \quad \text{on } \partial\Omega_1 \cap \partial\Omega
\end{array}
\right.
$$

and Ω_2

$$
\left\{
\begin{array}{l}
-(1-\nu)\Delta u_{2,k+1} = h + \nu \Delta u_{2,k} \quad \text{on } \Omega_2 \\
\\
(1+\nu)\frac{\partial u_{2,k+1}}{\partial \nu_2} + \mu u_{2,k+1} = \mu u_{1,k+1} + \nu \frac{\partial u_{2,k}}{\partial \nu_2} \quad \text{on } \Gamma \\
\\
u_{2,k+1} = 0 \quad \text{on } \partial\Omega_2 \cap \partial\Omega
\end{array}
\right.
$$

We have adopted the classical notations, $\frac{\partial z_i}{\partial \nu_i}$ is the derivative of z_i in the direction of ν_i which is the normal to Γ oriented outwards of Ω_i.

The above algorithm converges. The initial problem on Ω has been entirely decomposed into subproblems on Ω_1 and Ω_2.

6 Alternating proximal minimization: the analytic and subanalytic cases.

- $f : \mathbb{R}^n \mapsto \mathbb{R} \cup \{+\infty\}$, $g : \mathbb{R}^m \mapsto \mathbb{R} \cup \{+\infty\}$ lower semicontinuous,
- c: $\mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R} \cup \{+\infty\}$ is a smooth coupling function (differentiable with gradient lipschitz continuous on bounded sets).

Equilibrium: critical points of the bivariate function

$$L : \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R} \cup \{+\infty\}$$

$$L(\xi, \eta) = f(\xi) + g(\eta) + c(\xi, \eta)$$

Algorithm: Alternating minimization with anchoring (inertia) of L

$$\begin{cases} x_{k+1} \in \operatorname{argmin}\{f(\xi) + c(\xi, y_k) + \frac{1}{2\theta_k}|\xi - x_k|^2 : \xi \in \mathbb{R}^n\} \\ \text{then} \\ y_{k+1} \in \operatorname{argmin}\{g(\eta) + c(x_{k+1}, \eta) + \frac{1}{2\mu_k}|\eta - y_k|^2 : \eta \in \mathbb{R}^m\} \end{cases}$$

where $0 \prec \theta_- \le \theta_k \le \theta^+ \prec +\infty$, $0 \prec \mu_- \le \mu_k \le \mu^+ \prec +\infty$.

Let us assume that $\inf_{\mathbb{R}^n} f \succ -\infty$ and $\inf_{\mathbb{R}^m} g \succ -\infty$ and that f and g are respectively continuous on their domains dom f and dom g.

Theorem (Attouch-Bolte-Redont-Soubeyran, to appear)
Assume that $L : \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R} \cup \{+\infty\}$ satisfy the Lojasiewicz property. Let $(x_k, y_k)_{k\in\mathbb{N}}$ be a bounded sequence generated by the alternating proximal algorithm with anchoring, then

$$\sum_{k=0}^{+\infty} (|x_{k+1} - x_k| + |y_{k+1} - y_k|) \ < +\infty,$$

and the whole sequence $(x_k, y_k)_{k\in\mathbb{N}}$ converges to some critical point of L.

7 The case of n variables.

Pairwise quadratic coupling:

$$
(\mathcal{H}_n)\begin{cases}
\text{- } (\mathcal{X}_i)_{i\in\{1,\dots,n\}} \ n \text{ real Hilbert spaces;}\\
\text{- } \forall i \in \{1,...,n\},\ f_i : \mathcal{X}_i \mapsto\in \mathbb{R}\cup\{+\infty\} \text{ closed, convex, proper.}\\
\text{-}\forall 1\le i<j\le n,\ Q_{ij} : \mathcal{X}_i \times \mathcal{X}_j \mapsto \mathbb{R}^+ \text{ continuous quadratic form;}\\
\text{- } Q : (x_1,...,x_n) \in \prod_{i=1}^n \mathcal{X}_i \mapsto Q(x_1,...,x_n) \in \mathbb{R}^+\\
Q(x_1,...,x_n) = \sum_{1\le i<j\le n} Q_{ij}(x_i,x_j);\\
\text{- } L : \prod_{i=1}^n \mathcal{X}_i \mapsto \mathbb{R}\cup\{+\infty\}\\
L(x_1,...,x_n) = \sum_{i=1}^n f_i(x_i) + Q(x_1,...,x_n)\\
\text{has at least one minimum point.}
\end{cases}
$$

Alternate algorithm for successively computing the components $(x_{i,k+1})_{i\in\{1,..,n\}}$ at the $(k+1)$-th step:

$$
\begin{cases}
x_{1,k+1} = \text{argmin}_{\xi\in\mathcal{X}_1}\{f_1(\xi) + Q(\xi, x_{2,k}, ..., x_{n,k}) + \parallel \xi - x_{1,k} \parallel^2\}\\
\vdots\\
x_{i,k+1} = \text{argmin}\{f_i(\xi) + Q(x_{1,k+1},\dots,x_{i-1,k+1},\xi,x_{i+1,k},\dots,x_{n,k})\\
\qquad + \parallel \xi - x_{i,k} \parallel^2 : \xi \in \mathcal{X}_i\}\\
\vdots\\
x_{n,k+1} = \text{argmin}_{\xi\in\mathcal{X}_n}\{f_n(\xi) + Q(x_{1,k+1},\dots,x_{n-1,k+1},\xi)\\
\qquad + \parallel \xi - x_{n,k} \parallel^2\}
\end{cases}
$$

Theorem (ABRS, *J. Conv. Anal.*, 2008) Under assumptions $(\mathcal{H}_n)$, the sequence $k \mapsto (x_{i,k})_{i\in\{1,..,n\}}$ generated by the alternate algorithm is a minimizing sequence for L converging weakly in $\prod_{i=1}^n \mathcal{X}_i$ to a minimum point $(x_{i,\infty})_{i\in\{1,..,n\}}$ of L. Moreover $f_i(x_{i,k}) \to f_i(x_{i,\infty})$ and $Q(x_{1,k},\dots,x_{n,k}) \to Q(x_{1,\infty},\dots,x_{n,\infty})$ as $k \to +\infty$.

8 Splitting parallel algorithms with friction.

(Attouch-Briceno-Combettes, 2007, in progress)

Let $m \geq 2$ and $p \geq 1$ be integers and let $\mu > 0$.
For every $i \in \{1, \ldots, m\}$, $\mathcal{H}_i$ real Hilbert space, $f_i \in \Gamma_0(\mathcal{H}_i)$.
For every $k \in \{1, \ldots, p\}$, $\mathcal{G}_k$ real Hilbert space

$\varphi_k \in \Gamma_0(\mathcal{G}_k)$ differentiable with a τ_k^{-1}-lipschitz gadient.
For every $i \in \{1, \ldots, m\}$, $L_{ik}: \mathcal{H}_i \to \mathcal{G}_k$ be a bounded linear operator.

$$\boxed{\min\{\textstyle\sum_{i=1}^m f_i(x_i) + \mu \sum_{k=1}^p \varphi_k(\sum_{i=1}^m L_{ik} x_i) : x_i \in \mathcal{H}_i, i = 1, ..., m\}}$$

under the assumption that solutions exist.

Set $\nu_k = \sum_{i=1}^m \|L_{ik}\|^2$, $\beta = \frac{1}{p \max_{1 \leq k \leq p} \tau_k \nu_k}$.

Let $(\gamma_n)_{n \in \mathbb{N}}$ be a sequence in $]0, 2\beta/\mu[$ and let $(\lambda_n)_{n \in \mathbb{N}}$ be a sequence in $]0, 1]$ such that

$$\inf_{n \in \mathbb{N}} \gamma_n > 0, \sup_{n \in \mathbb{N}} \gamma_n < \frac{2\beta}{\mu}, \text{ and } \inf_{n \in \mathbb{N}} \lambda_n > 0.$$

Let $x_{1,0} \in \mathcal{H}_1, \ldots, x_{m,0} \in \mathcal{H}_m$. $\forall n \in \mathbb{N}, \forall i \in \{1, ..., m\}$, set

$$\boxed{x_{i,n+1} = (1 - \lambda_n) x_{i,n} + \lambda_n \operatorname{prox}_{\gamma_n f_i}(x_{i,n} - \mu \gamma_n B_i(x_{1,n}, ..., x_{m,n})).}$$

where

$g(x_1, ..., x_m) = \sum_{k=1}^p \varphi_k(\sum_{i=1}^m L_{ik} x_i)$, i.e.,

$B_i(x_1, ..., x_m) = \nabla_i g(x_1, ..., x_m)$,

$B_i(x_1, ..., x_m) = \sum_{k=1}^p L_{ik}^* \nabla \varphi_k(\sum_{i=1}^m L_{ik} x_i)$.

Convergence: $\forall i \in \{1, \ldots, m\}$, $(x_{i,n})_{n \in \mathbb{N}}$ converges weakly to a point $x_i \in \mathcal{H}_i$ and $(x_i)_{1 \leq i \leq m}$ is a solution to the above variational problem.