Part 1

5:18 PM

Corollary 27.6 (Douglas-Rachford algorithm)

$[\, f, g : \in \Gamma_0(\mathcal{H}),\ \operatorname{zer}(\partial f + \partial g) \neq \emptyset \ ;\ (\lambda_n)_{n\in\mathbb{N}} :\subseteq [0,2],\ \sum_{n\in\mathbb{N}} \lambda_n(2-\lambda_n) = +\infty \ ;\ \gamma \in \mathbb{R}_{++} \ ;\ x_0 \in \mathcal{H} \ ;$ set:

$$\forall_{n\in\mathbb{N}} \left| \begin{array}{l} y_n = \operatorname{Prox}_{\gamma g} x_n \\ z_n = \operatorname{Prox}_{\gamma f}(2y_n - x_n) \\ x_{n+1} = x_n + \lambda_n(z_n - y_n) \ \ ] \Rightarrow \exists_{x\in\mathcal{H}} \end{array} \right.$$

(i) $\operatorname{Prox}_{\gamma g} x \in \operatorname{Argmin}(f+g)$

(ii) $y_n - z_n \to 0$

(iii) $x_n \rightharpoonup x$

(iv) $y_n \rightharpoonup \operatorname{Prox}_{\gamma g} x$

$z_n \rightharpoonup \operatorname{Prox}_{\gamma g} x$

(v) (One of the following holds:

(a) f: uniformly convex on every nonempty bounded subset of $\operatorname{dom}\partial f$

(b) g: uniformly convex on every nonempty bounded subset of $\operatorname{dom}\partial g$)

$\Rightarrow$

$y_n \to \operatorname{Prox}_{\gamma g} x$: $\operatorname{Argmin}(f+g) = \{\operatorname{Prox}_{\gamma g} x\}$

$z_n \to \operatorname{Prox}_{\gamma g} x$: $\operatorname{Argmin}(f+g) = \{\operatorname{Prox}_{\gamma g} x\}$

Proof: Set $A := \partial f$, $B := \partial g$

$f, g \in \Gamma_0(\mathcal{H}) \Rightarrow \partial f, \partial g$: maximally monotone

$\Downarrow$

$f+g \in \Gamma_0(\mathcal{H}) \Rightarrow \partial(f+g) = \partial f + \partial g$: maximally monotone

/* set $K := \mathcal{H}$; $L = \overset{\Rightarrow L^* = \mathrm{Id}}{\mathrm{Id}}$; from Proposition 27.2 : $\operatorname{dom} f \cap \operatorname{dom} g \neq \emptyset$

Proposition 16.5. $[K$: real Hilbert space ; $f : \mathcal{H} \to]-\infty, +\infty]$; $g : K \to]-\infty, +\infty]$, proper ; $L \in \mathcal{B}(\mathcal{H}, K)$; $\lambda \in \mathbb{R}_{++}]$

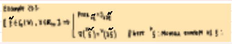
(i) $\partial(\lambda f) = \lambda \partial f$

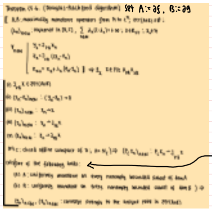
(ii) $\operatorname{dom} g \cap L(\operatorname{dom} f) \neq \emptyset \Rightarrow \partial f + L^* \circ (\partial g) \circ L \subseteq \partial(f + g \circ L)$

$\partial f + \partial g \subseteq \partial(f+g)$

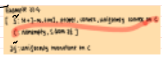
$\Rightarrow \operatorname{zer}(\partial f + \partial g) \subseteq \underbrace{\operatorname{zer} \partial(f+g) = \operatorname{Argmin}(f+g)}_{/*\text{ from Fermat's rule }*/}$

now the claims follow from the result below:





Set $A := \partial f$, $B := \partial g$



■

Proposition 27.8. (Parallel Splitting Algorithm) $[\, m : m \in \mathbb{N}, m \geq 2 \ ;\ I = \{1, \dots, m\} \ ;\ (f_i)_{i\in I} :\subseteq \Gamma_0(\mathcal{H})$;

Problem: $\min_{x\in\mathcal{H}} \sum_{i\in I} f_i(x)$: has atleast one solution ; one of the following holds;

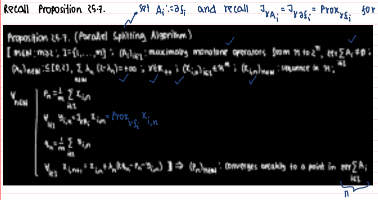
(i) $0 \in \bigcap_{i=2}^{m} \operatorname{sri}\left(\operatorname{dom} f_i - \bigcap_{j=1}^{i-1} \operatorname{dom} f_j\right)$, (ii) $\operatorname{dom} f_1 \cap \left(\bigcap_{i=2}^{m} \operatorname{int} \operatorname{dom} f_i\right) \neq \emptyset$; (iii) $\mathcal{H}$: finite-dimensional and $\bigcap_{i\in I} \operatorname{ri} \operatorname{dom} f_i \neq \emptyset$;

$(\lambda_n)_{n\in\mathbb{N}} :\subseteq [0,2],\ \sum_{n\in\mathbb{N}} \lambda_n(2-\lambda_n) = +\infty \ ;\ \gamma \in \mathbb{R}_{++} \ ;\ (x_{i,0})_{i\in I} \in \mathcal{H}^m$; set:

$$\forall_{n\in\mathbb{N}} \left| \begin{array}{l} p_n = \frac{1}{m} \sum_{i\in I} x_{i,n} \\ \forall_{i\in I}\ y_{i,n} = \operatorname{Prox}_{\gamma f_i} x_{i,n} \\ q_n = \frac{1}{m} \sum_{i\in I} y_{i,n} \\ \forall_{i\in I}\ x_{i,n+1} = x_{i,n} + \lambda_n(2q_n - p_n - y_{i,n}) \ \ ] \Rightarrow p_n \rightharpoonup p : p \in \operatorname{Argmin}_{x\in\mathcal{H}} \sum_{i\in I} f_i(x) \end{array} \right.$$

Proof:

Recall Proposition 25.7. set $A_i := \partial f_i$ and recall $J_{\gamma A_i} = J_{\gamma \partial f_i} = \operatorname{Prox}_{\gamma f_i}$ for A_i: maximally monotone



Proposition 25.7. (Parallel Splitting Algorithm)

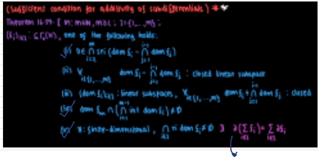
$[\, m \in \mathbb{N} : m \geq 2 \ ;\ I = \{1, \dots, m\} \ ;\ (A_i)_{i\in I}$: maximally monotone operators from $\mathcal{H}$ to $2^{\mathcal{H}}$, $\operatorname{zer} \sum_{i\in I} A_i \neq \emptyset$;

$(\lambda_n)_{n\in\mathbb{N}} :\subseteq [0,2],\ \sum_{n\in\mathbb{N}} \lambda_n(2-\lambda_n) = +\infty \ ;\ \gamma \in \mathbb{R}_{++} \ ;\ (x_{i,0})_{i\in I} \in \mathcal{H}^m \ ;\ (x_{i,n})_{n\in\mathbb{N}}$: sequence in $\mathcal{H}$;

$$\forall_{n\in\mathbb{N}} \left| \begin{array}{l} p_n = \frac{1}{m} \sum_{i\in I} x_{i,n} \\ \forall_{i\in I}\ y_{i,n} = J_{\gamma A_i} x_{i,n} = \operatorname{Prox}_{\gamma f_i} x_{i,n} \\ q_n = \frac{1}{m} \sum_{i\in I} y_{i,n} \\ \forall_{i\in I}\ x_{i,n+1} = x_{i,n} + \lambda_n(2q_n - p_n - y_{i,n}) \ \ ] \Rightarrow (p_n)_{n\in\mathbb{N}} : \text{converges weakly to a point in } \underbrace{\operatorname{zer} \sum_{i\in I} A_i}_{\operatorname{zer} \sum_{i\in I} \partial f_i} \end{array} \right.$$

$\therefore (p_n)_{n\in\mathbb{N}}$: converges weakly to a point in $\operatorname{zer} \sum_{i\in I} \partial f_i$

now



(Sufficient condition for additivity of subdifferentials) ✱

Theorem 16.39. $[\, m : m \in \mathbb{N}, m \geq 2 \ ;\ I = \{1, \dots, m\}$;

$(f_i)_{i\in I} :\subseteq \Gamma_0(\mathcal{H})$, one of the following holds:

(i) $0 \in \bigcap_{i=2}^{m} \operatorname{sri}\left(\operatorname{dom} f_i - \bigcap_{j=1}^{i-1} \operatorname{dom} f_j\right)$

(ii) $\forall_{i\in\{2,\dots,m\}}$ $\operatorname{dom} f_i - \bigcap_{j=1}^{i-1} \operatorname{dom} f_j$: closed linear subspace

(iii) $(\operatorname{dom} f_i)_{i\in I}$: linear subspaces, $\forall_{i\in\{2,\dots,m\}}$ $\operatorname{dom} f_i + \bigcap_{j=1}^{i-1} \operatorname{dom} f_j$: closed

(iv) $\operatorname{dom} f_m \cap \left(\bigcap_{i=1}^{m-1} \operatorname{int} \operatorname{dom} f_i\right) \neq \emptyset$

(v) $\mathcal{H}$: finite-dimensional, $\bigcap_{i\in I} \operatorname{ri} \operatorname{dom} f_i \neq \emptyset$ $]$ $\partial\left(\sum_{i\in I} f_i\right) = \sum_{i\in I} \partial f_i$

$\sum_{i\in I} \partial f_i = \partial \sum_{i\in I} f_i$

$(p_n)_{n\in\mathbb{N}}$: converges weakly to a point in $\operatorname{zer} \partial \sum_{i\in I} f_i = \operatorname{Argmin} \sum_{i\in I} f_i$

■

Corollary 27.9. (Forward-backward algorithm)

$[\, f \in \Gamma_0(\mathcal{H})$; $g : \mathcal{H} \to \mathbb{R}$, convex, differentiable, $\frac{1}{\beta}$-Lipschitz-continuous gradient with $\beta \in \mathbb{R}_{++}$;

$\gamma \in]0, 2\beta[$; $\delta = \min\{1, \beta/\gamma\} + \frac{1}{2}$; $(\lambda_n)_{n\in\mathbb{N}} :\subseteq [0, \delta],\ \sum_{n\in\mathbb{N}} \lambda_n(\delta - \lambda_n) = +\infty$; $x_0 \in \mathcal{H}$; $\operatorname{Argmin}(f+g) \neq \emptyset$;

$$\forall_{n\in\mathbb{N}} \left| \begin{array}{l} y_n = x_n - \gamma \nabla g(x_n) \\ x_{n+1} = x_n + \lambda_n(\operatorname{Prox}_{\gamma f} y_n - x_n) \ \ ] \Rightarrow \end{array} \right.$$

(i) $(x_n)_{n\in\mathbb{N}}$: converges weakly to a point in $\operatorname{Argmin}(f+g)$

(ii) $\left(\inf_{n\in\mathbb{N}} \lambda_n > 0 \ ;\ x \in \operatorname{Argmin}(f+g) \right) \Rightarrow \nabla g(x_n) \to \nabla g(x)$

(iii) ($\inf_{n\in\mathbb{N}} \lambda_n > 0$; one of the following holds:

(a) f: uniformly convex on every nonempty bounded subset of $\operatorname{dom} \partial f$

(b) g: uniformly convex on every nonempty bounded subset of $\mathcal{H}$) $\Rightarrow$

$x_n \to$ unique minimizer of $f+g$

Proof:

Set: $\left.\begin{array}{l} A = \partial f \\ B = \nabla g \end{array}\right\} \Rightarrow A, B$: maximally // subdifferential operator of a Γ_0 function is maximally monotone

now:

$\operatorname{dom} g = \mathcal{H}$ /* Per absurdum: $\exists_{\tilde{x}\in\mathcal{H}}$ $g(\tilde{x}) = +\infty$

g: cannot be differentiable at $\tilde{x}$ */

$\Rightarrow \operatorname{int} \operatorname{dom} g = \mathcal{H}$

$\therefore \operatorname{dom} f \cap \operatorname{int}(\operatorname{dom} g) \neq \emptyset$

$B = \nabla g \Big\} \Rightarrow A, B$: maximally // subdifferential operator of a Γ_0 function is maximally monotone

now:

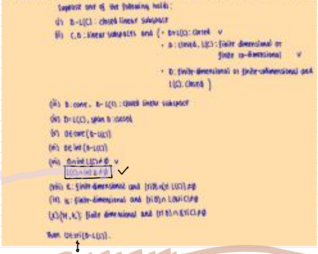
dom $g = \mathcal{H}$ /* Per absurdum: $\exists_{\tilde{x} \in \mathcal{H}}$ $g(\tilde{x}) = +\infty$ ↳ g : cannot be differentiable at $\tilde{x}$ */

$\Rightarrow$ int dom $g = \mathcal{H}$

$\therefore \underbrace{\text{dom} f}_{\text{convex}} \cap \underbrace{\text{int}(\text{dom}\, g)}_{\overset{\shortparallel}{\mathcal{H}} : \text{convex}} \neq \emptyset$

↳ as $f \in \Gamma_0(\mathcal{H})$

use



(sufficient condition for 0 belonging to the strong relative interior of a composite set) ★★

Set: $K := \mathcal{H}$; $L := \text{Id} : \mathcal{H} \to \mathcal{H}$; $C := \text{dom} f$, $D := \text{dom}\, g$

★ Proposition 6.19 [C: convex subset of $\mathcal{H}$, K: real Hilbert space, $L \in \mathcal{B}(\mathcal{H}, K)$, D: convex subset of K]

Suppose one of the following holds:

(i) $D - L(C)$: closed linear subspace

(ii) C, D : linear subspaces and { • $D + L(C)$: closed ✓

• D : closed, $L(C)$: finite dimensional or finite co-dimensional ✓

• D : finite-dimensional or finite-codimensional and $L(C)$: closed }

(iii) D : cone, $D - L(C)$: closed linear subspace

(iv) $D = L(C)$, span D : closed

(v) $0 \in \text{core}(D - L(C))$

(vi) $0 \in \text{int}(D - L(C))$

(vii) $D \cap \text{int}\, L(C) \neq \emptyset$ ✓

$\boxed{L(C) \cap \text{int}\, D \neq \emptyset}$ ✓

(viii) K : finite-dimensional and $(\text{ri}\, D) \cap (\text{ri}\, L(C)) \neq \emptyset$

(ix) K : finite-dimensional and $(\text{ri}\, D) \cap L(\text{ri}\, C) \neq \emptyset$

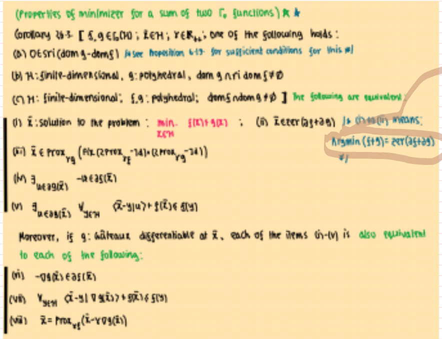
(x) $(\mathcal{H}, K)$: finite dimensional and $(\text{ri}\, D) \cap L(\text{ri}\, C) \neq \emptyset$

Then $0 \in \text{sri}(D - L(C))$.

$\updownarrow$

$0 \in \text{sri}(\text{dom}\, g - \text{dom} f)$

now recall:



(Properties of minimizer for a sum of two Γ_0 functions) ★★

Corollary 26.3 [$f, g \in \Gamma_0(\mathcal{H})$; $\bar{x} \in \mathcal{H}$; $\gamma \in \mathbb{R}_{++}$; one of the following holds:

(a) $0 \in \text{sri}(\text{dom}\, g - \text{dom} f)$ /* see Proposition 6.19 for sufficient conditions for this */

(b) $\mathcal{H}$: finite-dimensional, g : polyhedral, $\text{dom}\, g \cap \text{ri}\, \text{dom} f \neq \emptyset$

(c) $\mathcal{H}$: finite-dimensional ; f, g : polyhedral ; $\text{dom} f \cap \text{dom}\, g \neq \emptyset$] the following are equivalent:

(i) $\bar{x}$: solution to the problem : $\underset{x \in \mathcal{H}}{\text{min.}}\ f(x) + g(x)$; (ii) $\bar{x} \in \text{zer}(\partial f + \partial g)$ /* (i) ⇔ (ii) means:

Argmin $(f+g) = \text{zer}(\partial f + \partial g)$ */

(iii) $\bar{x} \in \text{Prox}_{\gamma g}\left(\text{Fix}\, (2\text{Prox}_{\gamma f} - \text{Id}) \circ (2\text{Prox}_{\gamma g} - \text{Id})\right)$

(iv) $\exists_{u \in \partial g(\bar{x})}$ $-u \in \partial f(\bar{x})$

(v) $\exists_{u \in \partial g(\bar{x})} \forall_{y \in \mathcal{H}}$ $\langle \bar{x} - y \mid u \rangle + f(\bar{x}) \leq f(y)$

Moreover, if g : Gâteaux differentiable at $\bar{x}$, each of the items (i)-(v) is also equivalent to each of the following:

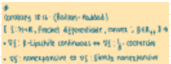
(vi) $-\nabla g(\bar{x}) \in \partial f(\bar{x})$

(vii) $\forall_{y \in \mathcal{H}}$ $\langle \bar{x} - y \mid \nabla g(\bar{x}) \rangle + f(\bar{x}) \leq f(y)$

(viii) $\bar{x} = \text{Prox}_{\gamma f}(\bar{x} - \gamma \nabla g(\bar{x}))$

$\rightsquigarrow \therefore$ Argmin $(f+g) = \text{zer}(A+B)$

now :



★ Corollary 18.16 (Baillon–Haddad)

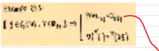
[$f : \mathcal{H} \to \mathbb{R}$, Fréchet differentiable, convex ; $\beta \in \mathbb{R}_{++}$] ⇒

• ∇f : β-Lipschitz continuous ⇔ ∇f : $\frac{1}{\beta}$-cocoercive

• ∇f : nonexpansive ⇔ ∇f : firmly nonexpansive

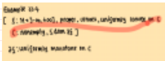
↳ $B = \nabla f$: β-cocoercive

↳ Finally we use the Theorem 25.8 which is the forward-backward algorithm in it's purest form:



Example 23.5

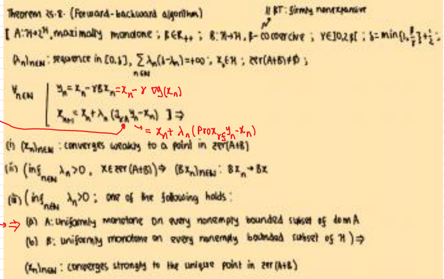
[$f \in \Gamma_0(\mathcal{H})$, $\gamma \in \mathbb{R}_{++}$] ⇒ | $\text{Prox}_{\gamma f} = J_{\gamma \partial f}$; $\partial(f^*) = (\partial f)^{-1}$



Example 22.4

[$f : \mathcal{H} \to]-\infty, +\infty]$, proper, convex, uniformly convex on C ; C : nonempty, $C \subset \text{dom}\, \partial f$]

∂f : uniformly monotone on C



Theorem 25.8 (Forward-backward algorithm) // βT : firmly nonexpansive

[$A : \mathcal{H} \to 2^{\mathcal{H}}$, maximally monotone ; $\beta \in \mathbb{R}_{++}$; $B : \mathcal{H} \to \mathcal{H}$, β-cocoercive ; $\gamma \in]0, 2\beta[$; $\delta = \min\{1, \frac{\beta}{\gamma}\} + \frac{1}{2}$;

$(\lambda_n)_{n \in \mathbb{N}}$: sequence in $[0, \delta]$, $\sum_{n \in \mathbb{N}} \lambda_n(\delta - \lambda_n) = +\infty$; $x_0 \in \mathcal{H}$; $\text{zer}(A+B) \neq \emptyset$;

$$\forall_{n \in \mathbb{N}} \left| \begin{array}{l} y_n = x_n - \gamma B x_n = x_n - \gamma \nabla g(x_n) \\ x_{n+1} = x_n + \lambda_n (J_{\gamma A} y_n - x_n) \end{array} \right.] \Rightarrow$$

$= x_n + \lambda_n(\text{Prox}_{\gamma f}\, y_n - x_n)$

(i) $(x_n)_{n \in \mathbb{N}}$: converges weakly to a point in $\text{zer}(A+B)$

(ii) $(\inf_{n \in \mathbb{N}} \lambda_n > 0$, $x \in \text{zer}(A+B)) \Rightarrow (Bx_n)_{n \in \mathbb{N}}$: $Bx_n \to Bx$

(iii) $(\inf_{n \in \mathbb{N}} \lambda_n > 0$; one of the following holds:

⇒ (a) A : Uniformly monotone on every nonempty bounded subset of dom A

(b) B : uniformly monotone on every nonempty bounded subset of $\mathcal{H}$) ⇒

$(x_n)_{n \in \mathbb{N}}$: converges strongly to the unique point in $\text{zer}(A+B)$

Proposition 27.15 (Primal-dual algorithm) [$\varphi \in \Gamma_0(\mathcal{H})$; $\psi \in \Gamma_0(K)$; $z \in \mathcal{H}$; $r \in K$; $L : \in \mathcal{B}(\mathcal{H}, K)$, $L \neq 0$, $r \in \text{sri}(L(\text{dom}\, \varphi) - \text{dom}\, \psi)$;

Primal problem:

$$\underset{x \in \mathcal{H}}{\text{min.}}\ \varphi(x) + \psi(Lx - r) + \tfrac{1}{2}\|x - z\|^2$$ [eq: 27.32]

Dual problem:

$$\underset{v \in K}{\text{min.}}\ (\varphi^*)(L^* v + z) + \psi^*(-v) - \langle v \mid r \rangle$$ [eq: 27.33]

$\gamma \in]0, \frac{2}{\|L\|^2}[$; $\delta = \min\{1, \frac{1}{\gamma} \cdot \frac{1}{\|L\|^2}\} + \frac{1}{2}$; $(\lambda_n)_{n \in \mathbb{N}}$: $\subseteq]0, \delta[$, $\inf_{n \in \mathbb{N}} \lambda_n > 0$, $\sup_{n \in \mathbb{N}} \lambda_n < \delta$; $v_0 \in K$;

Set:

$$\forall_{n \in \mathbb{N}} \left| \begin{array}{l} x_n = \text{Prox}_{\varphi}(L^* v_n + z) \\ v_{n+1} = v_n - \lambda_n\left(\text{Prox}_{\gamma \psi^*}(\gamma(Lx_n - r) - v_n) + v_n\right) \end{array} \right.$$ [eq: 27.34]

$\bar{x}$: unique solution to the primal problem] ⇒

(i) $v_n \rightharpoonup \bar{v}$: solution to the dual problem

$\bar{x} = \text{Prox}_{\varphi}(L^* \bar{v} + z)$

(ii) $x_n \to \bar{x}$

Proof:

Set $h : \mathcal{H} \to]-\infty, +\infty] : x \mapsto \varphi(x) + \frac{1}{2}\|x - z\|^2 \in \Gamma_0(\mathcal{H}) \Rightarrow \text{dom}\, \varphi = \text{dom}\, h$ (z : fixed vector)

$j : K \to]-\infty, +\infty] : y \mapsto \psi(y - r) \in \Gamma_0(K) \Rightarrow \text{dom}\, j = \text{dom}\, \psi$

↳ $\therefore j(y) = \psi(\underbrace{y - r}_{\tilde{y}}) \Leftrightarrow \psi(\tilde{y}) = j(\tilde{y} + r)$

then primal problem : $\left(\underset{x \in \mathcal{H}}{\min}\ \underbrace{\varphi(x) + \tfrac{1}{2}\|x - z\|^2}_{h(x)} + \underbrace{\psi(Lx - r)}_{j(Lx)}\right) = \left(\underset{x \in \mathcal{H}}{\text{min.}}\ h(x) + j(Lx)\right) = \min\, (h + j \circ L)(\mathcal{H})$ [eq: 27.35]

given $r \in \text{sri}(L(\text{dom}\, \varphi) - \text{dom}\, \psi)$ /* $\tilde{x} \in \text{sri}\, C \Leftrightarrow \text{cone}(C - \tilde{x}) = \overline{\text{span}}\, (C - x)$ */

$\Leftrightarrow \text{cone}(L(\text{dom}\, \varphi) - \text{dom}\, \psi - r) = \overline{\text{span}}(L(\text{dom}\, \varphi) - \text{dom}\, \psi - r)$

now $L(\text{dom}\, \varphi) - \text{dom}\, \psi - r$

$= L\{x \in \mathcal{H} \mid \varphi(x) < +\infty\} - \{y \in K \mid \psi(y) < +\infty\} - r$

$= \{Lx \in K \mid \varphi(x) < +\infty\} - \{y \in K \mid \psi(y) < +\infty\} - r$

$= \{Lx - y - r \mid \varphi(x) < +\infty, \psi(y) < +\infty\}$

$\quad \overset{\shortparallel}{\underbrace{\varphi(x) + \tfrac{1}{2}\|x - z\|^2}_{\shortparallel\, h(x)} < +\infty}$ $\quad \overset{\shortparallel}{j(y + r) < +\infty}$

$= \{Lx - (y + r) \mid h(x) < +\infty, j(y + r) < +\infty\}$

$= L\{x \mid h(x) < +\infty\} - \{\underbrace{y + r}_{\overset{\shortparallel}{\tilde{y}} \text{ say}} \mid j(y + r) < +\infty\}$

$= L\underbrace{\{x \mid h(x) < +\infty\}}_{\text{dom}\, h} - \underbrace{\{\tilde{y} \mid j(\tilde{y}) < +\infty\}}_{\text{dom}\, j} = L(\text{dom}\, h) - \text{dom}\, j$

$$\underset{\tilde{y}\text{ say}}{\underbrace{\quad}}$$

$$= L\{\underbrace{x \mid h(x) < +\infty}_{\text{dom } h}\} - \{\underbrace{\tilde{y} \mid j(\tilde{y}) < +\infty}_{\text{dom } j}\} = L(\text{dom } h) - \text{dom } j$$

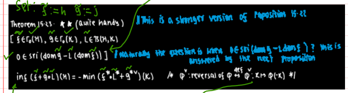
$$\text{cone}\left(L(\text{dom}h) - \text{dom}j\right) = \overline{\text{span}}\left(L(\text{dom}h) - \text{dom}j\right)$$

$$\Leftrightarrow \boxed{0 \in \text{sri}\left(L(\text{dom}h) - \text{dom}j\right)}$$ /* $\tilde{x} \in \text{sri}\, C \Leftrightarrow \text{cone}(C - \tilde{x}) = \overline{\text{span}}(C - x)$ */

$$\Leftrightarrow 0 \in \text{sri}(\text{dom}j - L(\text{dom}h))$$

now recall:

Set: $\tilde{f} := h \quad \tilde{g} := j$



Theorem 15.23 ★★ (Quite handy) //This is a stronger version of Proposition 15.22

[$f \in \Gamma_0(\mathcal{H})$, $g \in \Gamma_0(\mathcal{K})$, $L \in \mathcal{B}(\mathcal{H},\mathcal{K})$

$0 \in \text{sri}(\text{dom}\, g - L(\text{dom} f))$] // naturally the question is when $0 \in \text{sri}(\text{dom}\, g - L\,\text{dom} f)$? this is answered by the next proposition

$\inf(f + g \circ L)(\mathcal{H}) = -\min(f^* \circ L^* + g^{*\vee})(\mathcal{K})$ /* $\varphi^\vee$: reversal of $\varphi \overset{\text{def}}{\Leftrightarrow} \varphi^\vee : x \mapsto \varphi(-x)$ */

$$\inf(h + j \circ L)(\mathcal{H}) = -\min(\underset{\substack{\downarrow\\ \text{conjugate}\\ \text{function}}}{h^*} \circ \underset{\substack{\text{adjoint of the}\\ \text{linear operator}}}{L^*} + j^{*\vee})(\mathcal{K}) = -\min_{v \in \mathcal{K}}(h^*(L^*v) + \underbrace{j^{*\vee}(v)}_{j^*(-v)}) = -\min_{v \in \mathcal{K}}(h^*(L^*v) + j^*(-v))$$

minimizer exists in $\mathcal{K}$

so, inf on R.H.S has a minimizer in $\mathcal{H}$,

⇒ can be replaced with min.

[eq: 27.35.1]

[eq: 27.35] : primal problem : $\min(h + j \circ L)(\mathcal{H}) = \inf(h + j \circ L)(\mathcal{H}) = -\min_{v \in \mathcal{K}}(h^*(L^*v) + j^*(-v))$: has atleast one solution in $\mathcal{K}$

this is the dual problem in min. form, now we want to show that it coincides with [eq: 27.33]

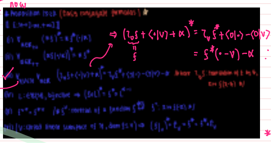
now: $h(x) = \varphi(x) + \frac{1}{2}\|x - z\|^2 = \varphi(x) + \frac{1}{2}\|x\|^2 + \frac{1}{2}\|z\|^2 - \cancel{2} \cdot \frac{1}{\cancel{2}}\langle x | z \rangle$

$= \varphi(x) + \frac{1}{2}\|x\|^2 + \langle x | -z \rangle + \frac{1}{2}\|z\|^2$

$\Leftrightarrow h = \varphi + \frac{1}{2}\|\cdot\|^2 + \langle \cdot | -z \rangle + \frac{1}{2}\|z\|^2$

$\therefore \boxed{h^* = {}^1(\varphi^*)(\cdot + z) - \frac{1}{2}\|z\|^2}$

/* now



$\Rightarrow (\underbrace{\tau_u f}_{f} + \langle \cdot | v \rangle + \alpha)^* = \tau_v f^* + \langle 0 | \cdot \rangle - \langle 0 | v \rangle - \alpha$

$= f^*(\cdot - v) - \alpha$ */

$\therefore h^* = (\varphi + \frac{1}{2}\|\cdot\|^2 + \langle \cdot | -z \rangle + \frac{1}{2}\|z\|^2)^*$

$= (\varphi + \frac{1}{2}\|\cdot\|^2)^*(\cdot + z) - \frac{1}{2}\|z\|^2)$ /*

*/

$= {}^1(\varphi^*)(\cdot + z) - \frac{1}{2}\|z\|^2$

also : $j(y) = \psi(y - r) = \tau_r \psi(y)$

$\Leftrightarrow j = \tau_r \psi$

$\therefore j^* = (\tau_r \psi)^* = (\tau_r \psi + \langle \cdot | 0 \rangle + 0)^*$

$= \underbrace{\tau_0 \psi^*}_{\psi^*(\cdot - 0) = \psi^*} + \langle \cdot | r \rangle - \langle 0 | r \rangle - 0$

$= \psi^* + \langle \cdot | r \rangle$

$\therefore \boxed{j^* = \psi^* + \langle \cdot | r \rangle}$

so, the problem $\left(\min_{v \in \mathcal{K}} h^*(L^*v) + j^*(-v)\right)$

$$= \min_{v \in \mathcal{K}} \left({}^1(\varphi^*)(\cdot + z) - \tfrac{1}{2}\|z\|^2\right)(L^*v) + (\psi^* + \langle \cdot | r \rangle)(-v)$$

$$= \left(\min_{v \in \mathcal{K}} {}^1(\varphi^*)(L^*v + z) - \underbrace{\tfrac{1}{2}\|z\|^2}_{\text{constant so can be dropped}} + \psi^*(-v) + \langle -v | r \rangle\right)$$

$$= \left(\min_{v \in \mathcal{K}} {}^1(\varphi^*)(L^*v + z) + \psi^*(-v) - \langle v | r \rangle\right)$$ // this coincides with [eq: 27.33]

and has atleast one solution

[eq: 27.35.2]

∴ [eq: 27.35], [eq: 27.35.1] and [eq: 27.35.2] ⇒

Primal problem :

$$\min_{x \in \mathcal{H}} \varphi(x) + \psi(Lx - r) + \frac{1}{2}\|x - z\|^2$$

Dual problem :

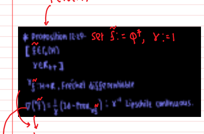
$$\left(\min_{v \in \mathcal{K}} \underbrace{{}^1(\varphi^*)(L^*v + z)}_{g(v) \text{ // say}} + \underbrace{\psi^*(-v) - \langle v | r \rangle}_{f(v) \text{ // say}}\right)$$ } has atleast one solution

: $g, f \in \Gamma_0(\mathcal{K})$

$$= \left(\min_{v \in \mathcal{K}} g(v) + f(v)\right) = \min(g + f)(\mathcal{K}) \longrightarrow \text{clearly Argmin}(f + g) \neq \emptyset$$

[eq:dual_problem]

now $\varphi^* \in \Gamma_0(\mathcal{H})$



Proposition 12.29 set $\tilde{f} := \varphi^*$, $\gamma := 1$

[$\tilde{f} \in \Gamma_0(\mathcal{H})$

$\gamma \in \mathbb{R}_{++}$]

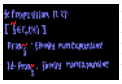
${}^\gamma \tilde{f} : \mathcal{H} \to \mathbb{R}$, Fréchet differentiable

$\nabla({}^\gamma \tilde{f}) = \frac{1}{\gamma}(\text{Id} - \text{Prox}_{\gamma \tilde{f}})$: γ^{-1} Lipschitz continuous.

- ${}^1(\varphi^*)$: Frechet differentiable ⇒ $g = {}^1(\varphi^*)(L^* \cdot + z)$. Frechet differentiable
- $\nabla\, {}^1(\varphi^*) = (\text{Id} - \text{Prox}_{\varphi^*})$: 1 Lipschitz continuous

$\Rightarrow \nabla\, {}^1(\varphi^*) = (\underbrace{\text{Id} - \text{Prox}_{\varphi^*}}_{\text{Prox}_\varphi})$: firmly nonexpansive

// using Moreau decomposition: // $\underset{\substack{\uparrow\\ \in \Gamma_0(\mathcal{H})}}{\text{Prox}_\varphi} + \text{Prox}_{\varphi^*} = \text{Id}$



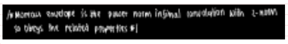
Proposition 12.27

[$\tilde{f} \in \Gamma_0(\mathcal{H})$]

$\text{Prox}_{\tilde{f}}$: firmly nonexpansive

$\text{Id} - \text{Prox}_{\tilde{f}}$: firmly nonexpansive

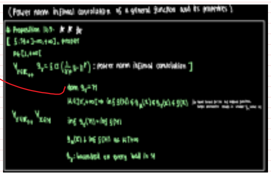
now ${}^1(\varphi^*)$ is 1-Moreau envelope



/* Moreau envelope is the power norm infimal convolution with 2-norm so obeys the related properties */

$\Rightarrow \text{dom}({}^1\varphi^*) = \mathcal{H}$

using



(Power norm infimal convolution of a general function and its properties)

Proposition 12.9 ★★★

[$f : \mathcal{H} \to]-\infty, +\infty]$, proper

$p \in [1, +\infty[$

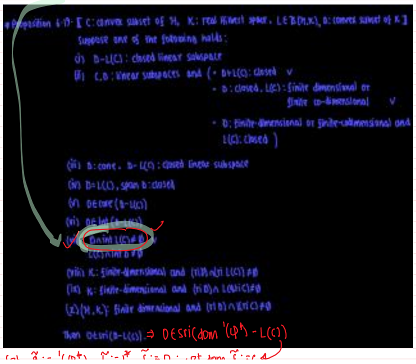
$\forall_{\gamma \in \mathbb{R}_{++}}$ $g_\gamma = f \,\square\, (\frac{1}{\gamma p}\|\cdot\|^p)$: power norm infimal convolution]

dom $g_\gamma = \mathcal{H}$

inf $g_\gamma(x)$ = inf $f(\mathcal{H})$

g_γ : bounded on every ball in $\mathcal{H}$

set : $D = \text{dom}\,{}^1(\varphi^*) = \mathcal{H} \Rightarrow \overset{\mathcal{H}}{\overset{\shortparallel}{D}} \cap \text{int}\, L(C) \neq \emptyset$



Proposition 6.19 [C: convex subset of $\mathcal{H}$, $\mathcal{K}$: real Hilbert space, $L \in \mathcal{B}(\mathcal{H},\mathcal{K})$, D: convex subset of $\mathcal{K}$]

Suppose one of the following holds:

(i) $D - L(C)$: closed linear subspace

(ii) C, D : linear subspaces and (• $D + L(C)$: closed ∨

• D : closed, $L(C)$: finite dimensional or finite co-dimensional ∨

• D : finite-dimensional or finite-codimensional and $L(C)$: closed)

(iii) D : cone, $D - L(C)$: closed linear subspace

(iv) $D = L(C)$, span D : closed

(v) $0 \in \text{core}(D - L(C))$

(vi) $0 \in \text{int}(D - L(C))$

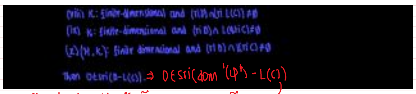
(vii) $D \cap \text{int}\, L(C) \neq \emptyset$ ∨ $L(C) \cap \text{int}\, D \neq \emptyset$

(viii) $\mathcal{K}$: finite-dimensional and $(\text{ri}\, D) \cap \text{ri}\, L(C) \neq \emptyset$

(ix) $\mathcal{K}$: finite-dimensional and $(\text{ri}\, D) \cap L(\text{ri}\, C) \neq \emptyset$

(x) $(\mathcal{H}, \mathcal{K})$: finite dimensional and $(\text{ri}\, D) \cap L(\text{ri}\, C) \neq \emptyset$

Then $0 \in \text{sri}(D - L(C))$. $\Rightarrow 0 \in \text{sri}(\text{dom}\,{}^1(\varphi^*) - L(C))$



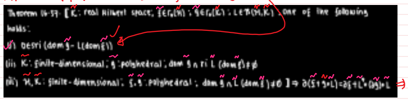
(viii) $\mathcal{K}$: finite-dimensional and $(\operatorname{ri} D) \cap (\operatorname{ri} L(C)) \neq \emptyset$

(ix) $\mathcal{K}$: finite-dimensional and $(\operatorname{ri} D) \cap L(\operatorname{ri} C) \neq \emptyset$

(x) $(\mathcal{H}, \mathcal{K})$: finite dimensional and $(\operatorname{ri} D) \cap L(\operatorname{ri} C) \neq \emptyset$

Then $0 \in \operatorname{sri}(D - L(C))$ $\Rightarrow 0 \in \operatorname{sri}(\operatorname{dom} {}^{1}(\varphi^{*}) - L(C))$

Set $\tilde{g} := {}^{1}(\varphi^{*})$, $\tilde{L} := L^{*}$, $\tilde{f} := 0$; set $\operatorname{dom} \tilde{f} := C$



Theorem 16.37. [$\tilde{\mathcal{K}}$: real Hilbert space; $\tilde{f} \in \Gamma_0(\mathcal{H})$; $\tilde{g} \in \Gamma_0(\tilde{\mathcal{K}})$; $L \in \mathcal{B}(\mathcal{H}, \tilde{\mathcal{K}})$; one of the following holds:

(i) $0 \in \operatorname{sri}(\operatorname{dom} \tilde{g} - L(\operatorname{dom} \tilde{f}))$

(ii) $\tilde{\mathcal{K}}$: finite-dimensional ; $\tilde{g}$: polyhedral ; $\operatorname{dom} \tilde{g} \cap \operatorname{ri} \tilde{L}(\operatorname{dom} \tilde{f}) \neq \emptyset$

(iii) $\tilde{\mathcal{H}}, \tilde{\mathcal{K}}$: finite-dimensional ; $\tilde{f}, \tilde{g}$: polyhedral ; $\operatorname{dom} \tilde{g} \cap \tilde{L}(\operatorname{dom} \tilde{f}) \neq \emptyset$] $\Rightarrow \partial(\tilde{f} + \tilde{g} \circ \tilde{L}) = \partial \tilde{f} + \tilde{L}^{*} \circ (\partial \tilde{g}) \circ \tilde{L}$

$\Rightarrow \nabla {}^{1}(\varphi^{*} \circ L) = \underbrace{L^{**}}_{\overset{\parallel}{L}} \circ \nabla {}^{1}(\varphi^{*}) \circ L$

$\Rightarrow \nabla {}^{1}(\varphi^{*}(Lv)) = L \nabla {}^{1}(\varphi^{*}) L^{*} v$

$\Rightarrow \nabla \underbrace{{}^{1}(\varphi^{*}(Lv + z))}_{g(v)} = L \nabla {}^{1}(\varphi^{*}) (L^{*} v + z)$

$$\boxed{\therefore \nabla g(v) = L \nabla {}^{1}(\varphi^{*}) (L^{*} v + z)}$$ [eq: 27.37]

L: linear operator

$\therefore \forall_{v \in \mathcal{K}} \forall_{w \in \mathcal{K}} \quad \| \nabla g(v) - \nabla g(w) \| = \| L \nabla {}^{1}(\varphi^{*}) (L^{*} v + z) - L \nabla {}^{1}(\varphi^{*}) (L^{*} w + z) \|$

$= \| L (\nabla {}^{1}(\varphi^{*}) (L^{*} v + z) - \nabla {}^{1}(\varphi^{*}) (L^{*} w + z)) \|$ /* use the def of linear-operator norm */

$\leq \| L \| \quad \underbrace{\| \nabla {}^{1}(\varphi^{*}) (L^{*} v + z) - \nabla {}^{1}(\varphi^{*}) (L^{*} w + z) \|}$ /* now $\nabla {}^{1}(\varphi^{*})$: 1 Lipschitz continuous */

/* $\leq \| L^{*} v + z - L^{*} w - z \| = \| L^{*} v - L^{*} w \|$ */

$\leq \| L \| \quad \underbrace{\| L^{*} v - L^{*} w \|}_{\parallel}$

/* $\| L^{*} (v - w) \| \leq \underbrace{\| L^{*} \|}_{\overset{\parallel}{\| L \|}} \| v - w \|$ */

// using Fact 2.18 $\| L \| = \| L^{*} \|$

$\leq \| L \|^{2} \| v - w \|$

$\therefore \forall_{v \in \mathcal{K}} \forall_{w \in \mathcal{K}} \quad \| \nabla g(v) - \nabla g(w) \| \leq \| L \|^{2} \| v - w \| \Rightarrow \nabla g : \| L \|^{2}$ Lipschitz continuous

$$\boxed{\therefore \nabla g(v) = L \underbrace{\nabla {}^{1}(\varphi^{*})}_{\operatorname{Prox}_{\varphi}} (L^{*} v + z) = L \operatorname{Prox}_{\varphi} (L^{*} v + z)}$$ [eq: 24.37.2]

[eq: 27.38]

[eq: 27.34], [eq: 27.37] and [eq: 27.38]

$x_n = \operatorname{Prox}_{\varphi} (L^{*} v_n + z) \Rightarrow L x_n = L \operatorname{Prox}_{\varphi} (L^{*} v_n + z) = \nabla g(v_n)$

$v_{n+1} = v_n - \lambda_n (\operatorname{Prox}_{\gamma \psi^{*}} (\gamma (L x_n - r) - v_n) + v_n)$ /* $\operatorname{Prox}_{\gamma \psi^{*}} (\gamma (\underbrace{L x_n}_{\nabla g(v_n)} - r) - v_n) = \operatorname{Prox}_{\gamma \psi^{*}} (\gamma (\nabla g(v_n) - r) - v_n) = \operatorname{Prox}_{\gamma \psi^{*}} (- (v_n - \gamma \nabla g(v_n) + \gamma r))$

$= - \operatorname{Prox}_{(\gamma \psi^{*})^{\vee}} (v_n - \gamma \nabla g(v_n) + \gamma r)$ /* using: */

$= - \operatorname{Prox}_{(\gamma \psi^{*})^{\vee}} ((v_n - \gamma \nabla g(v_n)) - (- \gamma r))$

$= - \operatorname{Prox}_{\underbrace{(\gamma \psi^{*})^{\vee} - \langle \cdot | \gamma r \rangle}} (v_n - \gamma \nabla g(v_n))$ /* using */

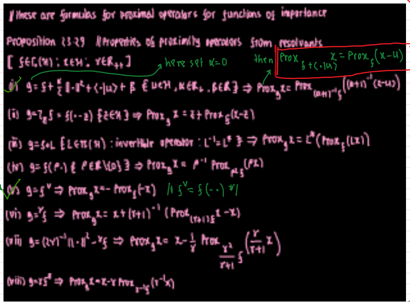
/* $\gamma (\psi^{*})^{\vee} - \gamma \langle \cdot | r \rangle$

$= \gamma (\underbrace{\psi^{* \vee} - \langle \cdot | r \rangle}_{f \text{ // by definition}}) = \gamma f$ */

$= - \operatorname{Prox}_{\gamma f} (v_n - \gamma \nabla g(v_n))$ */

$= v_n - \lambda_n (- \operatorname{Prox}_{\gamma f} (v_n - \gamma \nabla g(v_n)) - v_n)$ /* using */

$= v_n + \lambda_n (\operatorname{Prox}_{\gamma f} (v_n - \gamma \nabla g(v_n)) + v_n)$



// these are formulas for proximal operators for functions of importance

Proposition 23.29 // Properties of proximity operators from resolvents

[$f \in \Gamma_0(\mathcal{H})$; $x \in \mathcal{H}$; $\gamma \in \mathbb{R}_{++}$]

(i) $g = f + \frac{\alpha}{2} \| \cdot \|^{2} + \langle \cdot | u \rangle + \beta$ { $u \in \mathcal{H}$, $\alpha \in \mathbb{R}_{+}$, $\beta \in \mathbb{R}$ } $\Rightarrow \operatorname{Prox}_g x = \operatorname{Prox}_{(\alpha+1)^{-1} f} ((\alpha+1)^{-1} (x - u))$ here set $\alpha = 0$ then $\boxed{\operatorname{Prox}_{f + \langle \cdot | u \rangle} x = \operatorname{Prox}_f (x - u)}$

(ii) $g = \tau_z f = f(\cdot - z)$ { $z \in \mathcal{H}$ } $\Rightarrow \operatorname{Prox}_g x = z + \operatorname{Prox}_f (x - z)$

(iii) $g = f \circ L$ { $L \in \mathcal{B}(\mathcal{H})$: invertible operator : $L^{-1} = L^{*}$ } $\Rightarrow \operatorname{Prox}_g x = L^{*} (\operatorname{Prox}_f (L x))$

(iv) $g = f(\rho \cdot)$ { $\rho \in \mathbb{R} \setminus \{0\}$ } $\Rightarrow \operatorname{Prox}_g x = \rho^{-1} \operatorname{Prox}_{\rho^{2} f} (\rho x)$

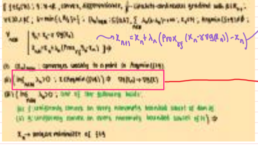
(v) $g = f^{\vee} \Rightarrow \operatorname{Prox}_g x = - \operatorname{Prox}_f (-x)$ /* $f^{\vee} = f(- \cdot)$ */

(vi) $g = {}^{\gamma} f \Rightarrow \operatorname{Prox}_g x = x + (\gamma + 1)^{-1} (\operatorname{Prox}_{(\gamma+1) f} x - x)$

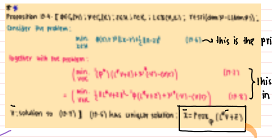
(vii) $g = (2\gamma)^{-1} \| \cdot \|^{2} - {}^{\gamma} f \Rightarrow \operatorname{Prox}_g x = x - \frac{1}{\gamma} \operatorname{Prox}_{\frac{\gamma^{2}}{\gamma+1} f} \left(\frac{\gamma}{\gamma+1} x \right)$

(viii) $g = \gamma f^{*} \Rightarrow \operatorname{Prox}_g x = x - \gamma \operatorname{Prox}_{\gamma^{-1} f} (\gamma^{-1} x)$

notice that this is the underlying recursion of forward-backward algorithm;



$x_{n+1} = x_n + \lambda_n (\operatorname{Prox}_{\gamma f} (x_n - \gamma \nabla g(x_n)) - x_n)$



this is the primal problem in our case

this is the dual problem in our case

$\boxed{\bar{x} = \operatorname{Prox}_{\varphi} (L^{*} \bar{v} + z)}$

(i) follows from :)

v_n: converges weakly to a point in $\operatorname{Argmin}(f + g)$ // $f : v \mapsto \psi^{*}(-v) - \langle v | r \rangle$ // $g : v \mapsto {}^{1}(\varphi^{*}) (L^{*} v + z)$ [eq:dual_problem] says solution of our dual problem is $\operatorname{Argmin}(f + g)$

denote it by $\bar{v}$

$\therefore \bar{x} = \operatorname{Prox}_{\varphi} (L^{*} \bar{v} + z)$:)

(ii) in (i) we have shown that $v_n \to \bar{v}$: $\bar{v} \in \operatorname{Argmin}(f + g)$ defined in [eq:dual_problem] (solution to dual problem)

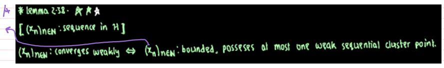
and $\bar{x} = \operatorname{Prox}_{\varphi} (L^{*} \bar{v} + z)$: unique solution to dual problem

set $\rho = \sup_{n \in \mathbb{N}} \| v_n - \bar{v} \|$

as $(v_n)_{n \in \mathbb{N}}$: bounded

$\rho < +\infty$

/*

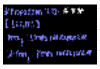


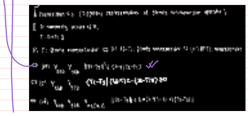
* lemma 2.38. ★★★

[$(x_n)_{n \in \mathbb{N}}$: sequence in $\mathcal{H}$]

$(x_n)_{n \in \mathbb{N}}$: converges weakly $\Leftrightarrow$ $(x_n)_{n \in \mathbb{N}}$: bounded, posseses at most one weak sequential cluster point. */

$\therefore \forall_{n \in \mathbb{N}} \quad \| \underset{\text{Prox}_{\varphi}(L^{*} v_n + z) \text{ // by definition}}{x_n} - \bar{x} \|^{2} = \| \operatorname{Prox}_{\varphi} (L^{*} v_n + z) - \operatorname{Prox}_{\varphi} (L^{*} \bar{v} + z) \|^{2}$ /*





*/

$\leq \langle L^{*} v_n + \cancel{z} - L^{*} \bar{v} - \cancel{z} | \operatorname{Prox}_{\varphi} (L^{*} v_n + z) - \operatorname{Prox}_{\varphi} (L^{*} \bar{v} + z) \rangle$

$= \langle L^{*} (v_n - \bar{v}) | \operatorname{Prox}_{\varphi} (L^{*} v_n + z) - \operatorname{Prox}_{\varphi} (L^{*} \bar{v} + z) \rangle$

/* linear operator */ /* now use : $\langle L^{*} \alpha | \beta \rangle = \langle \alpha | L \beta \rangle$ */

$= \langle v_n - \bar{v} | L (\operatorname{Prox}_{\varphi} (L^{*} v_n + z) - \operatorname{Prox}_{\varphi} (L^{*} \bar{v} + z)) \rangle$

$= \langle v_n - \bar{v} | \underbrace{L \operatorname{Prox}_{\varphi} (L^{*} v_n + z)}_{\nabla g(v_n)} - \underbrace{L \operatorname{Prox}_{\varphi} (L^{*} \bar{v} + z)}_{\nabla g(\bar{v})} \rangle$ /* from [eq: 24.37.2] . $\nabla g(v) = L \operatorname{Prox}_{\varphi} (L^{*} v + z)$ */

$= \langle v_n - \bar{v} | \nabla g(v_n) - \nabla g(\bar{v}) \rangle \leq \| v_n - \bar{v} \| \, \| \nabla g(v_n) - \nabla g(\bar{v}) \|$ /* Cauchy-Schwarz : $\langle x | y \rangle \leq \| x \| \, \| y \|$ */ $\forall_{n \in \mathbb{N}}$

$\leq \underbrace{\left(\sup_{n \in \mathbb{N}} \| v_n - \bar{v} \| \right)}_{\overset{\parallel}{\rho \text{ : finite, positive}}} \| \nabla g(v_n) - \nabla g(\bar{v}) \|$

$\therefore \forall_{n \in \mathbb{N}} \quad \| x_n - \bar{x} \|^{2} \leq \rho \underbrace{\| \nabla g(v_n) - \nabla g(\bar{v}) \|}_{\downarrow \; 0 \text{ as } n \to \infty}$

$\therefore \| x_n - \bar{x} \| \to 0$ as $n \to \infty$

$\Leftrightarrow \boxed{x_n \to x}$:)

■